Quantum Chemical Calculations of Drug-Like Molecules

Jung-Hsin Lin (林榮信)
Research Center for Applied Sciences &
Institute of Biomedical Sciences, Academia Sinica
School of Pharmacy, National Taiwan University
College of Engineering, Chang Gung University

http://www.rcas.sinica.edu.tw/faculty/jhlin.html

2018 Frontiers in Computational Drug Design, Academia Sinica, March 16-20, 2018

Rule of Five for Druglikeness

Lipinski et al, Adv. Drug. Deliv. Rev. **23:** 3-25 (1997)

In order to achieve good solubility and permeability for a novel drug:

- There should be less than 5 H-bond donors.
- There should be less than 10 H-bond acceptors.
- The molecular weight should be smaller than 500.
- The calculated Log P (CLogP) should be smaller than 5.

ELSEVIER

Advanced Drug Delivery Reviews 23 (1997) 3–25

advanced
drug delivery
reviews

Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings

Christopher A. Lipinski*, Franco Lombardo, Beryl W. Dominy, Paul J. Feeney

Central Research Division, Pfizer Inc., Groton, CT 06340, USA

Received 9 August 1996; accepted 14 August 1996

Abstract

Experimental and computational approaches to estimate solubility and permeability in discovery and development settings are described. In the discovery setting 'the rule of 5' predicts that poor absorption or permeation is more likely when there are more than 5 H-bond donors, 10 H-bond acceptors, the molecular weight (MWT) is greater than 500 and the calculated Log P (CLogP) is greater than 5 (or MlogP > 4.15). Computational methodology for the rule-based Moriguchi Log P (MLogP) calculation is described. Turbidimetric solubility measurement is described and applied to known drugs. High throughput screening (HTS) leads tend to have higher MWT and Log P and lower turbidimetric solubility than leads in the pre-HTS era. In the development setting, solubility calculations focus on exact value prediction and are difficult because of polymorphism. Recent work on linear free energy relationships and Log P approaches are critically reviewed. Useful predictions are possible in closely related analog series when coupled with experimental thermodynamic solubility measurements.

Keywords: Rule of 5; Computational alert; Poor absorption or permeation; MWT; MLogP; H-Bond donors and acceptors; Turbidimetric solubility; Thermodynamic solubility; Solubility calculation

like compounds. This review deals only with solubility and permeability as barriers to absorption. Intestinal wall active transporters and intestinal wall metabolic events that influence the measurement of drug bioavailability are beyond the scope of this review. We hope to spark lively debate with our

in the chemical structure was larger than 5. The 'rule of 5' states that: poor absorption or permeation are more likely when:

There are more than 5 H-bond donors (expressed as the sum of OHs and NHs);
The MWT is over 500;
The Log P is over 5 (or MLogP is over 4.15);
There are more than 10 H-bond acceptors (expressed as the sum of Ns and Os)
Compound classes that are substrates for biological transporters are exceptions to the rule.

clinical stages and Phase I safety evaluation. We expected that the most insoluble and poorly permeable compounds would have been eliminated in those compounds that survived to enter Phase II efficacy studies. We could use the presence of United States Adopted Name (USAN) or International Non-proprietary Name (INN) names to identify compounds entering Phase II since most drug companies (including Pfizer) apply for these names at entry to Phase II.

When we examined combinations of any two of the four parameters in the USAN data set, we found that combinations of two parameters outside the desirable range did not exceed 10%. The exact values from the USAN set are: sum of N and O + sum of NH and OH — 10%; sum of N and O + MWT — 7%; sum of NH and OH + MWT — 4% and sum of MWT + Log P — 1%. The rarity (1%) among USAN drugs of the combination of high MWT and high log P was striking because this particular combination of physico-chemical proper-

compounds. In our experience, most compounds failing the alert also will prove troublesome if they progress far enough to be studied experimentally. However, the converse is not true. Compounds passing the alert still can prove troublesome in experimental studies.

Black Body Radiation

A type of electromagnetic radiation with or surrounding a body in thermodynamic equilibrium with its environment, or emitted by a black body (an opague and non-radiating reflective body) held at constant, uniform temperature.

Rayleigh-Jeans equation:

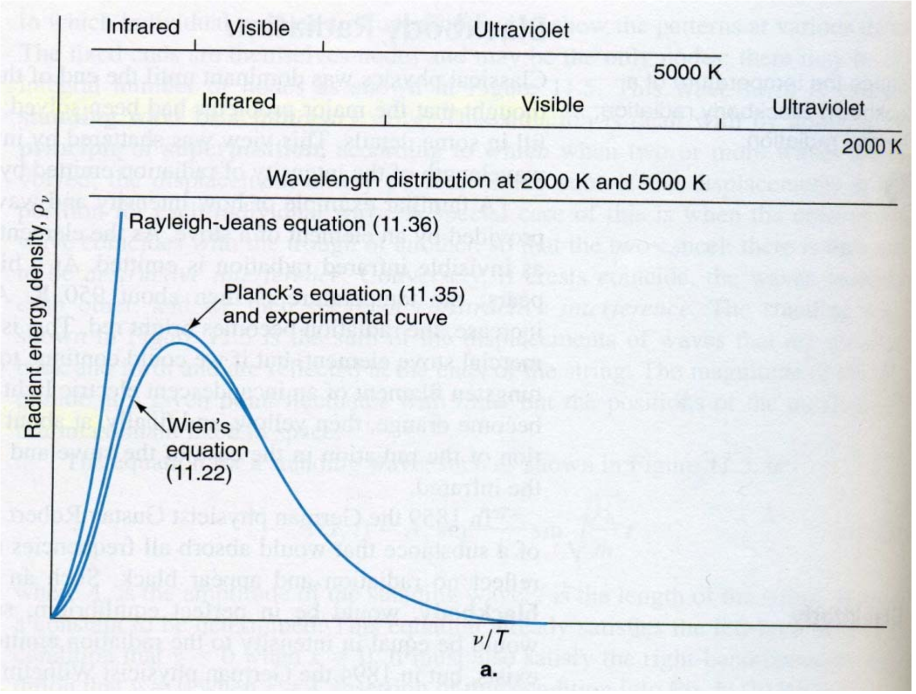
$$\frac{2\nu^2 k_{\mathrm{B}} T}{c^2}.$$

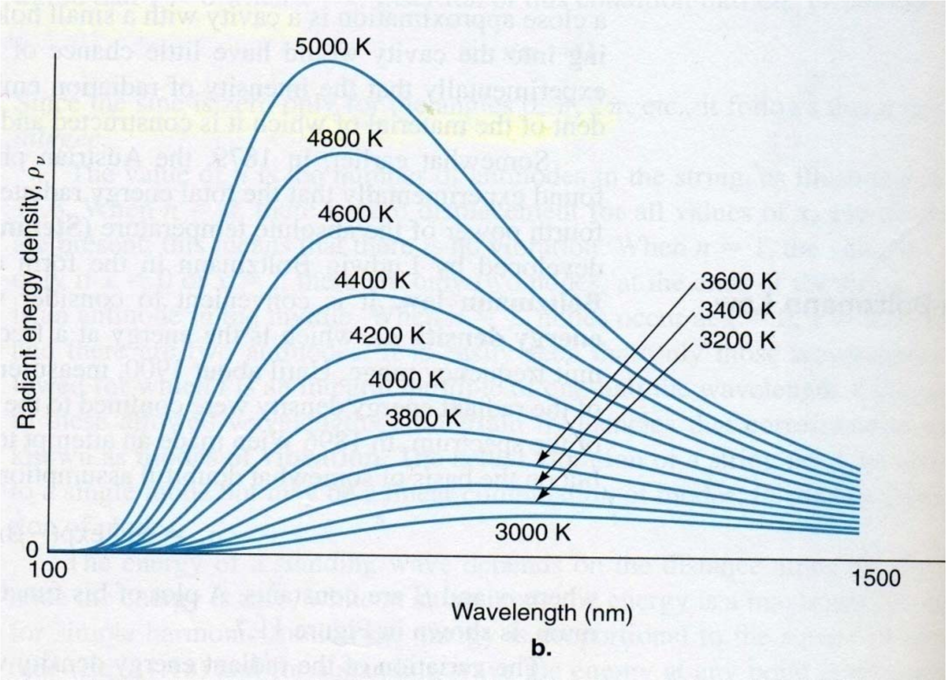
Wien's equation:

$$\nu_{\max} = \frac{\alpha}{h} kT \approx (5.879 \times 10^{10}\ \mathrm{Hz/K}) \cdot T$$

Planck's equation:

$$\frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{kT}} - 1}$$





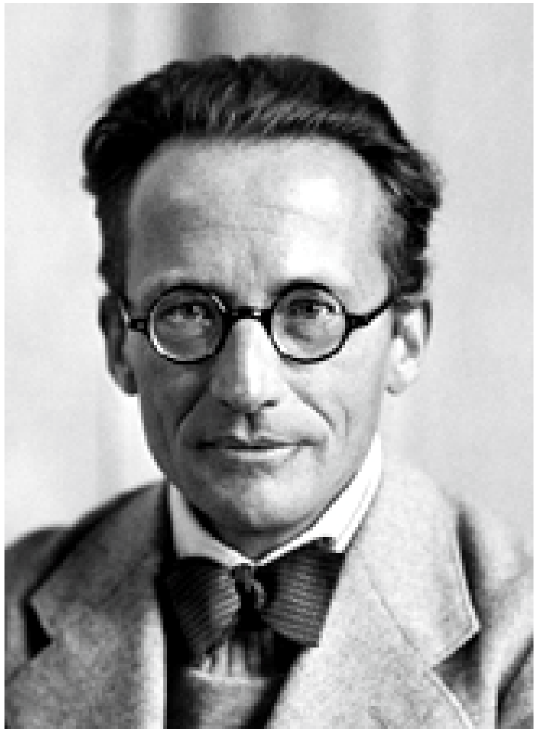
Wave-particle duality theory of matter (1924)

His 1924 **doctoral thesis**, *Recherches sur la théorie des quanta* (Research on Quantum Theory), introduced his theory of electron waves. This included the wave-particle duality theory of matter, based on the work of Max Planck and Albert Einstein on light. The thesis examiners, unsure of the material, passed his thesis to Einstein for evaluation who endorsed his wave-particle duality proposal wholeheartedly; de Broglie was awarded his doctorate. This research culminated in the de Broglie hypothesis stating that ***any moving particle or object had an associated wave***. De Broglie thus created a new field in physics, the *mécanique ondulatoire,* or **wave mechanics**, uniting the physics of energy (wave) and matter (particle). For this he won the **Nobel Prize in Physics in 1929.**

$$E = h\nu \qquad p = \frac{h}{\lambda}$$

Schrödinger's wave equation (1926)

$$\frac{\partial^2\Psi(x,t)}{\partial x^2}=\frac{1}{c^2}\frac{\partial^2\Psi(x,t)}{\partial t^2}$$



$$\Psi(x,t)=Ce^{\frac{2\pi i x}{\lambda}-2\pi i \nu t}$$

$$\Psi(x,t)=Ce^{\frac{2\pi i x p_x}{h}-\frac{2\pi i E t}{h}}$$

$$\frac{\partial\Psi(x,t)}{\partial t}=-\frac{iEC}{\hbar}e^{\frac{i x p_x}{\hbar}}e^{-\frac{iEt}{\hbar}}=-\frac{iE}{\hbar}\Psi(x,t)$$

Time-dependent Schrödinger Equation

$$\left\{-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V\right\}\Psi(\mathbf{r},t)=i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t}$$

$$\hbar=\frac{h}{2\pi}$$

h : Planck constant

m : mass

V : potential energy

Time-independent Schrödinger Equation

$$\left\{-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V\right\}\Psi(\mathbf{r},t)=i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t}$$

When the external potential V is independent of time

$$\Psi(\mathbf{r},t)=\psi(\mathbf{r})T(t)$$

$$\left\{-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{r})\right\}\psi(\mathbf{r})=E\psi(\mathbf{r})$$

$$\mathcal{H}\psi(\mathbf{r})=E\psi(\mathbf{r})$$

$$\mathcal{H}=-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{r})$$

Hamiltonian Operator

$$\mathcal{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})$$

$$\mathcal{H}\psi(\mathbf{r}) = E\psi(\mathbf{r})$$

$$E = \frac{\int \psi^* \mathcal{H} \psi \, d\tau}{\int \psi^* \psi \, d\tau}$$

Normalization of wave functions

$$\int \psi^* \psi \, d\tau = 1$$

total probability should be 1

Orthonormal condition: orthogonal to each other and normalized

$$\int \psi_m{}^* \psi_n \, d\tau = \begin{cases} 1 & , m = n \\ 0 & , m \neq n \end{cases}$$

$$= \delta_{mn}$$

Kronecker delta

One-electron atoms

$$\mathcal{H} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\varepsilon_0 r}$$

Z=+1: hydrogen atom H

Z=+2: helium cation He$^+$

Separation of variables in spherical polar coordinates:

$$\psi_{nlm} = R_{nl}(r)Y_{lm}(\theta,\phi)$$

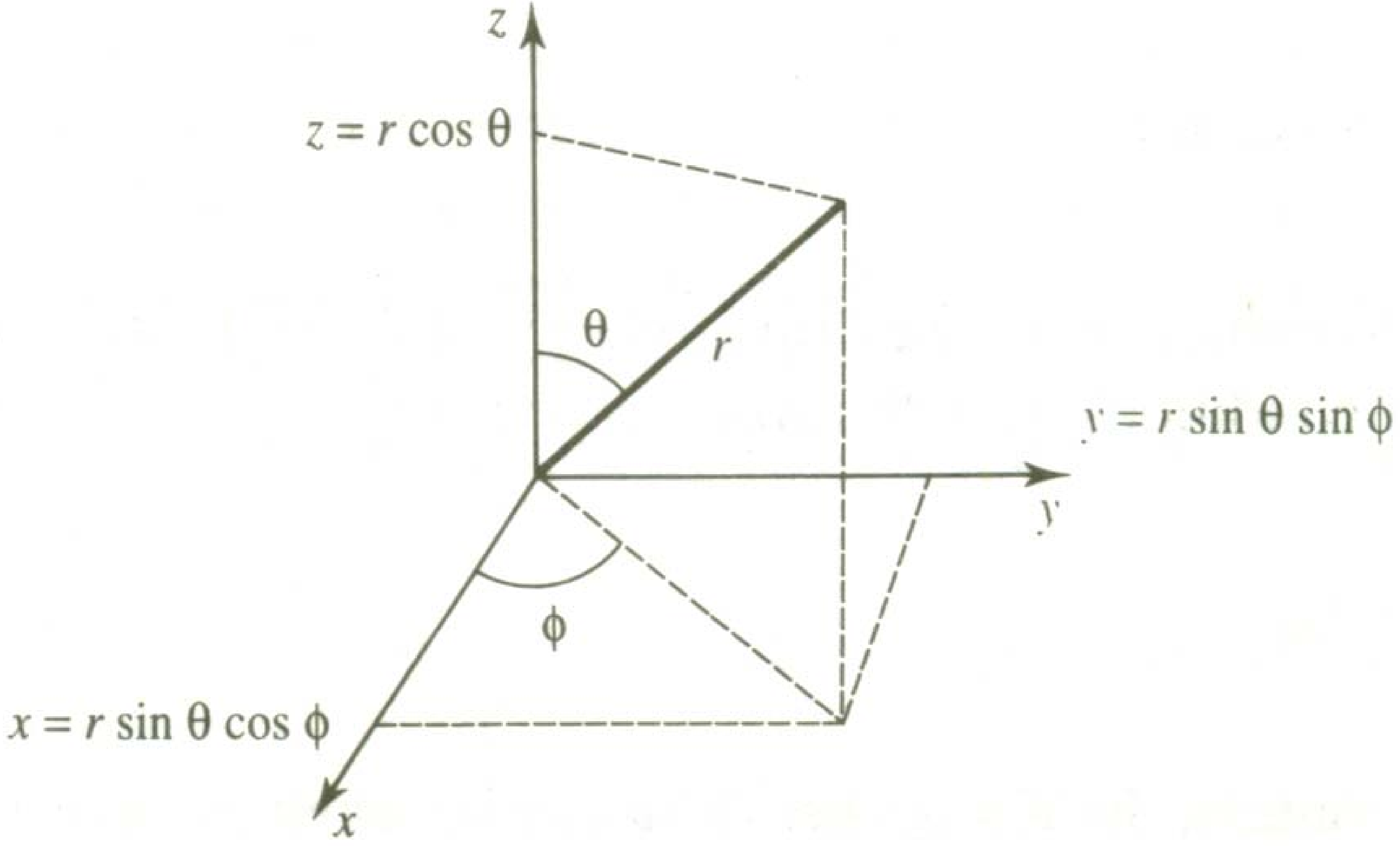
also called **orbitals**

n: principal quantum number: *1, 2, …*

l: azimuthal quantum number: *0, 1, … (n-1)*

m: magnetic quantum number: *-l, -(l-1), …,0,…(l-1), l.*

Spherical Polar Coordinate



Radial Function

$$R_{nl}(r) = -\sqrt{\left(\frac{2Z}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}}\, e^{-\frac{\rho}{2}} \rho^l L_{n+1}^{2l+1}(\rho)$$

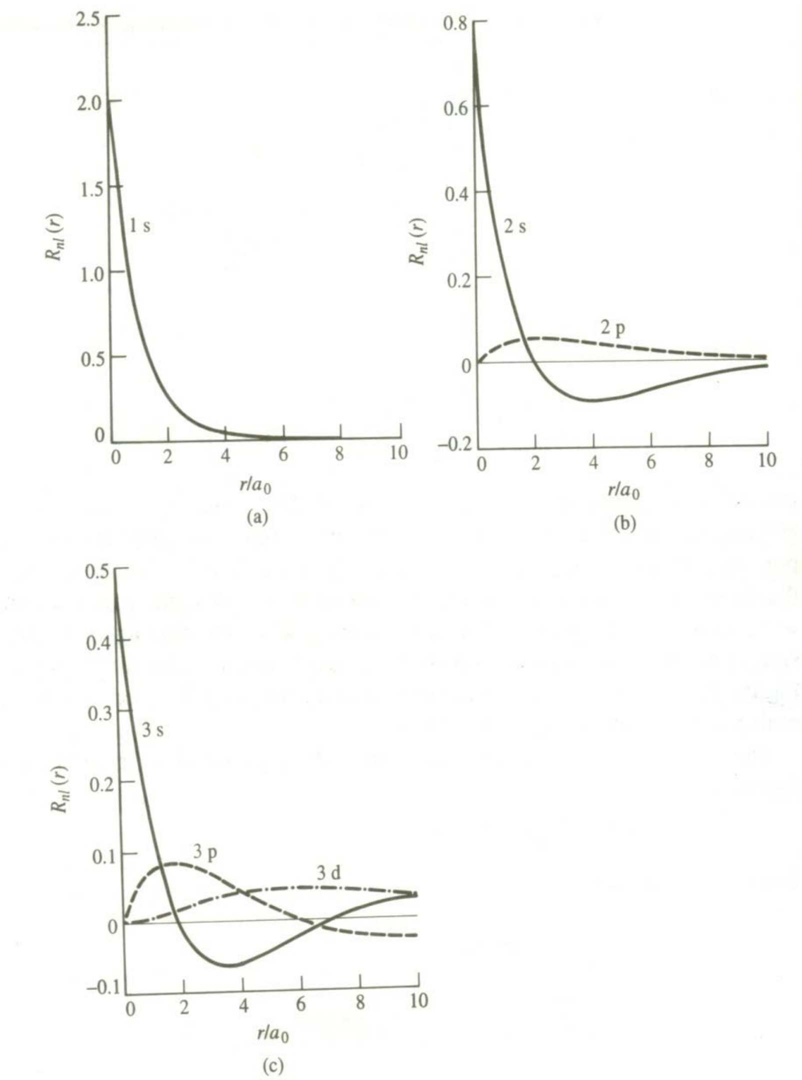
$$\rho = \frac{2Zr}{na_0}$$

a_0 : Bohr radius

$L_{n+1}^{2l+1}(\rho)$: Laguerre polynomial

Examples of radial functions

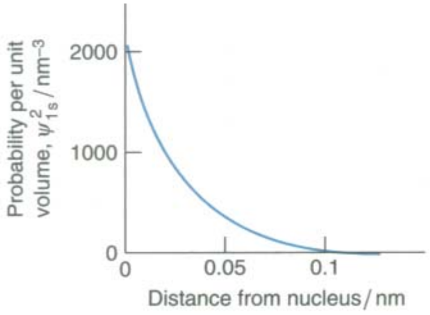
n	l	$R_{nl}(r)$
1	0	$2\zeta^{3/2}e^{-\zeta r}$
2	0	$2\zeta^{3/2}(1-\zeta r)e^{-\zeta r}$
2	1	$\sqrt{\frac{4}{3}}\zeta^{5/2}re^{-\zeta r}$
3	0	$\sqrt{\frac{2}{3}}\zeta^{3/2}(3-6\zeta r+2\zeta^2 r^2)e^{-\zeta r}$
3	1	$\sqrt{\frac{8}{9}}\zeta^{5/2}(2-\zeta r)re^{-\zeta r}$
3	2	$\sqrt{\frac{8}{45}}\zeta^{7/2}r^2e^{-\zeta r}$

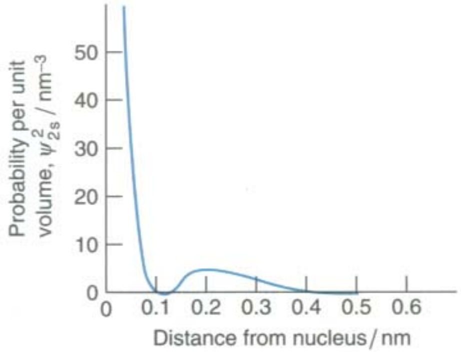


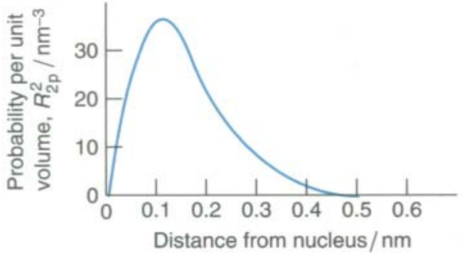
Polynomial multiplied by a decaying exponential!

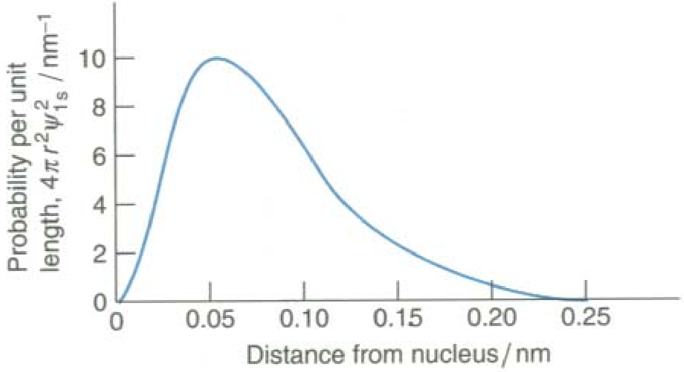
$$\zeta = \frac{Z}{n}$$

Radial Probability Per Unit Length

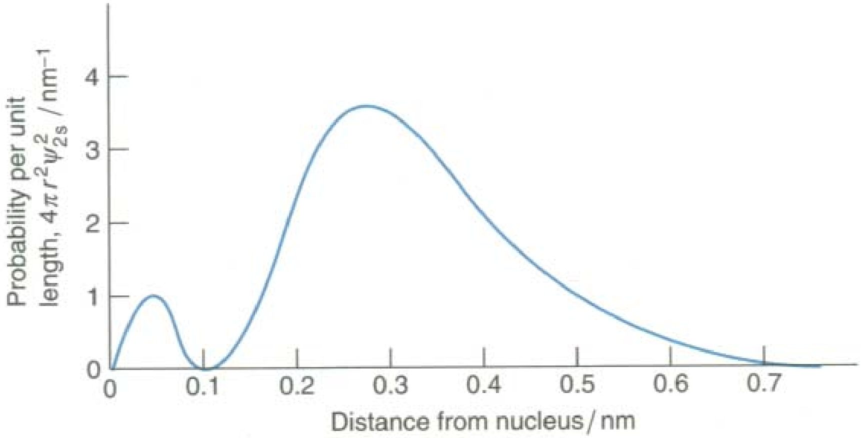




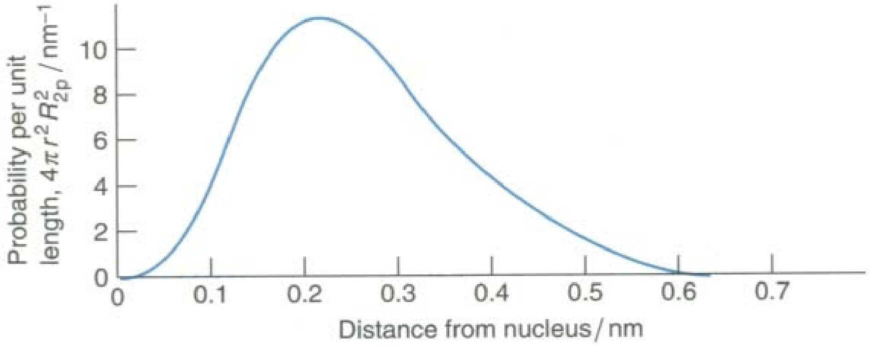




a.



b.



c.

Spherical Harmonics

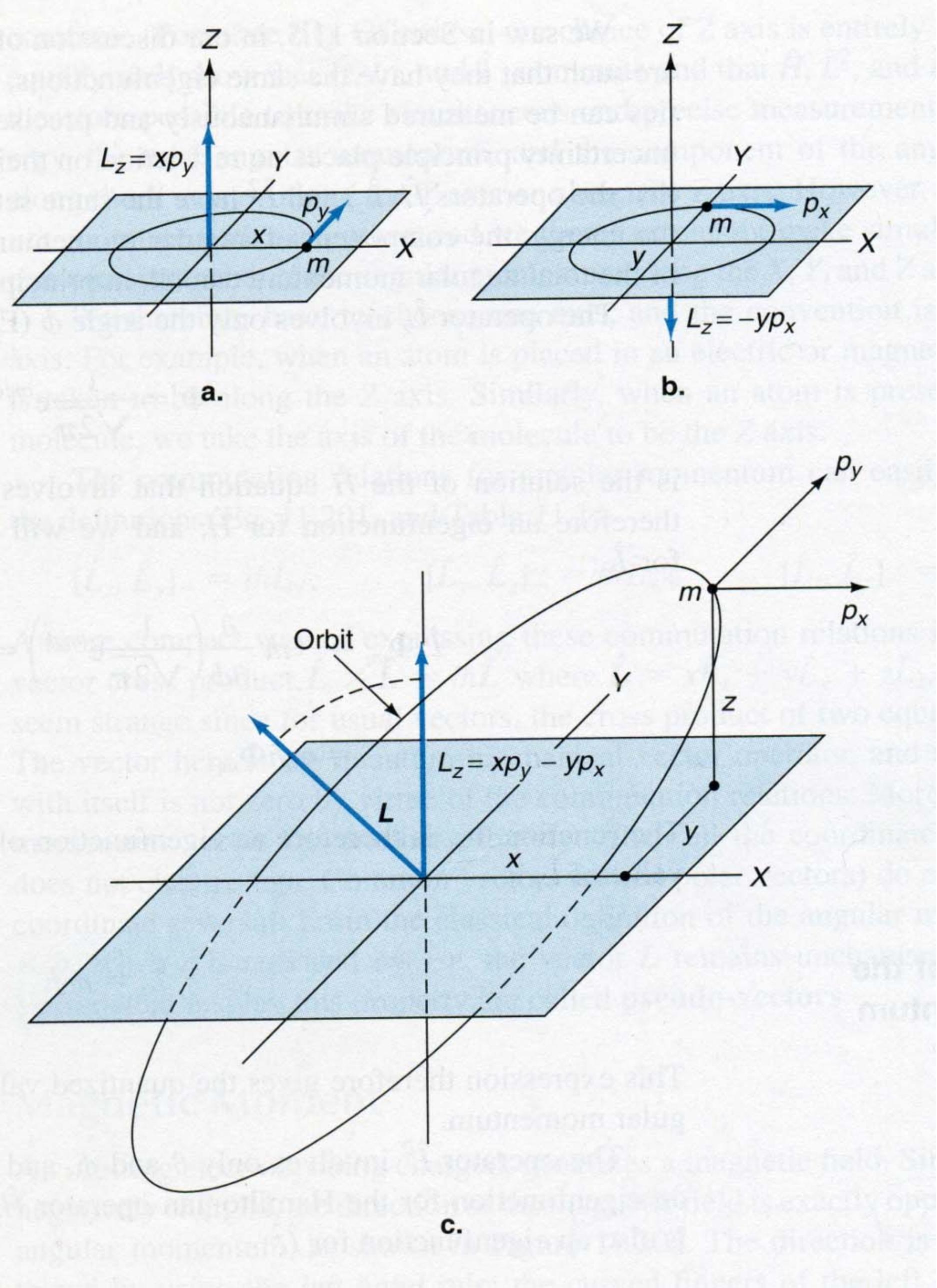
$$Y_{lm}(\theta,\phi) = \Theta_{lm}(\theta)\Phi_m(\phi)$$

$$\Theta_{lm}(\theta) = \sqrt{\frac{(2l+1)}{2}\frac{(l-|m|)!}{(l+|m|)!}}P_l^{|m|}(\cos\theta)$$

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}}e^{im\phi}$$

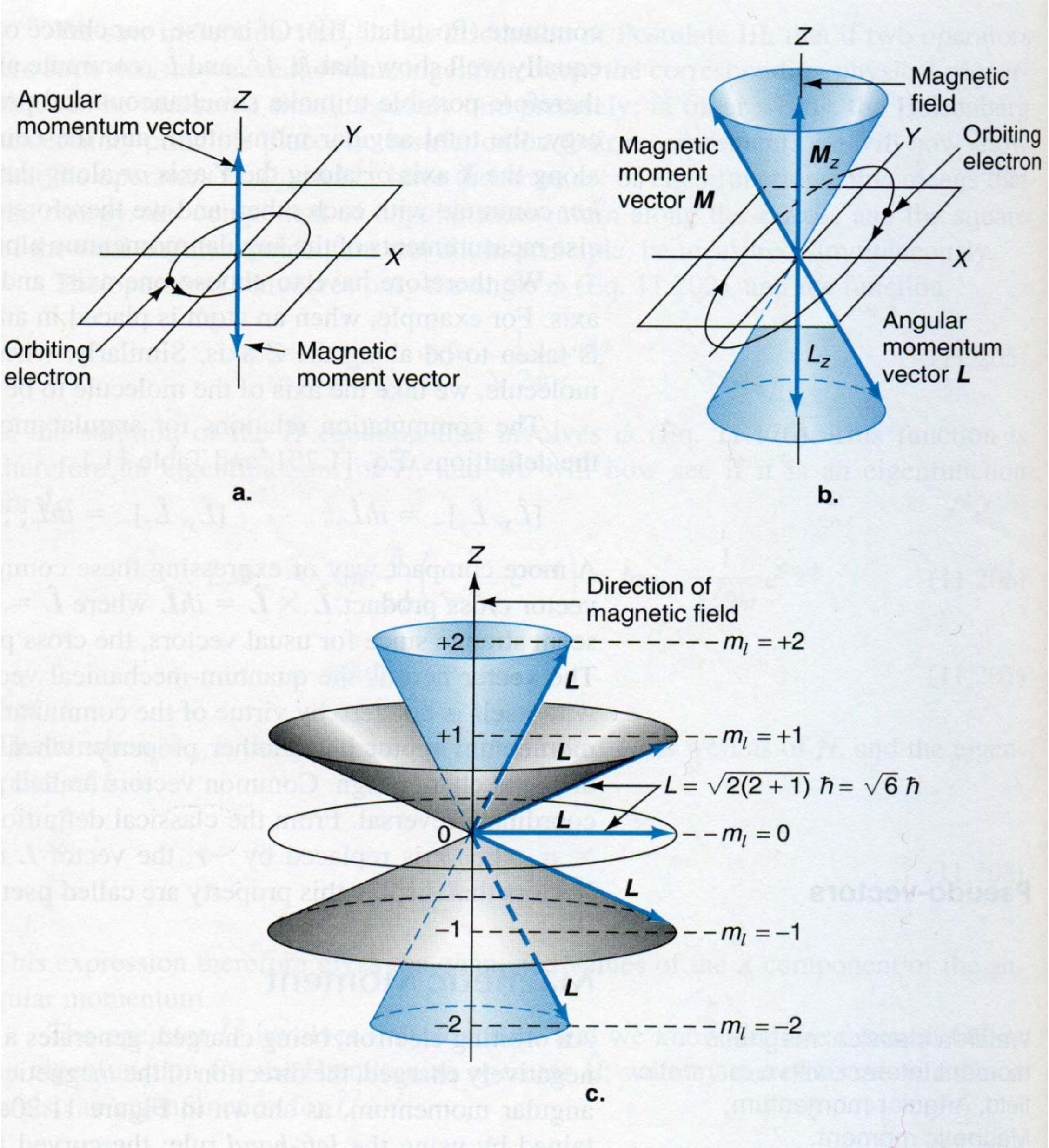
$P_l^{|m|}(\cos\theta)$: associated Legendre polynomial

Classical Angular Momentum



$$\vec{L} = \vec{r} \times \vec{p}$$

Quantum Angular Momentum



Total orbital angular momentum:

$$L = l(l+1)\hbar$$

Z component of total orbital angular momentum:

$$L_z = m\hbar$$

TABLE 11.3 Solution of the Θ Equation

l	m_l	Function
0	0	$\Theta_{00} = \dfrac{\sqrt{2}}{2} = \dfrac{1}{\sqrt{2}}$
1	0	$\Theta_{10} = \dfrac{\sqrt{6}}{2} \cos\theta$
1	+1, −1	$\Theta_{1\pm1} = \dfrac{\sqrt{3}}{2} \sin\theta$
2	0	$\Theta_{20} = \dfrac{\sqrt{10}}{4}(3\cos^2\theta - 1)$
2	+1, −1	$\Theta_{2\pm1} = \dfrac{\sqrt{15}}{2} \sin\theta\cos\theta$
2	+2, −2	$\Theta_{2\pm2} = \dfrac{\sqrt{15}}{4} \sin^2\theta$

TABLE 11.2 Solution of the Φ Equation

Value of m_l	Solution in Complex Form	Real Form
0	—	$\Phi_0 = \frac{1}{\sqrt{2\pi}}$
1	$\Phi_1 = \frac{1}{\sqrt{2\pi}} e^{i\phi}$	$\Phi_x = \frac{1}{\sqrt{2}}(\Phi_1 + \Phi_{-1}) = \frac{\cos\phi}{\sqrt{\pi}}$
-1	$\Phi_{-1} = \frac{1}{\sqrt{2\pi}} e^{-i\phi}$	$\Phi_y = \frac{1}{\sqrt{2}}(\Phi_1 - \Phi_{-1}) = \frac{\sin\phi}{\sqrt{\pi}}$
2	$\Phi_2 = \frac{1}{\sqrt{2\pi}} e^{i2\phi}$	$\Phi_{x^2-y^2} = \frac{1}{\sqrt{2}}(\Phi_2 + \Phi_{-2}) = \frac{\cos 2\phi}{\sqrt{\pi}}$
-2	$\Phi_{-2} = \frac{1}{\sqrt{2\pi}} e^{-i2\phi}$	$\Phi_{xy} = \frac{1}{\sqrt{2}}(\Phi_2 - \Phi_{-2}) = \frac{\sin 2\phi}{\sqrt{\pi}}$

The reasons for the $x^2 - y^2$ and xy subscripts are as follows:

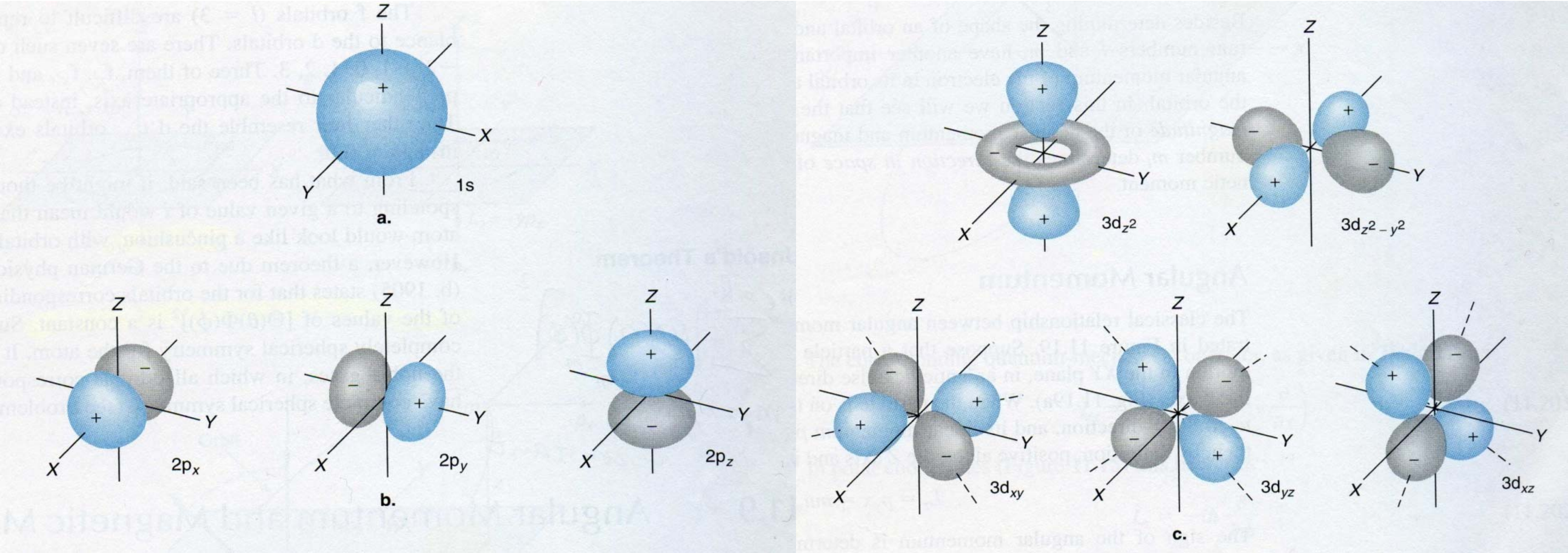
1. $\cos 2\phi = \cos^2\phi - \sin^2\phi$; from Figure 11.15, $\cos^2\phi = x^2/(r^2 \sin^2\theta)$ and $\sin^2\phi = y^2/(r^2\sin^2\theta)$; thus $\cos 2\phi$ has the same dependence on ϕ as $x^2 - y^2$.
2. $\sin 2\phi = 2\sin\phi\cos\phi = xy/(r^2\sin^2\theta)$; thus $\sin 2\phi$ has the same dependence on ϕ as xy.

TABLE 11.5 Selected Complete Hydrogen Atom Wave Functions $\psi_{n,l,m}$ for the Hydrogen Atom ($Z = 1$)

Quantum Numbers			
n	l	m_l	Function
1	0	0	$\psi_{1\mathrm{s}} = \psi_{100} = \dfrac{1}{\sqrt{\pi}}\left(\dfrac{1}{a_0}\right)^{3/2} \mathrm{e}^{-r/a_0}$
2	0	0	$\psi_{2\mathrm{s}} = \psi_{200} = \dfrac{1}{4\sqrt{2\pi}}\left(\dfrac{1}{a_0}\right)^{3/2}\left(2 - \dfrac{r}{a_0}\right) \mathrm{e}^{-r/2a_0}$
2	1	0	$\psi_{2\mathrm{p}_z} = \psi_{210} = \dfrac{1}{4\sqrt{2\pi}}\left(\dfrac{1}{a_0}\right)^{3/2} \dfrac{r}{a_0}\mathrm{e}^{-r/2a_0} \cos\theta$
2	1	+1	$\psi_{2\mathrm{p}_x} = \dfrac{1}{4\sqrt{2\pi}}\left(\dfrac{1}{a_0}\right)^{3/2} \dfrac{r}{a_0}\mathrm{e}^{-r/2a_0} \sin\theta \cos\phi$
2	1	−1	$\psi_{2\mathrm{p}_y} = \dfrac{1}{4\sqrt{2\pi}}\left(\dfrac{1}{a_0}\right)^{3/2} \dfrac{r}{a_0}\mathrm{e}^{-r/2a_0} \sin\theta \sin\phi$
3	0	0	$\psi_{3\mathrm{s}} = \psi_{300} = \dfrac{1}{18\sqrt{3\pi}}\left(\dfrac{1}{a_0}\right)^{3/2}\left[6 - 6\left(\dfrac{2r}{3a_0}\right) + \left(\dfrac{2r}{3a_0}\right)^2\right] \mathrm{e}^{-r/3a_0}$

The significance of the notation 1s, 2s, $2\mathrm{p}_z$, etc., is considered later.

Visualization of Some Typical Orbitals



Moving from atoms to molecules

- How to reasonable simply the Hamiltonian of molecular systems?
- How to express the molecular wave functions?
- How to solve the corresponding mathematical equations effectively?

Born-Oppenheimer Approximation

Electrons can adjust almost instantaneously to any changes in the positions of the nuclei and not their momentum:

$$\psi_{total}(nuclei, electrons) = \psi(electrons)\psi(nuclei)$$

The total energy is the sum of the nuclear energy and the electronic energy:

$$E_{total}(nuclei, electrons) = E(electrons) + E(nuclei)$$

The energy of a general many-electron many-nucleus system

$$\mathcal{H} = -\frac{1}{2}\sum_{\eta}\nabla_{\eta}^{2} - \frac{1}{r_{1A}} - \frac{1}{r_{1B}} - \frac{1}{r_{1C}} \cdots - \frac{1}{r_{2A}} - \frac{1}{r_{2B}} - \frac{1}{r_{2C}} \cdots$$

$$+ \frac{1}{r_{12}} + \frac{1}{r_{13}} + \frac{1}{r_{14}} \ldots \frac{1}{r_{23}} + \frac{1}{r_{24}} + \frac{1}{r_{25}} \cdots$$

$\mathrm{A, B, C} \cdots$ nuclei

$1,2,3 \cdots$ electron

No Electron-Electron Repulsion and Hartree-Product Wavefunctions

If only one-electron kinetic energy and nuclear attraction are considered:

$$\mathcal{H} = \sum_{i=1}^{N} h_i$$

$$h_i = -\frac{\hbar^2}{2m}\nabla_i^{\,2} - \sum_{k=1}^{M}\frac{Z_k}{r_{ik}}$$

$$h_i\psi_i = \varepsilon_i\psi_i$$

The Hartree Product

$$\Psi_{HP} = \psi_1\psi_2\ldots\psi_N$$

$$\mathcal{H}\,\Psi_{HP} = \mathcal{H}\psi_1\psi_2\ldots\psi_N = \left(\sum_{i=1}^{N}\varepsilon_i\right)\Psi_{HP}$$

The Hartree-Fock Approximation

- Each electron is moving in an averaged field (mean-field) caused by other electrons.
- Electrons can therefore be considered independent to each other in motion.
- The effective interaction in the mean field approximation will be calculated using a self-consistent field (SCF) scheme, i.e., an iterative procedure until the convergence of some quantity is achieved.
- Electron-electron correlations are thus **ignored**.

The Hartree Hamiltonian

If the electron-electron repulsion is considered in an average way:

$$\mathcal{H} = \sum_{i=1}^{N} h_i$$

$$h_i = -\frac{\hbar^2}{2m}\nabla_i^{\,2} - \sum_{k=1}^{M}\frac{Z_k}{r_{ik}} + \sum_{j\neq i}^{N}\int\frac{\rho_j}{r_{ij}}d\mathbf{r}$$

$$= -\frac{\hbar^2}{2m}\nabla_i^{\,2} - \sum_{k=1}^{M}\frac{Z_k}{r_{ik}} + V_i\{j\}$$

$$h_i\psi_i = \varepsilon_i\psi_i$$

$$E = \sum_{i=1}^{N}\varepsilon_i - \frac{1}{2}\sum_{i\neq j}\boxed{\iint\frac{\left|\psi_i\right|^2\left|\psi_j\right|^2}{r_{ij}}d\mathbf{r}_i d\mathbf{r}_j}$$

The Coulomb integral

Molecular Orbitals (MO) Expressed as Linear Combination of Atomic Orbitals (LCAO)

Each molecular orbital is expressed as a linear combination of atomic orbitals:

$$\psi_i = \sum_{\mu=1}^{K} c_{\mu i}\phi_\mu$$

ψ_i : spatial molecular orbital

ϕ_μ : one of K atomic orbitals

$c_{\mu i}$: coefficient

Spin Orbital

$$\chi_k(n) = \psi_i(n)s(n)$$

$\chi_k(n)$: a spin orbital of electron n

$\psi_i(n)$: spatial orbital of electron n, e.g., 1s, 2s, etc.

$s(n)$: spin state of electron n; α (up) or β(down)

Slater Determinant is a representation of wave function that satisfies Pauli exclusion principle automatically

$$\psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_i(1) & \chi_j(1) & \cdots & \chi_k(1) \\ \chi_i(2) & \chi_j(2) & \cdots & \chi_k(2) \\ \vdots & \vdots & & \vdots \\ \chi_i(N) & \chi_j(N) & \cdots & \chi_k(N) \end{vmatrix}$$

$$\equiv \left| \chi_i(1) \quad \chi_j(2) \quad \cdots \quad \chi_k(N) \right\rangle$$

Noted that this is still a determinant! Just a short-hand notation.

$\chi_i(1)$: a spin orbital filled with the electron labeled '1'.

The energy of a general polyelectronic system

$$\mathcal{H} = -\frac{1}{2}\sum_{\eta}\nabla_{\eta}^{2} - \frac{1}{r_{1A}} - \frac{1}{r_{1B}} - \frac{1}{r_{1C}}\cdots - \frac{1}{r_{2A}} - \frac{1}{r_{2B}} - \frac{1}{r_{2C}}\cdots$$
$$+\frac{1}{r_{12}} + \frac{1}{r_{13}} + \frac{1}{r_{14}}\cdots\frac{1}{r_{23}} + \frac{1}{r_{24}} + \frac{1}{r_{25}}\cdots$$

$A, B, C\cdots$	nuclei
$1,2,3\cdots$	electron

$$\int \psi^{*}\mathcal{H}\psi = \int\cdots\int d\tau_{1}d\tau_{2}\cdots d\tau_{N}$$
$$\left\{\left\langle \chi_{i}(1)\chi_{j}(2)\chi_{k}(3)\cdots\right| \left[-\frac{1}{2}\sum_{\eta}\nabla_{\eta}^{2} - \frac{1}{r_{1A}} - \frac{1}{r_{1B}} - \frac{1}{r_{1C}}\cdots - \frac{1}{r_{2A}} - \frac{1}{r_{2B}} - \frac{1}{r_{2C}}\cdots + \frac{1}{r_{12}} + \frac{1}{r_{13}} + \frac{1}{r_{14}}\cdots\frac{1}{r_{23}} + \frac{1}{r_{24}} + \frac{1}{r_{25}}\cdots\right] \left|\chi_{i}(1)\chi_{j}(2)\chi_{k}(3)\cdots\right\rangle\right\}$$

Complex conjugate of the Slater determinant

Slater determinant

The numerator in the energy expression can be broken down into one-electron and two-electron integrals

$$E = \frac{\int \psi^* \mathcal{H} \psi \, d\tau}{\int \psi^* \psi \, d\tau}$$

$$\int \psi^* \psi \, d\tau = N!$$

$$\int \psi^* \mathcal{H} \psi = \int \cdots \int d\tau_1 d\tau_2 \cdots d\tau_N$$

$$\left\{ \left\langle \chi_i(1)\chi_j(2)\chi_k(3)\cdots \right| \left[\begin{array}{l} -\frac{1}{2}\sum_{\eta} \nabla_{\eta}^2 - \frac{1}{r_{1A}} - \frac{1}{r_{1B}} - \frac{1}{r_{1C}} \cdots - \frac{1}{r_{2A}} - \frac{1}{r_{2B}} - \frac{1}{r_{2C}} \cdots \\ \qquad + \frac{1}{r_{12}} + \frac{1}{r_{13}} + \frac{1}{r_{14}} \cdots \frac{1}{r_{23}} + \frac{1}{r_{24}} + \frac{1}{r_{25}} \cdots \end{array} \right] \left| \chi_i(1)\chi_j(2)\chi_k(3)\cdots \right\rangle \right\}$$

$$= \sum_i \int d\tau_i \chi_i^* O_i \chi_i + \sum_{i,j} \iint d\tau_i d\tau_j \chi_i^* \chi_j^* O_{ij} \chi_i \chi_j$$

Splitting the Total Hamiltonian into Core Hamiltonian, Coulomb integral, and Exchange integral

$$\mathcal{H}_{ii}^{core} = \int d\tau_1 \left\{ \chi_i^*(1) \left[-\frac{1}{2}\nabla_1^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{1A}} \right] \chi_i(1) \right\}$$

$$\mathcal{J}_{ij} = \int d\tau_1 d\tau_2 \quad \chi_i^*(1)\chi_i(1)\frac{1}{r_{12}}\chi_j^*(2)\chi_j(2)$$

$$= \int d\tau_1 d\tau_2 \quad \left|\chi_i(1)\right|^2 \frac{1}{r_{12}} \left|\chi_j(2)\right|^2$$

$$\mathcal{K}_{ij} = \int d\tau_1 d\tau_2 \quad \chi_i^*(1)\chi_j(1)\frac{1}{r_{12}}\chi_j^*(2)\chi_i(2)$$

For closed shell systems:

$$E = 2\sum_{i=1}^{N/2} H_{ii}^{core} + \sum_{i=1}^{N/2}\sum_{j=i+1}^{N/2}\left(4J_{ij} - 2K_{ij}\right) + \sum_{i=1}^{N/2} J_{ii}$$

$$= 2\sum_{i=1}^{N/2} H_{ii}^{core} + \sum_{i=1}^{N/2}\sum_{j=1}^{N/2}\left(2J_{ij} - K_{ij}\right)$$

Core Hamiltonian

$$\mathcal{H}_{ii}^{core} = \int d\tau_1 \left\{ \chi_i^*(1) \left[-\frac{1}{2}\nabla_1^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{1A}} \right] \chi_i(1) \right\}$$

- The kinetic and potential energy of an electron moving in the field of the nuclei.
- The potential energy in the core Hamiltonian makes a favorable (i.e., negative) contribution to the total energy.

Coulomb integral

$$
\begin{aligned}
\mathcal{J}_{ij} &= \int d\tau_1 d\tau_2 \quad \chi_i^*(1)\chi_i(1)\frac{1}{r_{12}}\chi_j^*(2)\chi_j(2) \\
&= \int d\tau_1 d\tau_2 \quad \left|\chi_i(1)\right|^2 \frac{1}{r_{12}}\left|\chi_j(2)\right|^2
\end{aligned}
$$

- The Coulomb integral arises from the electrostatic repulsion between pairs of electrons.
- This term makes unfavorable contribution to the total energy.

Exchange integral

$$\mathcal{K}_{ij} = \int d\tau_1 d\tau_2 \quad \chi_i^*(1)\chi_j^*(2)\frac{1}{r_{12}}\chi_j(1)\chi_i(2)$$

- Electrons with the same spin tend to 'avoid' each other, and therefore experience a lower Columbic repulsion.
- There is no classical counterpart because the origin of this term comes from the Pauli exclusion principle.
- The exchange integral corresponds to the correlated motions of electrons: whereas there is a finite probability of finding two electrons with opposite spins in the same region of space, the probability will be zero when the two spins are the same.
- This integral is non-zero only when the spins of the electrons in the spin orbital χ_i and χ_j are the same (parallel).
- This term ($-K_{ij}$) makes favorable contribution to the total energy.

Short-hand Notations

$$\langle i|\mathcal{O}|j\rangle=\langle\chi_i|\mathcal{O}|\chi_j\rangle=\int d\tau_1\quad\chi_i^*(1)\mathcal{O}(1)\chi_j(1)=[i|\mathcal{O}|j]$$

$$\langle ij|kl\rangle=\langle\chi_i\chi_j|\chi_k\chi_l\rangle=\int d\tau_1 d\tau_2\quad\chi_i^*(1)\chi_j^*(2)\frac{1}{r_{12}}\chi_k(1)\chi_l(2)=[ik|jl]$$

$$[ij|kl]=[\chi_i\chi_j|\chi_k\chi_l]=\int d\tau_1 d\tau_2\quad\chi_i^*(1)\chi_j(1)\frac{1}{r_{12}}\chi_k^*(2)\chi_l(2)=\langle ik|jl\rangle$$

$$\langle ij||kl\rangle=\langle ij|kl\rangle-\langle ij|lk\rangle=\int d\tau_1 d\tau_2\quad\chi_i^*(1)\chi_j^*(2)\frac{1}{r_{12}}(1-\mathcal{P}_{12})\chi_k(1)\chi_l(2)$$

$$(i|\mathcal{O}|j)=\mathcal{O}_{ij}=\int d\tau_1\quad\psi_i^*(1)\mathcal{O}(1)\psi_j(1)$$

$$(ij|kl)=(\chi_i\chi_j|\chi_k\chi_l)=\int d\tau_1 d\tau_2\quad\psi_i^*(1)\psi_j(1)\frac{1}{r_{12}}\psi_k^*(2)\psi_l(2)$$

$$J_{ij}=(ii|jj)$$

$$K_{ij}=(ij|ji)$$

Linear Combination of Atomic Orbitals (LCAO)

Each molecular orbital is expressed as a linear combination of atomic orbitals:

$$\psi_i = \sum_{\mu=1}^{K} c_{\mu i} \phi_\mu$$

ψ_i : spatial molecular orbital

ϕ_μ : one of K atomic orbitals

$c_{\mu i}$: coefficient

Minimal Basis Set

- In principle, a minimal basis set contains just the number of functions that are required to accommodate all the filled orbitals in each atom.
- In practice, a minimal basis set normally includes all of the atomic orbitals in the shell.
- Examples: 1s for H and He; 1s, 2s, and 2p for Li, Be, B, C, N, O, F, Ne.

Slater-type Orbitals (STO)

$$R_{nl}(r) = (2\zeta)^{n+1/2}[(2n)!]^{-1/2} r^{n-1} e^{-\zeta r}$$

$$R_{1s}(r) = (2\zeta)^{3/2} e^{-\zeta r}$$

$$R_{2s}(r) = R_{2p}(r) = \sqrt{\frac{4\zeta^5}{3}}\, r e^{-\zeta r}$$

$$R_{3s}(r) = R_{3p}(r) = R_{3d}(r) = \sqrt{\frac{8\zeta^7}{45}}\, r^2 e^{-\zeta r}$$

$$\phi_{1s}(\mathbf{r}) = \sqrt{\frac{\zeta^3}{\pi}}\, e^{-\zeta r}$$

$$\phi_{2s}(\mathbf{r}) = \sqrt{\frac{\zeta^5}{3\pi}}\, \mathbf{r}\, e^{-\zeta r}$$

$$\phi_{2p}(\mathbf{r}) = \sqrt{\frac{\zeta^5}{\pi}}\, e^{-\zeta r} \cos\theta$$

$$\zeta = \frac{Z - \sigma}{n^*}$$

ζ: orbital exponent

σ: shielding constant

*n**: effective principal quantum number

Slater Rules

$$\zeta = \frac{Z - \sigma}{n^*}$$

ζ: orbital exponent

σ: shielding constant

n^*: effective principal quantum number

$$n^*(n) = n, \quad \text{for } n = 1,2,3;$$
$$n^*(4) = 3.7;$$
$$n^*(5) = 4.0;$$
$$n^*(6) = 4.2$$

For a given orbital, σ is obtained by summing up the following contributions:

a) From an orbital further from the nucleus than those in the group, 0;

b) From each other electron in the same group, 0.35, but if the other orbital is the 1s, 0.3;

c) For each electron in a group with a principal quantum number 2 or more fewer than the current orbital, 1.0;

d) For each electron with a principal quantum number 1 fewer than the current orbital: 1.0 if the current orbital is d or f; 0.85 if the current orbital is s or p.

Gaussian-type Orbitals (GTO)

$$x^a y^b z^c \, e^{-\alpha r^2}$$

$$g_s(\alpha, r) = \left(\frac{2\alpha}{\pi}\right)^{4/3} e^{-\alpha r^2}$$

$$g_x(\alpha, r) = \left(\frac{128\alpha^5}{\pi^3}\right)^{1/4} x e^{-\alpha r^2}$$

$$g_y(\alpha, r) = \left(\frac{128\alpha^5}{\pi^3}\right)^{1/4} y e^{-\alpha r^2}$$

$$g_z(\alpha, r) = \left(\frac{128\alpha^5}{\pi^3}\right)^{1/4} z e^{-\alpha r^2}$$

$$g_{xx}(\alpha, r) = \left(\frac{2048\alpha^7}{9\pi^3}\right)^{1/4} x^2 e^{-\alpha r^2}$$

$$g_{xy}(\alpha, r) = \left(\frac{2048\alpha^7}{9\pi^3}\right)^{1/4} xy \, e^{-\alpha r^2}$$

$$g_{3zz-rr}(\alpha, r) = \frac{1}{2}\left(2g_{zz} - g_{xx} - g_{yy}\right)$$

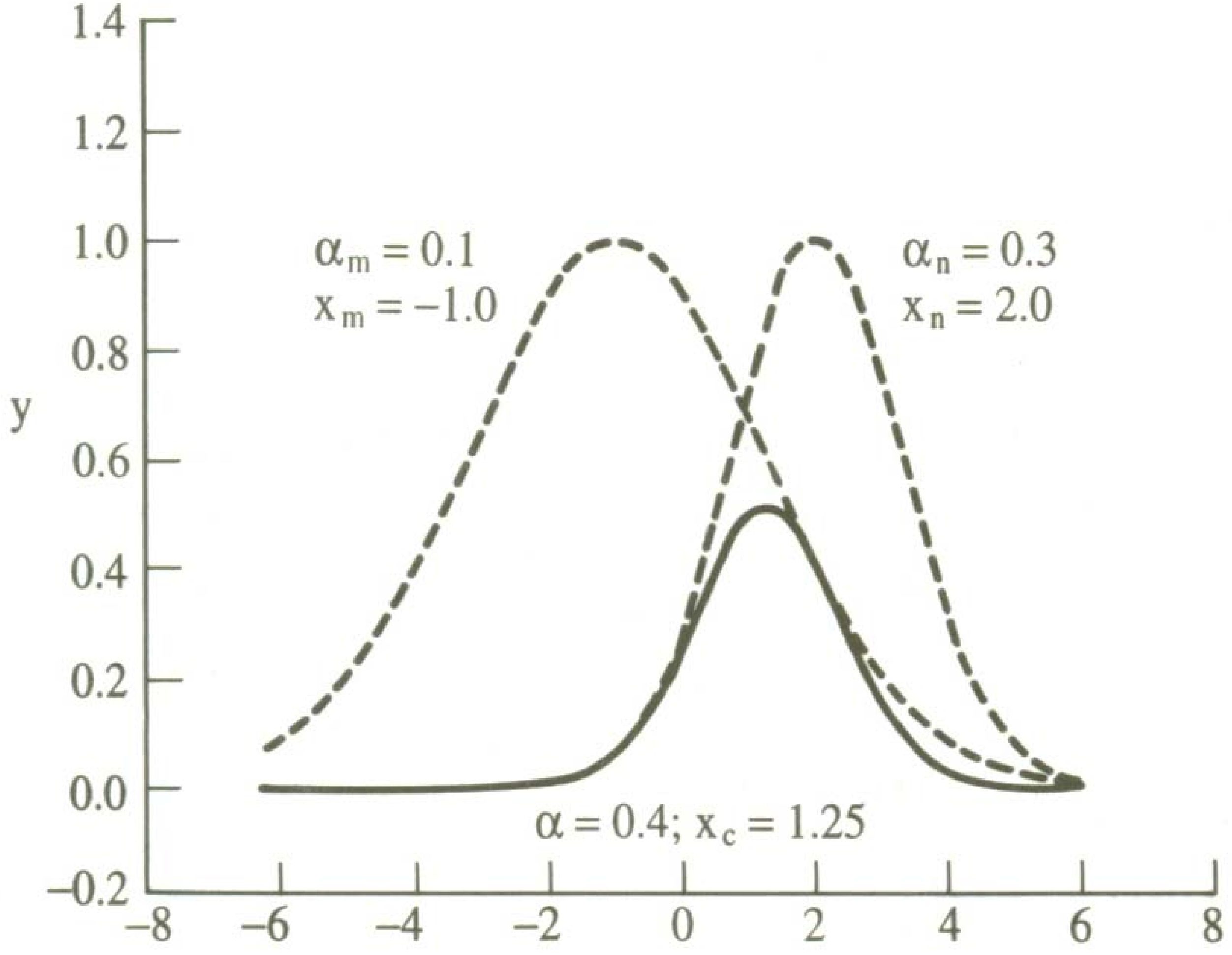
$$g_{xx-yy}(\alpha, r) = \sqrt{\frac{3}{4}}\left(g_{xx} - g_{yy}\right)$$

$$g_{rr}(\alpha, r) = \sqrt{5}\left(g_{xx} + g_{yy} + g_{zz}\right)$$

Primitive GTOs and Contracted GTOs

- In the Cartesian scheme shown in the previous slide, there are $(l+1)/(l+2)/2$ components of a given angular momentum l, whereas the number of independent spherical harmonics is only $2l+1$.
- For a more compact and accurate description of the electronic structure, the GTOs are not used individually as primitive GTOs but mostly as contracted GTOs (i.e., as fixed, linear combinations of primitive GTOs with different exponents a.)

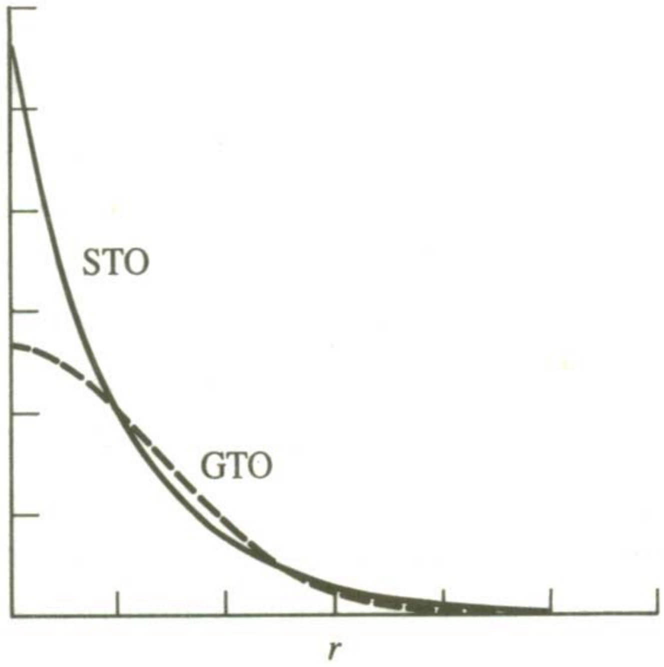
Product of two Gaussian Functions is another Gaussian

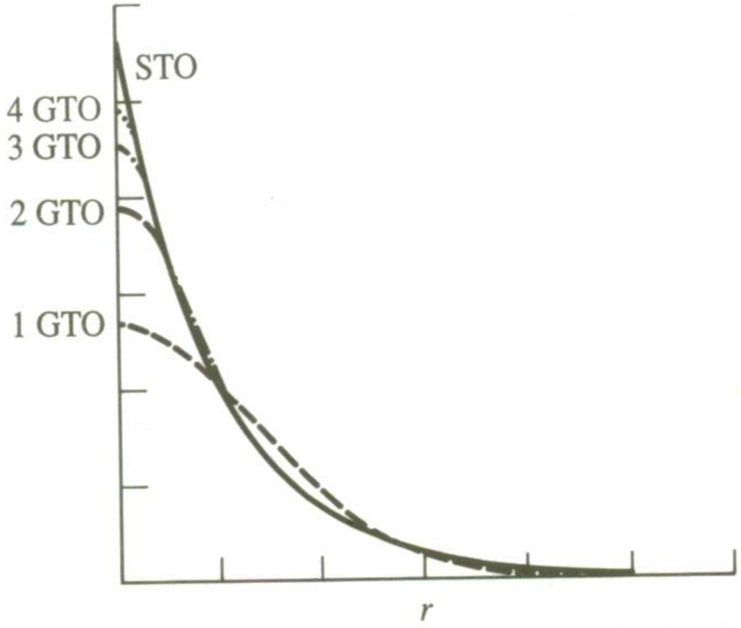


STO-*L*G

- *L* Gaussian functions are used to represent the Slater-type orbitals.
- It is found that at least three Gaussian functions are required to properly represent each STO and so the **STO-3G** is the **absolute minimum** for *ab initio* MO calculations.

Approximating STO by GTO





Deficiencies of Minimal Basis Sets

- Compounds with atoms at the end of a period, e.g., oxygen or fluorine, cannot be described properly.
- The functions cannot contract or expand in size in accordance with the molecular environment.
- Non-spherical aspects of electronic distribution cannot be described properly.

Double ζ (zeta) basis set

- A basis set which doubled the number in the minimum basis set is described as a ***double ζ (zeta)*** basis.
- The anisotropic electron distribution can be described more properly because different linear combinations for p_x, p_y, p_z orbitals are possible.

Split valence double ζ (zeta) basis set

For valence electrons:
2 contracted basis functions
1 diffuse basis function

3-21G

For core orbitals:
3 Gaussian basis functions

- The rationale behind is that the core orbitals should not affect the chemical properties very much and are similar for different elements.
- Other examples of split valence double ζ (zeta) basis set include 4-31G, 6-31G,etc.

Polarization functions

- Introduction of polarization functions, i.e., orbitals with higher angular quantum numbers, helps to describe the anisotropic electron distribution more accurately.
- For first row elements, this means to introduce *p* orbitals; for second row elements, *d* orbitals.
- 6-31 G*: in addition to 6-31G, introduce polarization functions only on the heavy (i.e., non-hydrogen) atoms.
- 6-31 G**: in addition to 6-31G, introduce polarization functions also on hydrogen and helium atoms.

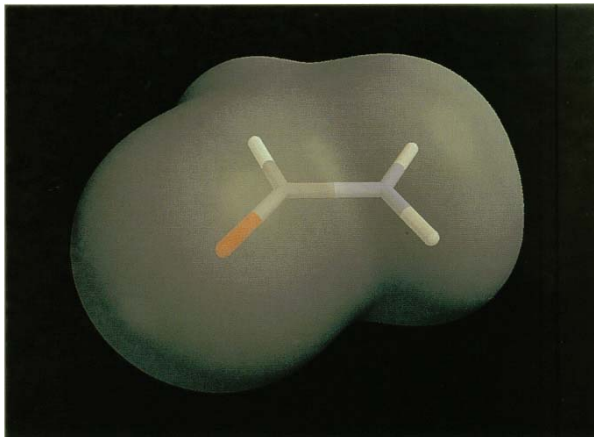
cc-pCV*n*Z basis sets

- **‘correlation-consistent’ polarized Core and Valence** (**D**ouble/**T**riple/**Q**uadruple etc.) **Zeta** basis sets. Introduction by Dunning and co-workers, aiming to have ‘balanced’ basis set.
- Balanced double-ζ basis sets should put *d* functions on heavy atoms and *p* functions on H; balanced triple- ζ basis sets should include 1 set of *f* and 2 sets of *d* functions on heavy atoms, and 1 set of *d* and 2 sets of *p* functions on H, etc.
- An ideal approach to push to the **Hartree-Fock** (infinite basis set) **limit**.

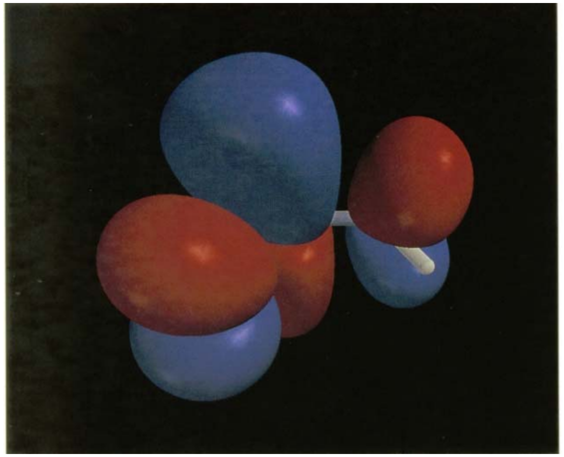
Basis set superposition error (BSSE)

- The dissociation energy of a molecular complex (also called supermolecule) is always overestimated, which is due to the basis set superposition error.
- BSSE will be particularly significant when small, inadequate basis sets are used.
- BSSE can be estimated by the counterpoise correction method.
- The relevance of the BSSE and its dependence on the basis set and the level of theory employed remains a subject of much research.

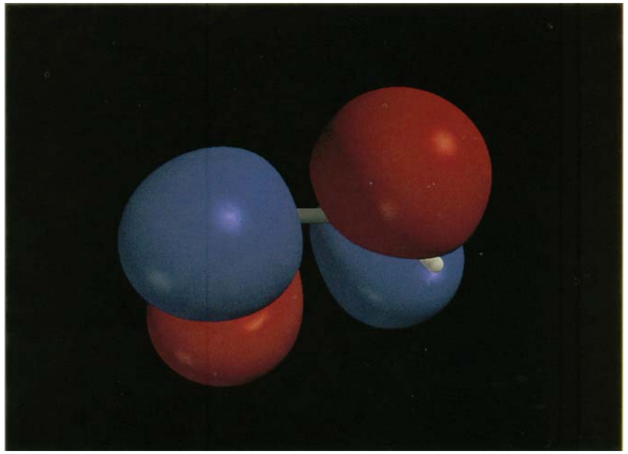
HOMO and LUMO



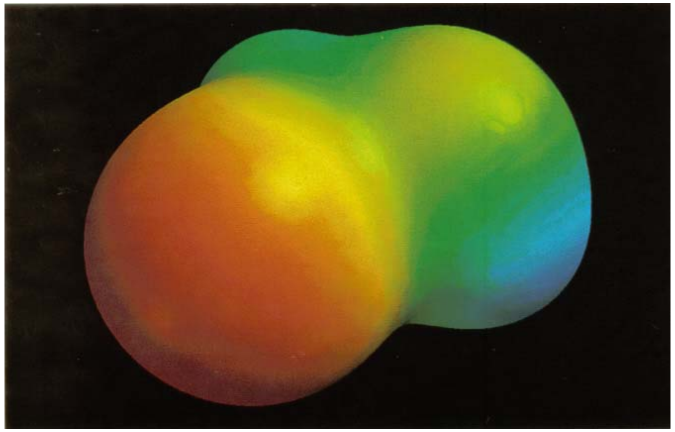
Surface representation of electron density around formamide at a contour of 0.0001 au (eletrons/bohr^3)



Lowest unoccupied molecular orbital (LUMO) of formamide.



Highest occupied molecular orbital (HOMO) of formamide. Red: negative; blue the positive.



Electrostatic potential mapped onto the electron density surface of formamide. Red: negative; blue the positive.

Koopman’s theorem

Koopman’s theorem: the energy of an electron in an orbital is equated with the energy required to remove the electron to give the corresponding ion.

Electron affinities predicted by Koopman’s theorem are always **positive** when Hartee-Fock calculations are used, because virtual orbital always have a higher energy.

Ionization Potential, Electron Affinity, and Electronegativity

Ionization potential:

$$-I = E_N - E_{N-1} = \int_0^1 \varepsilon_{HOMO}(n)dn \approx \varepsilon_{HOMO}\Big|_{n=\frac{1}{2}}$$

Electron affinity:

$$-A = E_{N+1} - E_N = \int_0^1 \varepsilon_{LUMO}(n)dn \approx \varepsilon_{LUMO}\Big|_{n=\frac{1}{2}}$$

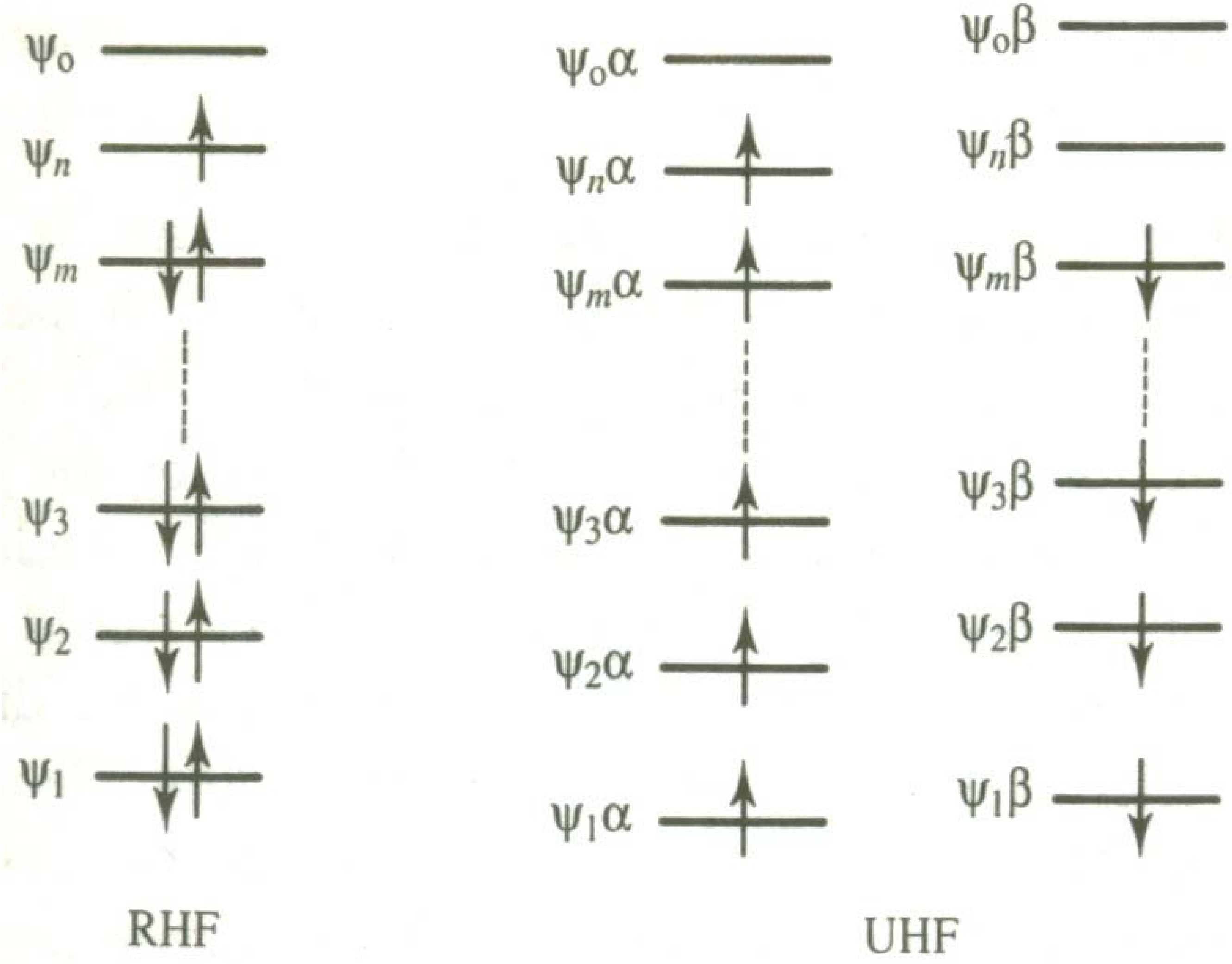
Electronegativity:

$$\chi = \frac{I+A}{2} \approx -\frac{\varepsilon_{HOMO}\Big|_{n=\frac{1}{2}} + \varepsilon_{LUMO}\Big|_{n=\frac{1}{2}}}{2}$$

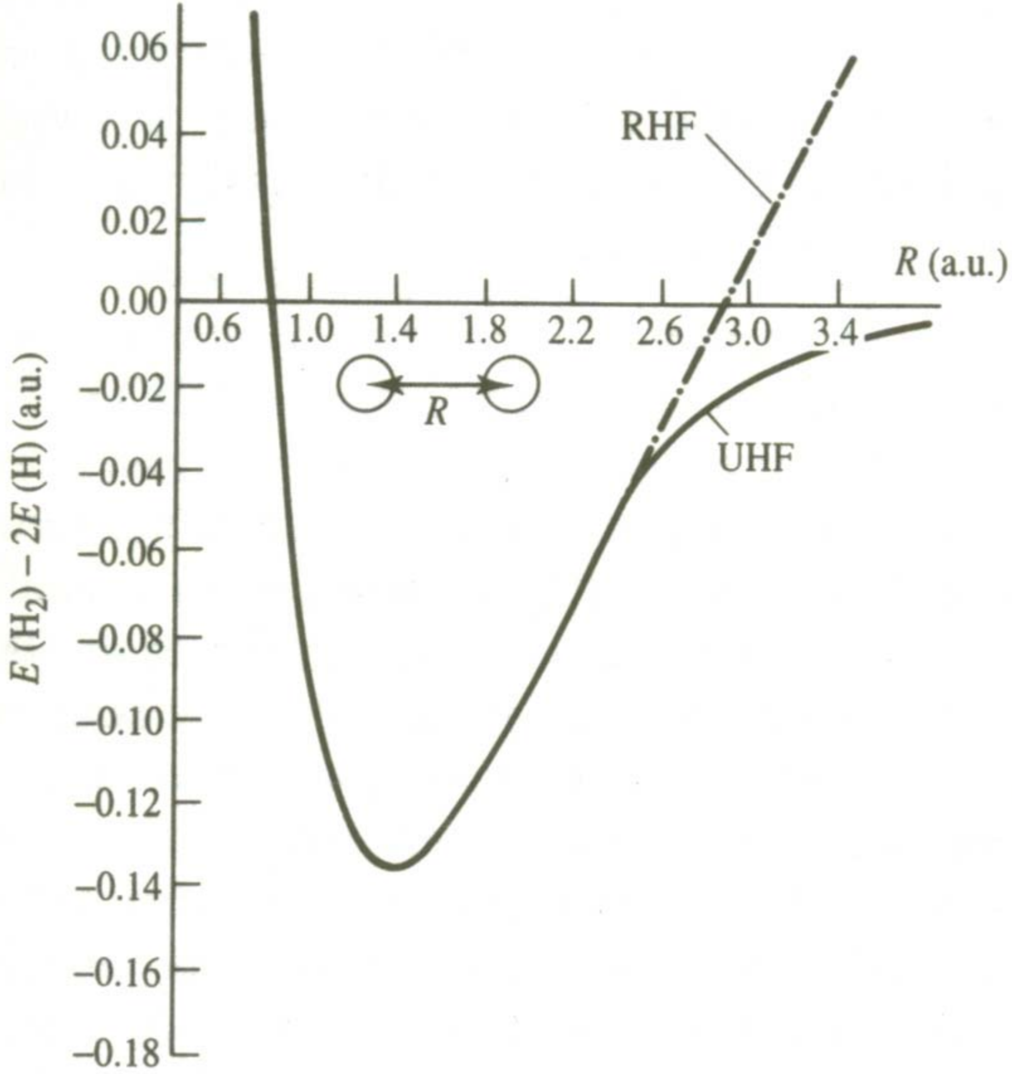
Open-Shell Systems

- The Roothaan-Hall equations are not applicable to open-shell systems, which contains one or more unpaired electrons.
- **Radicals** are by definition open-shell systems as are some ground-state molecules such as NO and O_2.
- Spin-**r**estricted **H**artree-**F**ock (RHF) theory and spin-**u**nrestricted **H**artree-**F**ock (UHF) theory are possible theoretical approaches to treat the open-shell systems.

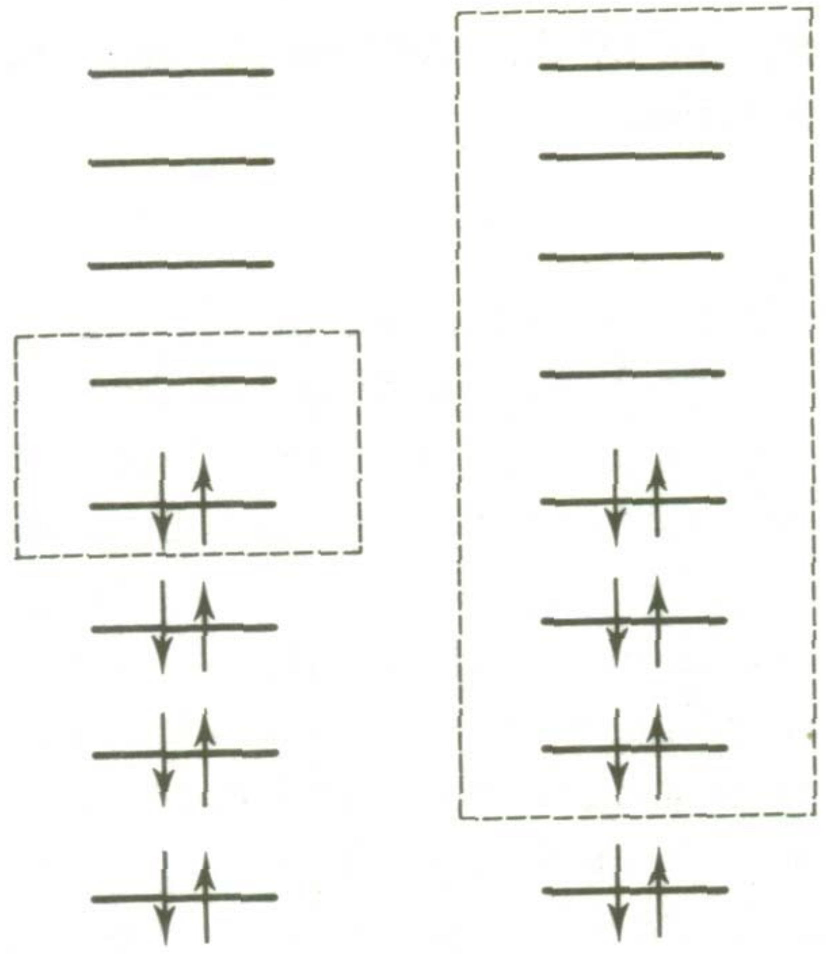
Conceptual Differences between RHF and UHF



H_2 Dissociation Curve using RHF and UHF



Treating Excited States



HOMO + LUMO only

Valence orbitals only

Electron Correlation

- The Hartree-Fock theory does not describe the electron correlation adequately.
- The correlation energy is defined as the **difference between the Hartree-Fock energy and the exact energy**.
- An uncorrelated calculation would predict that the electrons in H_2 spend equal time on both nuclei, even when they are infinitely separated. This is obviously wrong!
- Inclusion of correlation effect is important, especially when quantitative information is required.
- Electron correlation is crucial in the study of **dispersive effect**, which plays a major role in intermolecular interactions.

Møller-Plesset Perturbation Theory (MPPT)

- In the Møller-Plesset approach, the unperturbed Hamiltonian is the sum of the one-electron Fock operators for the *N* electrons.
- To obtain an improvement on the Hartree-Fock energy it is necessary to use Møller-Plesset theory to at least second order, which is referred to as MP2.
- The advantage of such perturbation theory is that it is size-independent.
- However, MPPT is not variational and sometimes the calculated energy can be lower than the ‘true’ energy.
- Møller-Plesset calculations are computationally intensive and usually only used in ‘single-point’ (SP) calculations.
- MP2/6-31G* indicates a second order Møller-Plesset calculation with the 6-31G* basis set.

Configuration Interaction (CI)

- In the CI approach, the excited states are included in the calculations of an electronic state, even in its ground state.
- For example, the ground state of lithium can be written as $1s^22s^1$, and excitation of outer valence electron gives states such as $1s^23s^1$.
- If a Hartree-Fock calculation is performed with K basis functions, then $2K$ spin orbitals are obtained. If these $2K$ spin orbitals are filled with N electrons ($N<2K$), there will be $2K-N$ unoccupied, virtual orbitals.
- In the full CI calculation, all the permutations to place N electrons in $2K$ spin orbitals will be included, and there will be totally $2K!/[(N!(2K-N)!]$ possible configurations.

H_2 Dissociation with Various Theories

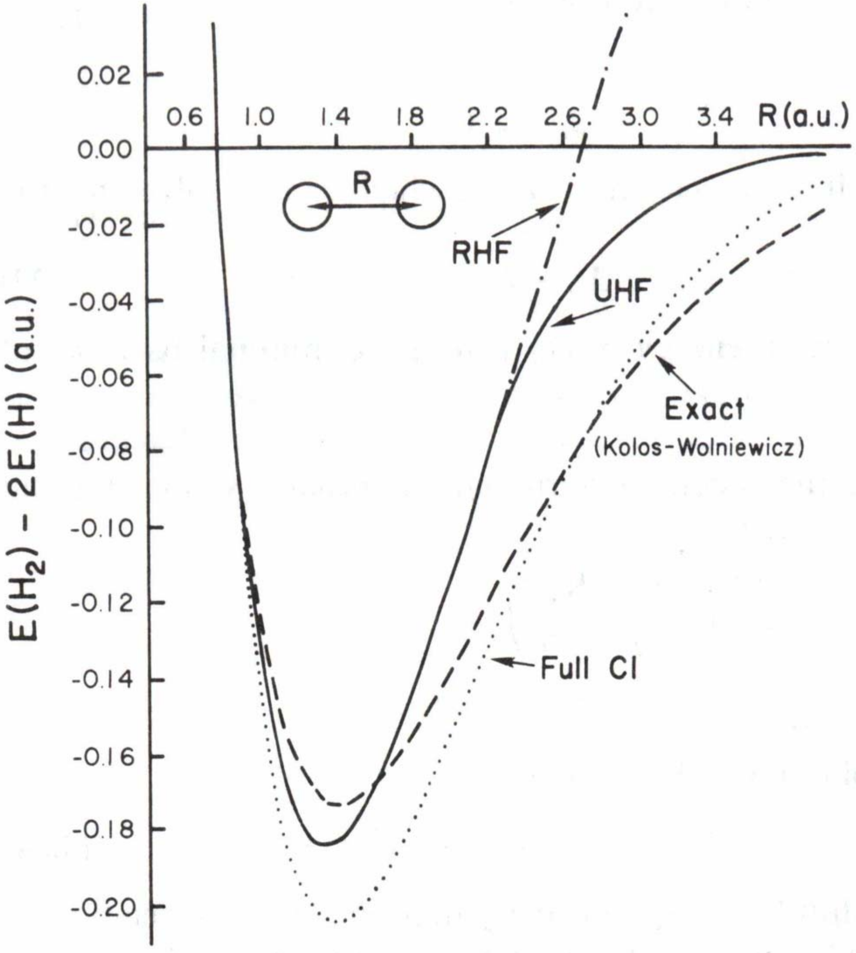


Figure 4.2 STO-3G potential energy curves for H_2.

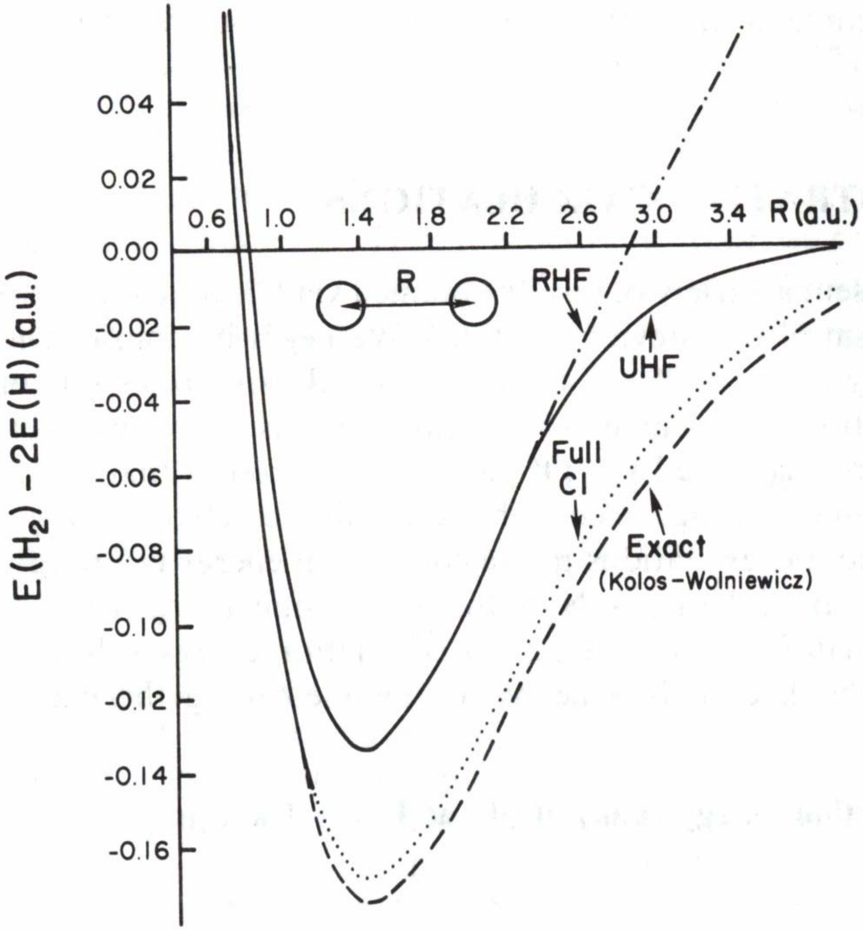
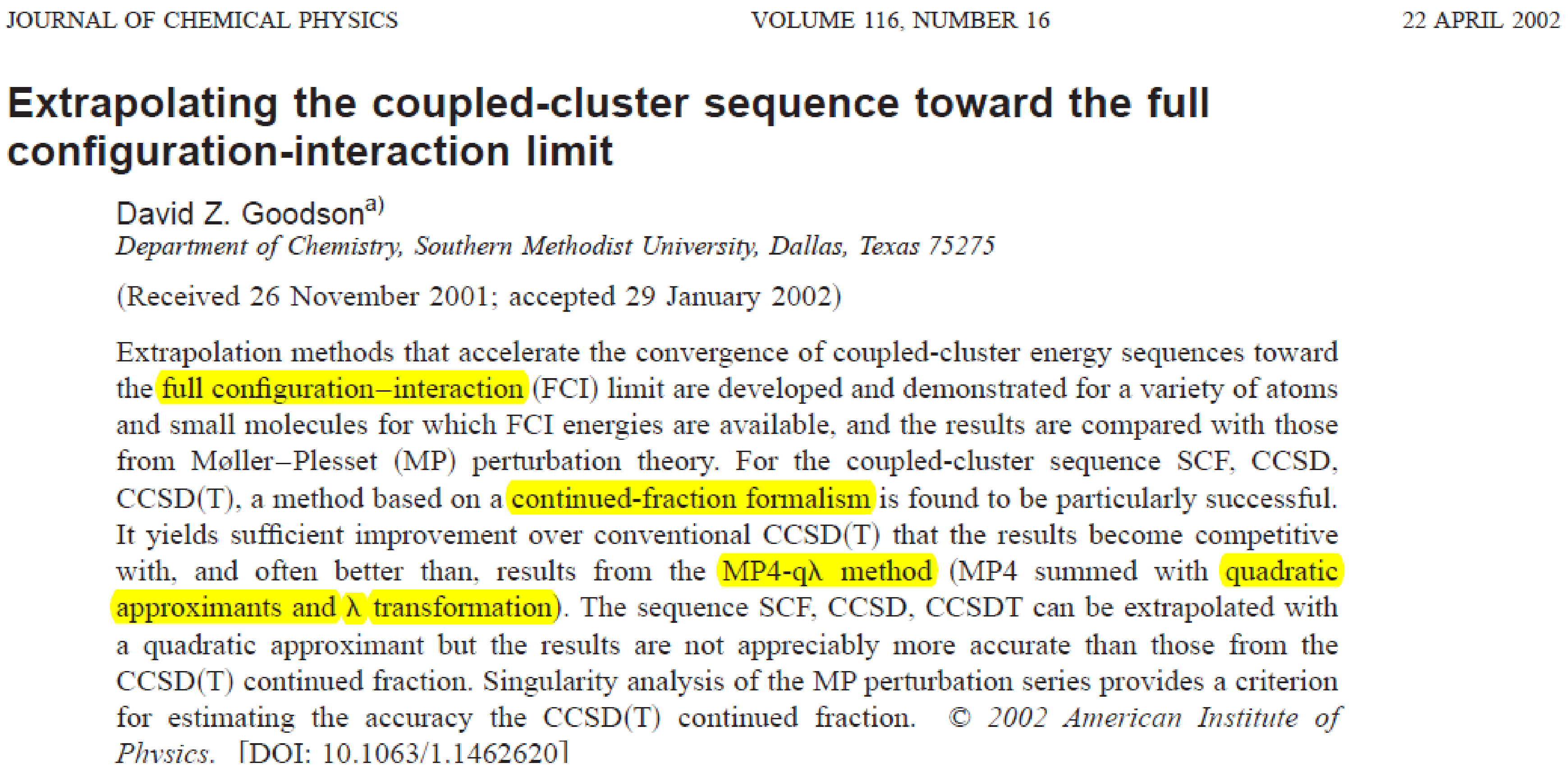


Figure 4.3 6-31G** potential energy curves for H_2.



JOURNAL OF CHEMICAL PHYSICS VOLUME 116, NUMBER 16 22 APRIL 2002

Extrapolating the coupled-cluster sequence toward the full configuration-interaction limit

David Z. Goodson[a)]
Department of Chemistry, Southern Methodist University, Dallas, Texas 75275

(Received 26 November 2001; accepted 29 January 2002)

Extrapolation methods that accelerate the convergence of coupled-cluster energy sequences toward the full configuration–interaction (FCI) limit are developed and demonstrated for a variety of atoms and small molecules for which FCI energies are available, and the results are compared with those from Møller–Plesset (MP) perturbation theory. For the coupled-cluster sequence SCF, CCSD, CCSD(T), a method based on a continued-fraction formalism is found to be particularly successful. It yields sufficient improvement over conventional CCSD(T) that the results become competitive with, and often better than, results from the MP4-qλ method (MP4 summed with quadratic approximants and λ transformation). The sequence SCF, CCSD, CCSDT can be extrapolated with a quadratic approximant but the results are not appreciably more accurate than those from the CCSD(T) continued fraction. Singularity analysis of the MP perturbation series provides a criterion for estimating the accuracy the CCSD(T) continued fraction. *© 2002 American Institute of Physics.* [DOI: 10.1063/1.1462620]

Goodson, *J. Chem. Phys.* **116**: 6948-6956 (2002)

Hohenberg-Kohn Theorem

Hohenberg and Kohn, *Phys. Rev.* **136**: B864-B871 (1964)

The total energy and *all* observable information of a system are uniquely defined by the ground-state electron density:

$$E[\rho(\mathbf{r})] = \int V_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + F[\rho(\mathbf{r})]$$

$V_{ext}(\mathbf{r})$: external potential, e.g., Coulumb interaction with the nuclei

$F[\rho(\mathbf{r})]$: kinetic energy + interelectron energy

Why DFT is attractive?

- The electron density is a much simpler entity than the wave function, which depends only on three spatial coordinates rather than on the *4n* spatial coordinates and spin configurations of *n* electrons.

- The correlation of electron motions is included, and even extremely simple approximation works for many systems.

Variational Principle

The energy should be minimum if the electron density function is correct:

$$E\left[\rho^*(\mathbf{r})\right] \leq E\left[\rho(\mathbf{r})\right]$$

$\rho(\mathbf{r})$: electron density

$\rho^*(\mathbf{r})$: correct electron density

↓

$$\frac{\delta}{\delta\rho^*(\mathbf{r})} E\left[\rho^*(\mathbf{r})\right] = 0$$

Introducing a Lagrange Multiplier

$$N = \int \rho(\mathbf{r})\, d\mathbf{r}$$

$$\frac{\delta}{\delta\rho(\mathbf{r})}\int \rho(\mathbf{r})\, d\mathbf{r} = 0$$

Lagrange multiplier

$$(-\mu)\cdot\frac{\delta}{\delta\rho(\mathbf{r})}\int \rho(\mathbf{r})\, d\mathbf{r} = 0$$

DFT Equivalent of the Schrödinger Equation

$$\frac{\delta}{\delta\rho(\mathbf{r})}E[\rho(\mathbf{r})]-\mu\int\rho(\mathbf{r})d\mathbf{r}=0$$

⬇

$$\left(\frac{\delta E[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}\right)_{V_{ext}}=\mu$$

The Kohn-Sham Scheme

Kohn and Sham, *Phys. Rev.* **140**: A1133-A1138 (1965)

The functional for the electron part:

$$F\left[\rho(\mathbf{r})\right] = E_{KE}\left[\rho(\mathbf{r})\right] + E_{H}\left[\rho(\mathbf{r})\right] + E_{XC}\left[\rho(\mathbf{r})\right]$$

$E_{KE}\left[\rho(\mathbf{r})\right]$: kinetic energy

$E_{H}\left[\rho(\mathbf{r})\right]$: Hartree electrostatic energy

$E_{XC}\left[\rho(\mathbf{r})\right]$: exchange - correlation energy

Kohn-Sham Kinetic Energy

The kinetic energy of a system of non-interacting electrons:

$$E_{KE}[\rho(\mathbf{r})] = \sum_{i=1}^{N} \int \psi_i^*(\mathbf{r}) \left(-\frac{\nabla^2}{2} \right) \psi_i(\mathbf{r}) d\mathbf{r}$$

Hartree Electrostatic Energy

The electrostatic energy of a system of non-interacting electrons:

$$E_H\left[\rho(\mathbf{r})\right] = \frac{1}{2}\iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2$$

Total Energy of an N-electron System within the Kohn-Sham Scheme

$$E\left[\rho(\mathbf{r})\right]=\sum_{i=1}^{N}\int\psi_i^{*}(\mathbf{r})\left(-\frac{\hbar^2}{2m_e}\nabla^2\right)\psi_i(\mathbf{r})d\mathbf{r}+\frac{1}{2}\iint\frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_1d\mathbf{r}_2$$

$$+E_{XC}[\rho(\mathbf{r})]-\sum_{A=1}^{M}\int\frac{Z_A}{\left|\mathbf{r}-\mathbf{R}_A\right|}\rho(\mathbf{r})d\mathbf{r}$$

Kohn-Sham Density

Kohn and Sham wrote the density as the sum of the square moduli of a set of one-electron orthonormal orbitals:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} \psi_i^{*}(\mathbf{r})\psi_i(\mathbf{r})$$

$$= \sum_{i=1}^{N} \left|\psi_i(\mathbf{r})\right|^2$$

One-electron Kohn-Sham Equations

$$\left\{\left(-\frac{\nabla_1^2}{2}\right)-\left(\sum_{A=1}^{M}\frac{Z_A}{\left|\mathbf{r}_{1A}\right|}\right)+\int\frac{\rho(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_2+V_{XC}\left[\mathbf{r}_1\right]\right\}\psi_i(\mathbf{r}_1)=\varepsilon_i\psi_i(\mathbf{r}_1)$$

$$V_{XC}\left[\mathbf{r}_1\right]=\left(\frac{\delta E_{XC}\left[\rho(\mathbf{r})\right]}{\delta\rho(\mathbf{r})}\right)$$

Spin-polarized Kohn-Sham Equations

Two sets of wave functions, one for each spin:

$$\left\{\left(-\frac{\nabla_1^2}{2}\right)-\left(\sum_{A=1}^{M}\frac{Z_A}{|\mathbf{r}_{1A}|}\right)+\int\frac{\rho(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_2+V_{XC}[\mathbf{r}_1,\sigma]\right\}\psi_i^{\sigma}(\mathbf{r}_1)=\varepsilon_i^{\sigma}\psi_i^{\sigma}(\mathbf{r}_1)$$

$$\sigma=\alpha,\beta$$

$$\sigma(\mathbf{r})=\rho_{\uparrow}(\mathbf{r})-\rho_{\downarrow}(\mathbf{r})$$

The Local Density Approximation for the Exchange-Correlation Energy

One reason why DFT is so appealing is that even relatively simple approximations to the exchange-correlation functional can give favorable results.

In the so-called local density approximation (LDA), which is based on the uniform electron gas model, assumes that the electron density is “constant” throughout all space, even when a system (e.g., a molecule) with inhomogeneous electron distributions is considered.

LDA Exchange-Correlation Energy

The total exchange-correlation energy of a system under the local density approximation (LDA):

$$E_{XC}[\rho(\mathbf{r})] = \frac{1}{2}\int \rho(\mathbf{r})\varepsilon_{XC}[\rho(\mathbf{r})]d\mathbf{r}$$

$\varepsilon_{XC}[\rho(\mathbf{r})]$: exchange correlation energy per electron

LDA Exchange-Correlation Functional

The exchange-correlation functional of a system under the local density approximation (LDA):

$$V_{XC}\left[\rho(\mathbf{r})\right] = \rho(\mathbf{r}) \frac{\delta \varepsilon_{XC}\left[\rho(\mathbf{r})\right]}{\delta \rho(\mathbf{r})} + \varepsilon_{XC}\left[\rho(\mathbf{r})\right]$$

Gunnarsson-Lundqvist Exchange-Correlation Functional

Gunnarsson and Lundqvist, *Phys. Rev. B* **13**: 4274-4298 (1976)

$$\varepsilon_{XC}\left[\rho(\mathbf{r})\right] = -\frac{0.458}{r_s} - 0.0666\,G\left(\frac{r_s}{11.4}\right)$$

$$G(x) = \frac{1}{2}\left[(1+x)\log\left(1+x^{-1}\right) - x^2 + \frac{x}{2} - \frac{1}{3}\right]$$

$$r_s = \sqrt[3]{\frac{3}{4\pi\rho(\mathbf{r})}}$$

Slater Exchange-Only Functional

Slater, *Quantum Theory of Molecular and Solids.*
Vol. 4: The Self-Consistent Field for Molecular and Solids
(McGraw-Hill, New York, 1974)

$$E_X\left[\rho_\alpha(\mathbf{r}),\rho_\beta(\mathbf{r})\right]=-\frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3}\int\left(\rho_\alpha^{4/3}(\mathbf{r})+\rho_\beta^{4/3}(\mathbf{r})\right)d\mathbf{r}$$

Perdew-Zunger Correlation Functional

Perdew and Zunger, *Phys. Rev.* B **23**: 5048-5079 (1981)

$$\varepsilon_C[\rho(\mathbf{r})]=\begin{cases} -0.1423/(1+1.9529r_s^{1/2}+0.3334r_s), & r_s \geq 1 \\ -0.0480+0.0311\ln r_s-0.0116r_s+0.0020r_s\ln r_s, & r_s<1 \end{cases}$$

$$r_s=\sqrt[3]{\frac{3}{4\pi\rho(\mathbf{r})}}$$

Vosko-Wilk-Nusair Correlation Functional

Vosko, Wilk, and Nusair, *Can. J. Phys.* **58**: 1200-1211 (1980)

$$\varepsilon_C[\rho(\mathbf{r})] = \left\{ \frac{A}{2} \ln \frac{x^2}{X(x)} + \frac{2b}{Q} \tan^{-1} \frac{Q}{2x+b} - \frac{bx_0}{X(x_0)} \left[\ln \frac{(x-x_0)^2}{X(x)} + \frac{2(b+2x_0)}{Q} \tan^{-1} \frac{Q}{2x+b} \right] \right\}$$

$$r_s = \sqrt[3]{\frac{3}{4\pi\rho(\mathbf{r})}}, \quad x = r_s^{1/2}, \quad X(x) = x^2 + bx + c, \quad Q = \left(4c - b^2\right)^{1/2};$$

$$A = 0.0621814, \quad x_0 = -0.409286, \quad b = 13.0720, \quad c = 42.7198$$

Linear Combination of Atom-Centered Basis Functions

Each Kohn-Sham orbital is expressed as a linear combination of atom-centered basis functions:

$$\psi_i(\mathbf{r}) = \sum_{\nu=1}^{K} c_{\nu i}\phi_\nu(\mathbf{r})$$

ψ_i : Kohn - Sham orbital

ϕ_ν : one of K basis functions

$c_{\nu i}$: coefficient

$$\left\{\left(-\frac{\nabla_1^2}{2}\right)-\left(\sum_{A=1}^{M}\frac{Z_A}{|\mathbf{r}_{1A}|}\right)+\int\frac{\rho(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_2+V_{XC}[\mathbf{r}_1]\right\}\psi_i(\mathbf{r}_1)=\varepsilon_i\psi_i(\mathbf{r}_1)$$

Becke Gradient-Corrected Exchange Functional

Becke, *Phys. Rev.* A**38**: 3098-3100 (1988);
Becke, *J. Chem. Phys.* **96**: 2155-2160 (1992)

$$E_X\left[\rho(\mathbf{r})\right] = E_X^{LDA}\left[\rho(\mathbf{r})\right] - b\sum_{\sigma=\alpha,\beta}\int \rho_\alpha^{4/3}(\mathbf{r})\frac{x_\sigma^2}{\left(1+6bx_\sigma \sinh^{-1} x_\sigma\right)}d\mathbf{r}$$

$$x_\sigma = \frac{\left|\nabla\rho_\sigma\right|}{\rho_\sigma^{4/3}}$$

$$b = 0.0042\ \text{a.u.}$$

Lee-Yang-Parr Gradient-Corrected Correlation Functional

Lee, Yang, Parr, *Phys. Rev.* B**37**: 785-789 (1988)

$$E_C\left[\rho(\mathbf{r})\right] = -a\int \frac{1}{1+d\rho^{-1/3}}\left\{r + b\rho^{-2/3}\left[C_F\rho^{5/3} - 2t_W + \left(\frac{1}{9}t_W + \frac{1}{18}\nabla^2\rho\right)e^{-cr^{-1/3}}\right]\right\}d\mathbf{r}$$

$$t_W(\mathbf{r}) = \sum_{i=1}^{N}\frac{\left|\nabla\rho_i(\mathbf{r})\right|^2}{\rho_i(\mathbf{r})} - \frac{1}{8}\nabla^2\rho_i(\mathbf{r})$$

$$C_F = \frac{3}{10}\left(3\pi^2\right)^{2/3}$$

Hybrid Functional

Becke, *J. Chem. Phys.* **98**: 1372-1377 (1993)
Becke, *J. Chem. Phys.* **98**: 5648-5652 (1993)

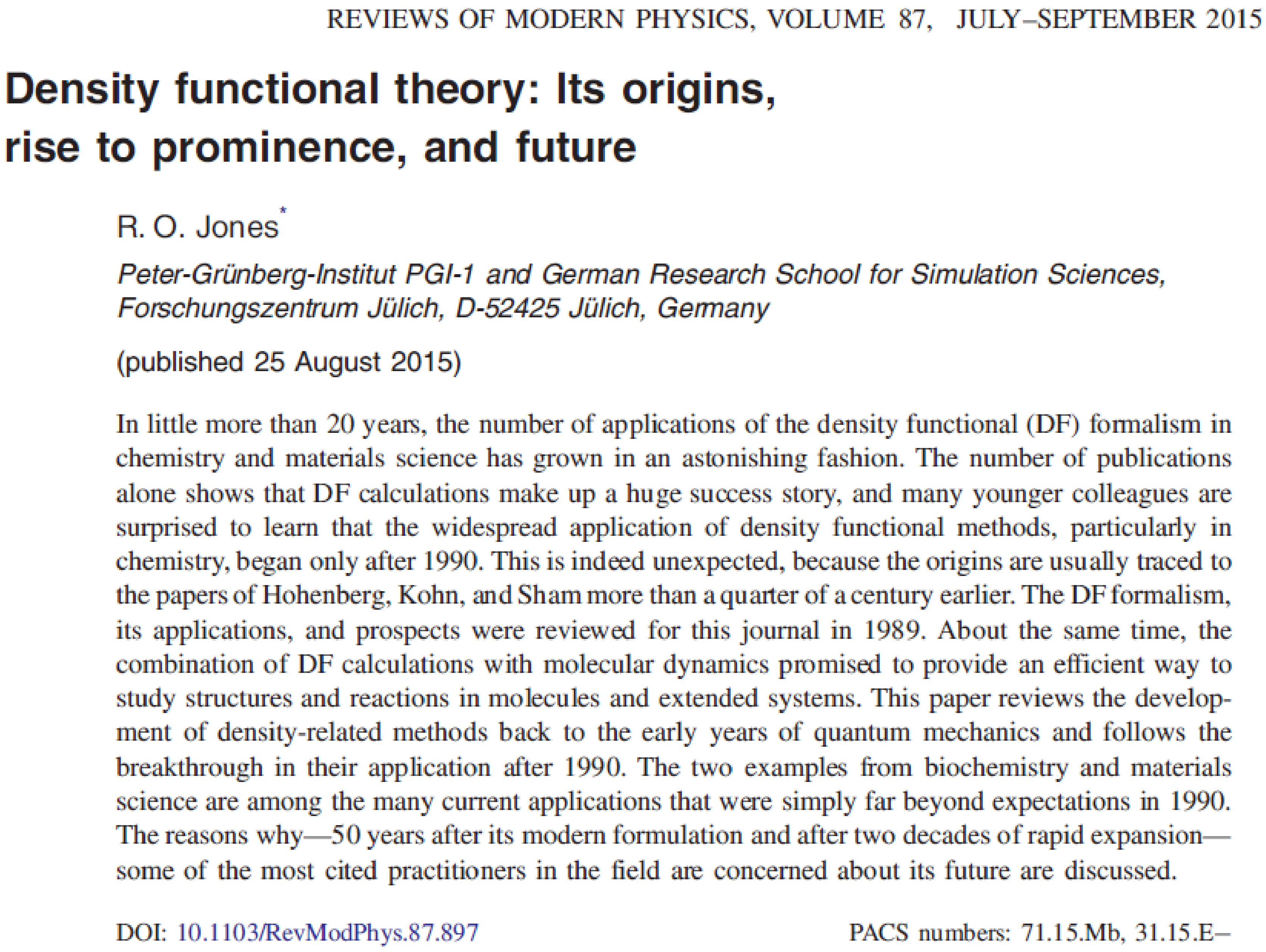
$$E_{XC} = \int_{\lambda=0}^{\lambda=1} U_{XC}^{\lambda} \, d\lambda$$

B3LYP Gradient-Corrected Exchange-Correlation Functional

Becke, *J. Chem. Phys.* **98**: 5648-5652 (1993)

$$E_{XC}^{B3LYP} = (1 - a_0)E_X^{LDA} + a_0 E_X^{HF} + a_X \Delta E_X^{B88} + a_C E_C^{LYP} + (1 - a_C)E_C^{VWN}$$

$$a_0 = 0.20, \quad a_X = 0.72, \quad a_C = 0.81$$



REVIEWS OF MODERN PHYSICS, VOLUME 87, JULY–SEPTEMBER 2015

Density functional theory: Its origins, rise to prominence, and future

R. O. Jones*

Peter-Grünberg-Institut PGI-1 and German Research School for Simulation Sciences, Forschungszentrum Jülich, D-52425 Jülich, Germany

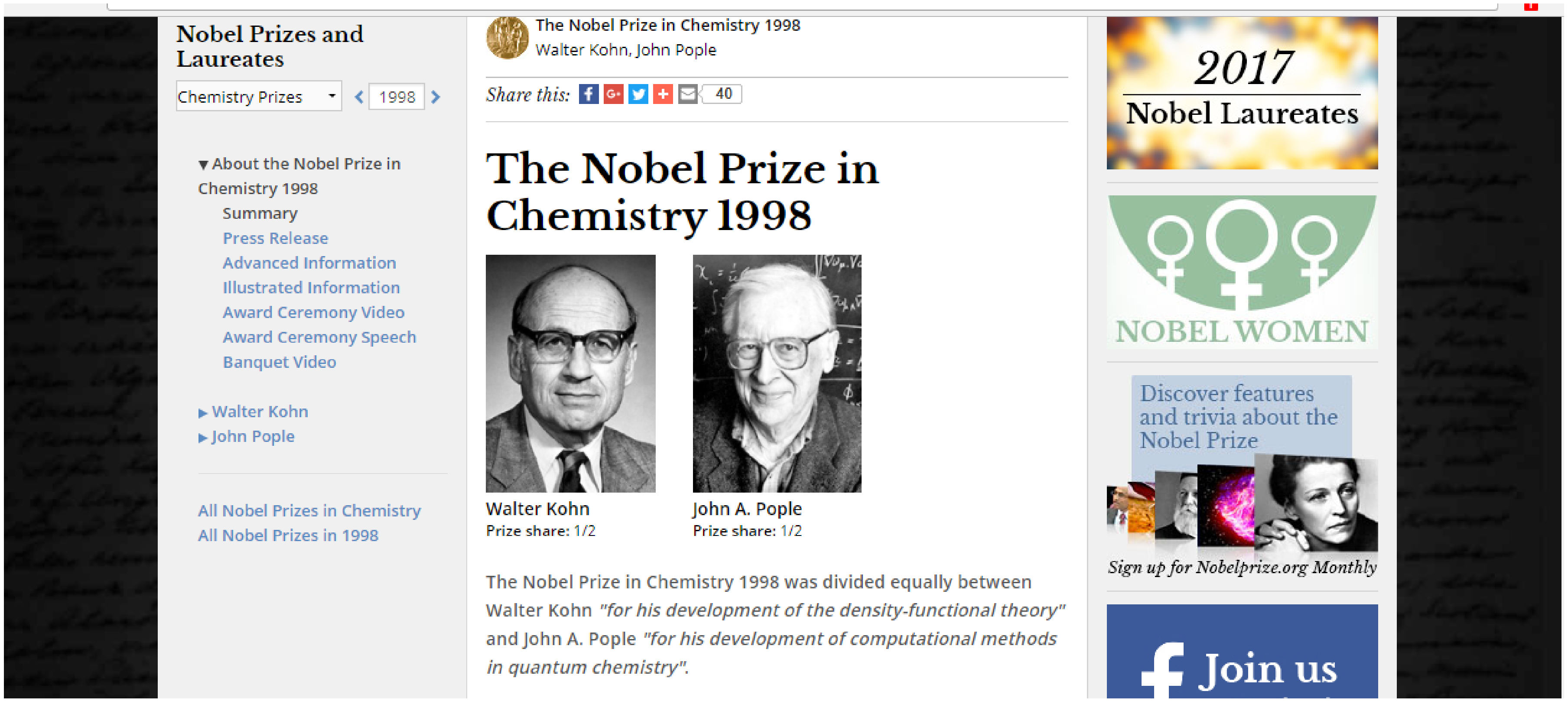
(published 25 August 2015)

In little more than 20 years, the number of applications of the density functional (DF) formalism in chemistry and materials science has grown in an astonishing fashion. The number of publications alone shows that DF calculations make up a huge success story, and many younger colleagues are surprised to learn that the widespread application of density functional methods, particularly in chemistry, began only after 1990. This is indeed unexpected, because the origins are usually traced to the papers of Hohenberg, Kohn, and Sham more than a quarter of a century earlier. The DF formalism, its applications, and prospects were reviewed for this journal in 1989. About the same time, the combination of DF calculations with molecular dynamics promised to provide an efficient way to study structures and reactions in molecules and extended systems. This paper reviews the development of density-related methods back to the early years of quantum mechanics and follows the breakthrough in their application after 1990. The two examples from biochemistry and materials science are among the many current applications that were simply far beyond expectations in 1990. The reasons why—50 years after its modern formulation and after two decades of rapid expansion—some of the most cited practitioners in the field are concerned about its future are discussed.

DOI: 10.1103/RevModPhys.87.897

PACS numbers: 71.15.Mb, 31.15.E−

Jones, *Rev. Modern. Phys.* **87**: 897-923 (2015)



The Nobel Prize in Chemistry 1998

Walter Kohn
Prize share: 1/2

John A. Pople
Prize share: 1/2

The Nobel Prize in Chemistry 1998 was divided equally between Walter Kohn *"for his development of the density-functional theory"* and John A. Pople *"for his development of computational methods in quantum chemistry"*.

Why Semi-Empirical Methods?

- Either Molecular Orbital approaches or Density Functional Theory calculations are computational too demanding.
- In some cases, the semi-empirical approaches have been shown to provide similar accuracy as the more sophisticated approaches.
- Semi-empirical methods can give insights into the qualitative trends.

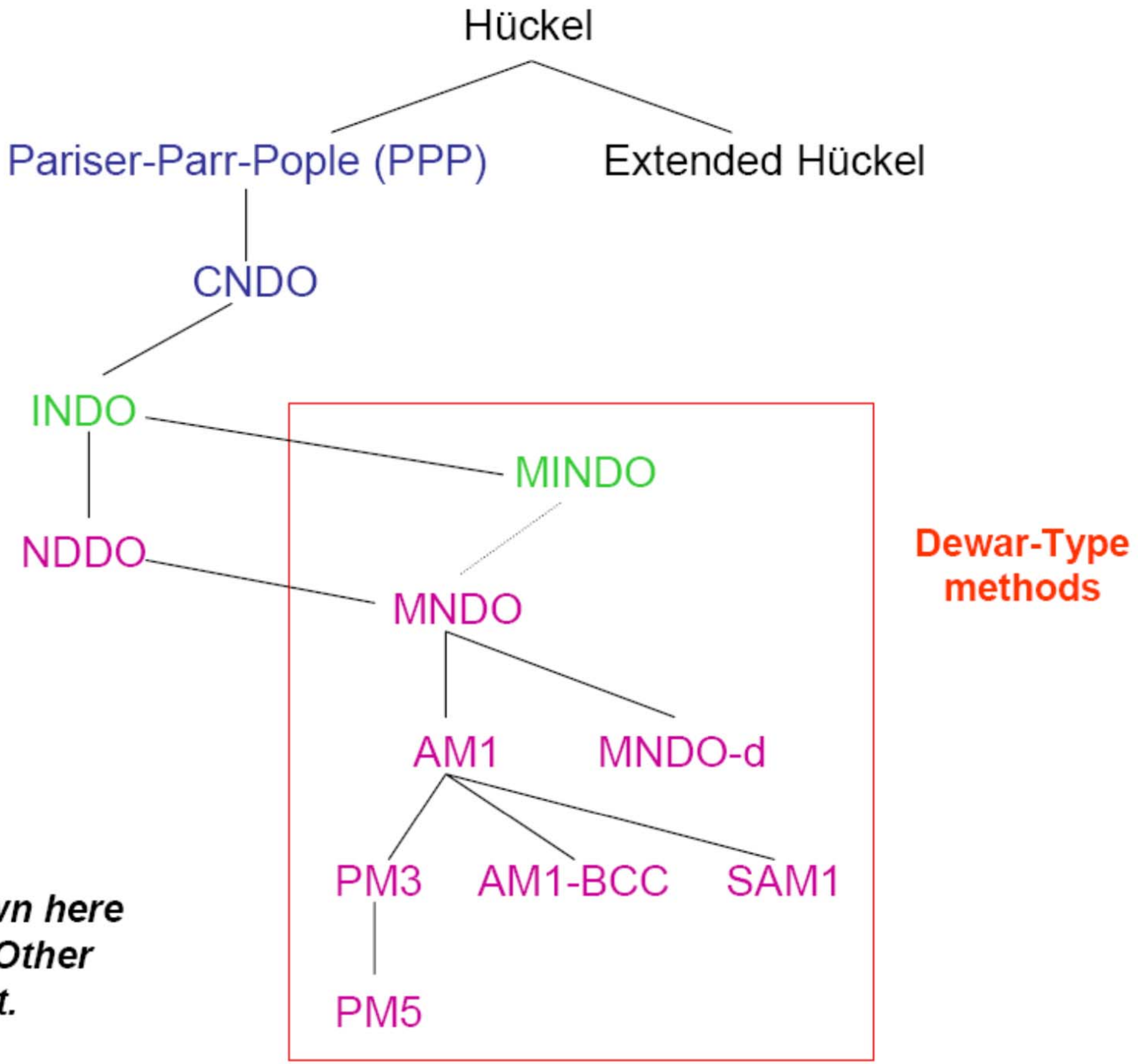
Simplifications Made in the Semi-Empirical Methods

- In the construction of the single Slater-determinant, usually only the valence electrons are treated explicitly. In some cases, only π electrons are taken into account.
- In the construction of the transformed Fock matrix, a large part of the interactions is neglected, particular in the two-electron part. Certain classes of integrals are replaced by parameterized functions.

Semi-Empirical Methods

- ZDO
- CNDO
- INDO
- NDDO
- MINDO/3
- MNDO
- AM1
- PM3
- PM6

Family Tree of Semi-Empirical Methods



Hückel

Pariser-Parr-Pople (PPP)

Extended Hückel

CNDO

INDO

MINDO

NDDO

MNDO

AM1

MNDO-d

PM3

AM1-BCC

SAM1

PM5

Dewar-Type methods

zero differential overlap

intermediate neglect of differential overlap

neglect of diatomic differential overlap

The family tree relationships shown here are approximate. Other methods also exist.

Zero-differential Overlap (ZDO)

Parsier and Parr, *J. Chem. Phys.* **21**: 466-471 (1953)

Parsier and Parr, *J. Chem. Phys.* **21**: 767-776 (1953)

Pople, *Trans. Faraday. Soc.* **49**: 1375-1385 (1953)

$$\phi_\mu \phi_\nu d\upsilon = 0$$

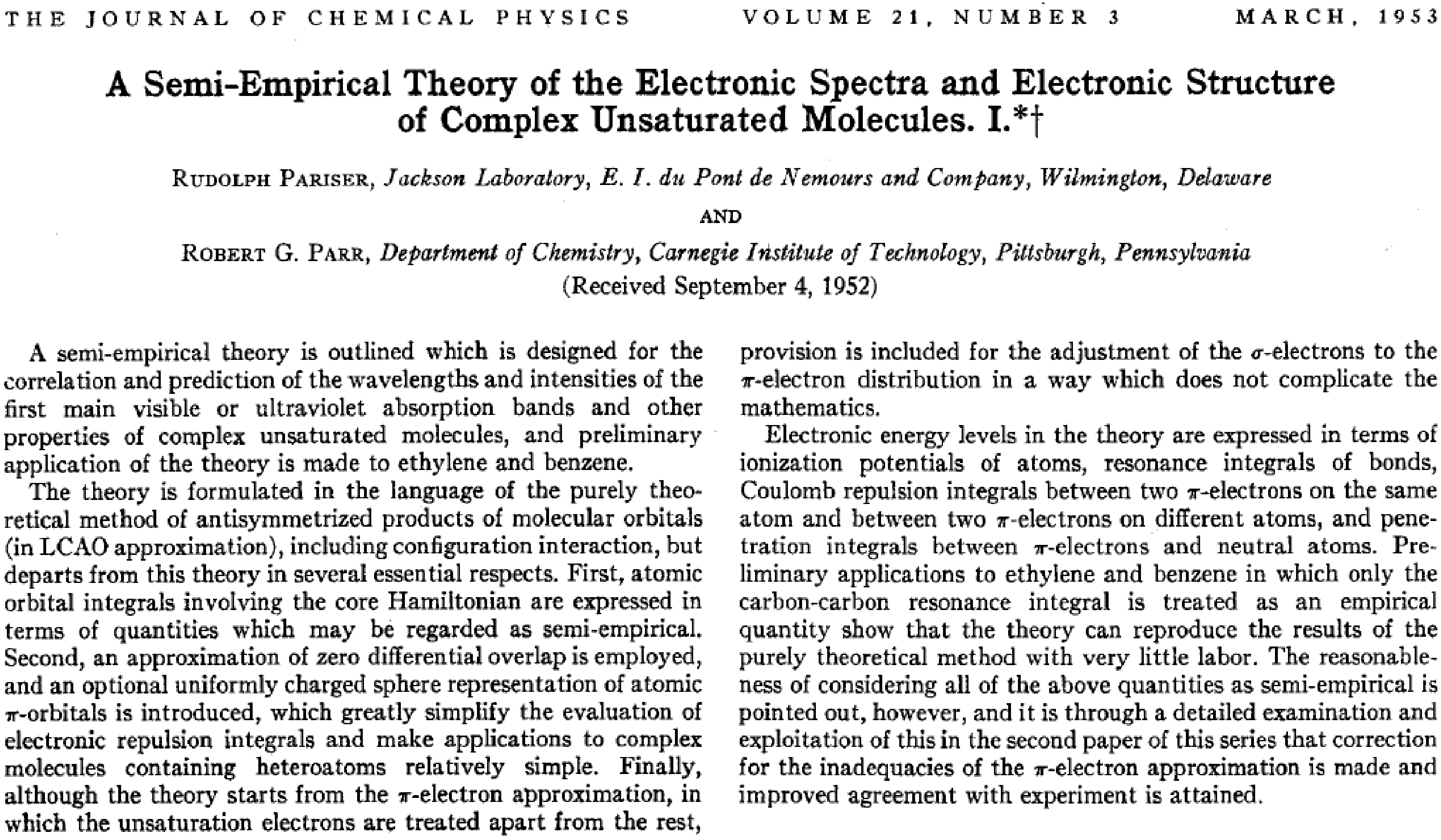
$$S_{\mu\nu} = \delta_{\mu\nu}$$

$$(\mu\nu \mid \lambda\sigma) = \int d\upsilon_1 \; \phi_\mu(1)\phi_\nu(1)\frac{1}{r_{12}}\phi_\lambda(2)\phi_\sigma(2) - \int d\upsilon_1 \; \phi_\mu(1)\phi_\lambda(1)\frac{1}{r_{12}}\phi_\nu(2)\phi_\sigma(2)$$

$$(\mu\nu \mid \lambda\sigma) = (\mu\mu \mid \lambda\lambda)\delta_{\mu\nu}\delta_{\lambda\sigma}$$

Simplified Roothaan-Hall equations:

$$F_{\mu\mu} = H_{\mu\mu}^{core} + \sum_{\lambda=1}^{K} P_{\lambda\lambda}(\mu\mu \mid \lambda\lambda) - \frac{1}{2}\sum_{\lambda=1}^{K} P_{\mu\mu}(\mu\mu \mid \mu\mu)$$



THE JOURNAL OF CHEMICAL PHYSICS VOLUME 21, NUMBER 3 MARCH, 1953

A Semi-Empirical Theory of the Electronic Spectra and Electronic Structure of Complex Unsaturated Molecules. I.*†

RUDOLPH PARISER, *Jackson Laboratory, E. I. du Pont de Nemours and Company, Wilmington, Delaware*
AND
ROBERT G. PARR, *Department of Chemistry, Carnegie Institute of Technology, Pittsburgh, Pennsylvania*

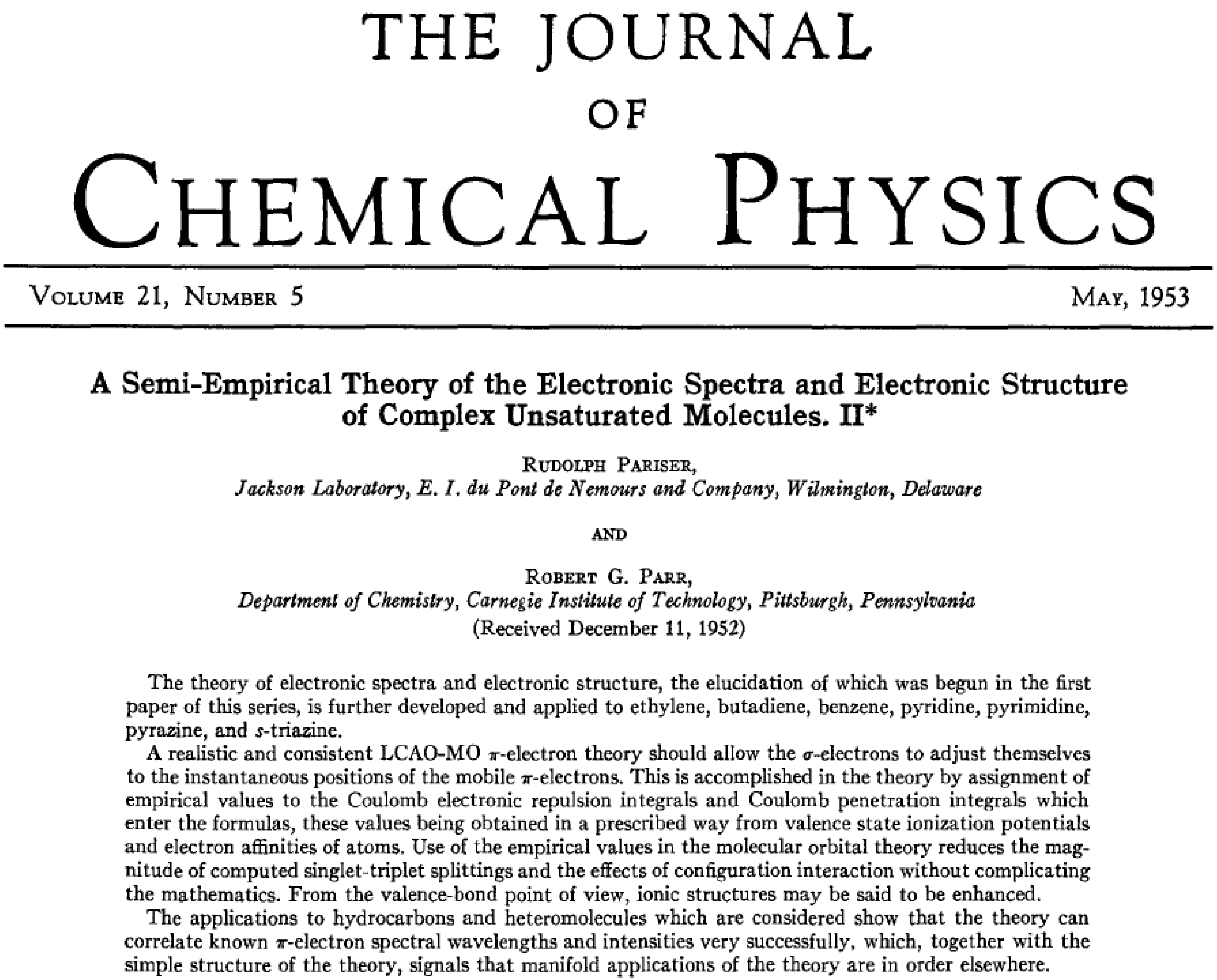
(Received September 4, 1952)

A semi-empirical theory is outlined which is designed for the correlation and prediction of the wavelengths and intensities of the first main visible or ultraviolet absorption bands and other properties of complex unsaturated molecules, and preliminary application of the theory is made to ethylene and benzene.

The theory is formulated in the language of the purely theoretical method of antisymmetrized products of molecular orbitals (in LCAO approximation), including configuration interaction, but departs from this theory in several essential respects. First, atomic orbital integrals involving the core Hamiltonian are expressed in terms of quantities which may be regarded as semi-empirical. Second, an approximation of zero differential overlap is employed, and an optional uniformly charged sphere representation of atomic π-orbitals is introduced, which greatly simplify the evaluation of electronic repulsion integrals and make applications to complex molecules containing heteroatoms relatively simple. Finally, although the theory starts from the π-electron approximation, in which the unsaturation electrons are treated apart from the rest, provision is included for the adjustment of the σ-electrons to the π-electron distribution in a way which does not complicate the mathematics.

Electronic energy levels in the theory are expressed in terms of ionization potentials of atoms, resonance integrals of bonds, Coulomb repulsion integrals between two π-electrons on the same atom and between two π-electrons on different atoms, and penetration integrals between π-electrons and neutral atoms. Preliminary applications to ethylene and benzene in which only the carbon-carbon resonance integral is treated as an empirical quantity show that the theory can reproduce the results of the purely theoretical method with very little labor. The reasonableness of considering all of the above quantities as semi-empirical is pointed out, however, and it is through a detailed examination and exploitation of this in the second paper of this series that correction for the inadequacies of the π-electron approximation is made and improved agreement with experiment is attained.

Parsier and Parr, *J. Chem. Phys.* **21**: 466-471 (1953)



THE JOURNAL
OF
CHEMICAL PHYSICS

VOLUME 21, NUMBER 5 MAY, 1953

A Semi-Empirical Theory of the Electronic Spectra and Electronic Structure of Complex Unsaturated Molecules. II*

RUDOLPH PARISER,
Jackson Laboratory, E. I. du Pont de Nemours and Company, Wilmington, Delaware

AND

ROBERT G. PARR,
Department of Chemistry, Carnegie Institute of Technology, Pittsburgh, Pennsylvania

(Received December 11, 1952)

The theory of electronic spectra and electronic structure, the elucidation of which was begun in the first paper of this series, is further developed and applied to ethylene, butadiene, benzene, pyridine, pyrimidine, pyrazine, and *s*-triazine.

A realistic and consistent LCAO-MO π-electron theory should allow the σ-electrons to adjust themselves to the instantaneous positions of the mobile π-electrons. This is accomplished in the theory by assignment of empirical values to the Coulomb electronic repulsion integrals and Coulomb penetration integrals which enter the formulas, these values being obtained in a prescribed way from valence state ionization potentials and electron affinities of atoms. Use of the empirical values in the molecular orbital theory reduces the magnitude of computed singlet-triplet splittings and the effects of configuration interaction without complicating the mathematics. From the valence-bond point of view, ionic structures may be said to be enhanced.

The applications to hydrocarbons and heteromolecules which are considered show that the theory can correlate known π-electron spectral wavelengths and intensities very successfully, which, together with the simple structure of the theory, signals that manifold applications of the theory are in order elsewhere.

Parsier and Parr, *J. Chem. Phys.* **21**: 767-776 (1953)

ELECTRON INTERACTION IN UNSATURATED HYDROCARBONS

By J. A. Pople
Dept. of Theoretical Chemistry, University of Cambridge

Received 24th July, 1953

An approximate form of the molecular orbital theory of unsaturated hydrocarbon molecules in their ground states is developed. The molecular orbital equations rigorously derived from the correct many-electron Hamiltonian are simplified by a series of systematic approximations and reduce to equations comparable with those used in the semi-empirical method based on an incompletely defined one-electron Hamiltonian. The two sets of equations differ, however, in that those of this paper include certain important terms representing electronic interaction. The theory is used to discuss the resonance energies, ionization potentials, charge densities, bond orders and bond lengths of some simple hydrocarbons. The electron interaction terms introduced in the theory are shown to play an important part in determining the ionization potentials. It is also shown that the uniform charge density theorem, proved by Coulson and Rushbrooke [1] for the simpler theory, holds also for the self-consistent orbitals derived by the method of this paper.

Complete Neglect of Differential Overlap (CNDO)

First practical method to implement ZDO

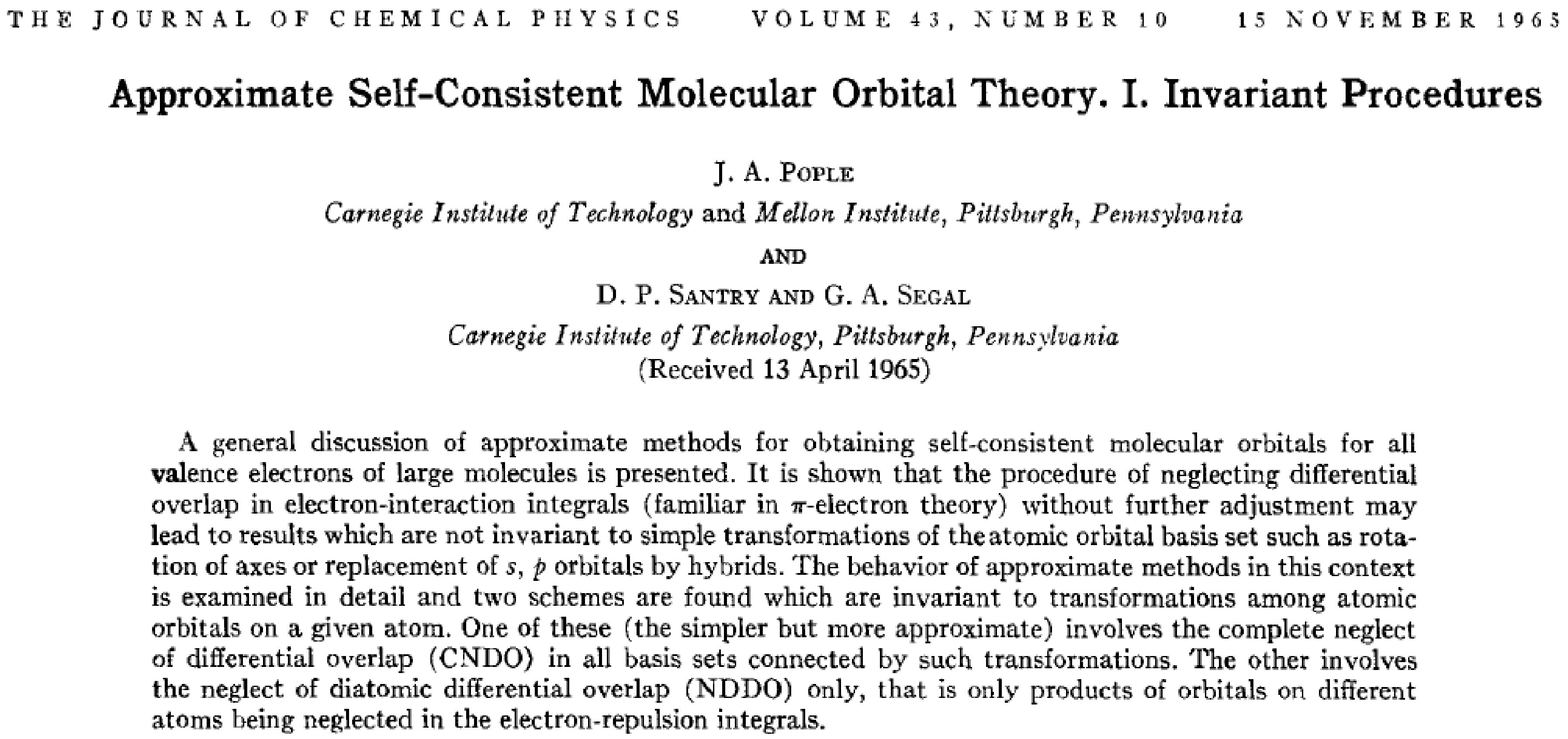
Pople, Santry and Segal, *J. Chem. Phys.* **43**: S129-S135 (1966)

Pople and Segal, *J. Chem. Phys.* **43**: S136-S149 (1966)

$$F_{\mu\mu} = H_{\mu\mu}^{core} + \sum_{\lambda=1;\lambda \text{ on A}}^{K} P_{\lambda\lambda}\gamma_{AA} - \frac{1}{2}P_{\mu\mu}\gamma_{AA} + \sum_{\lambda=1;\lambda \text{ not on A}}^{K} P_{\lambda\lambda}\gamma_{AB}$$

$$F_{\mu\nu} = H_{\mu\nu}^{core} - \frac{1}{2}P_{\mu\nu}\gamma_{AA};\ \mu \text{ and } \nu \text{ both on atom A}$$

$$F_{\mu\nu} = H_{\mu\nu}^{core} - \frac{1}{2}P_{\mu\nu}\gamma_{AB};\ \mu \text{ and } \nu \text{ on different atoms, A and B}$$



THE JOURNAL OF CHEMICAL PHYSICS VOLUME 43, NUMBER 10 15 NOVEMBER 1965

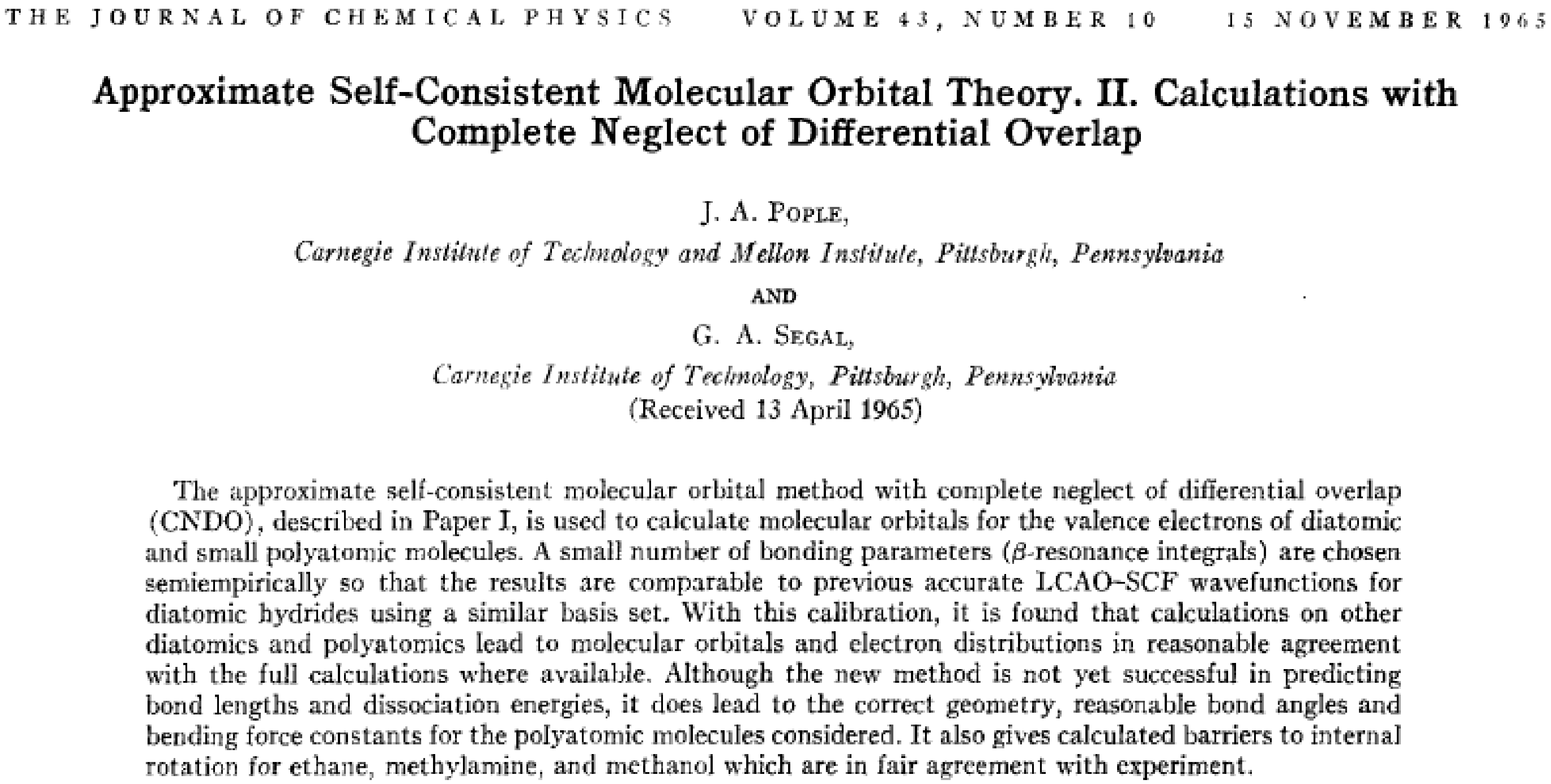
Approximate Self-Consistent Molecular Orbital Theory. I. Invariant Procedures

J. A. Pople
Carnegie Institute of Technology and *Mellon Institute, Pittsburgh, Pennsylvania*
AND
D. P. Santry and G. A. Segal
Carnegie Institute of Technology, Pittsburgh, Pennsylvania

(Received 13 April 1965)

A general discussion of approximate methods for obtaining self-consistent molecular orbitals for all valence electrons of large molecules is presented. It is shown that the procedure of neglecting differential overlap in electron-interaction integrals (familiar in π-electron theory) without further adjustment may lead to results which are not invariant to simple transformations of the atomic orbital basis set such as rotation of axes or replacement of s, p orbitals by hybrids. The behavior of approximate methods in this context is examined in detail and two schemes are found which are invariant to transformations among atomic orbitals on a given atom. One of these (the simpler but more approximate) involves the complete neglect of differential overlap (CNDO) in all basis sets connected by such transformations. The other involves the neglect of diatomic differential overlap (NDDO) only, that is only products of orbitals on different atoms being neglected in the electron-repulsion integrals.

Pople, Santry and Segal, *J. Chem. Phys.* **43**: S129-S135 (1966)



THE JOURNAL OF CHEMICAL PHYSICS VOLUME 43, NUMBER 10 15 NOVEMBER 1965

Approximate Self-Consistent Molecular Orbital Theory. II. Calculations with Complete Neglect of Differential Overlap

J. A. Pople,
Carnegie Institute of Technology and Mellon Institute, Pittsburgh, Pennsylvania
AND
G. A. Segal,
Carnegie Institute of Technology, Pittsburgh, Pennsylvania

(Received 13 April 1965)

The approximate self-consistent molecular orbital method with complete neglect of differential overlap (CNDO), described in Paper I, is used to calculate molecular orbitals for the valence electrons of diatomic and small polyatomic molecules. A small number of bonding parameters (β-resonance integrals) are chosen semiempirically so that the results are comparable to previous accurate LCAO–SCF wavefunctions for diatomic hydrides using a similar basis set. With this calibration, it is found that calculations on other diatomics and polyatomics lead to molecular orbitals and electron distributions in reasonable agreement with the full calculations where available. Although the new method is not yet successful in predicting bond lengths and dissociation energies, it does lead to the correct geometry, reasonable bond angles and bending force constants for the polyatomic molecules considered. It also gives calculated barriers to internal rotation for ethane, methylamine, and methanol which are in fair agreement with experiment.

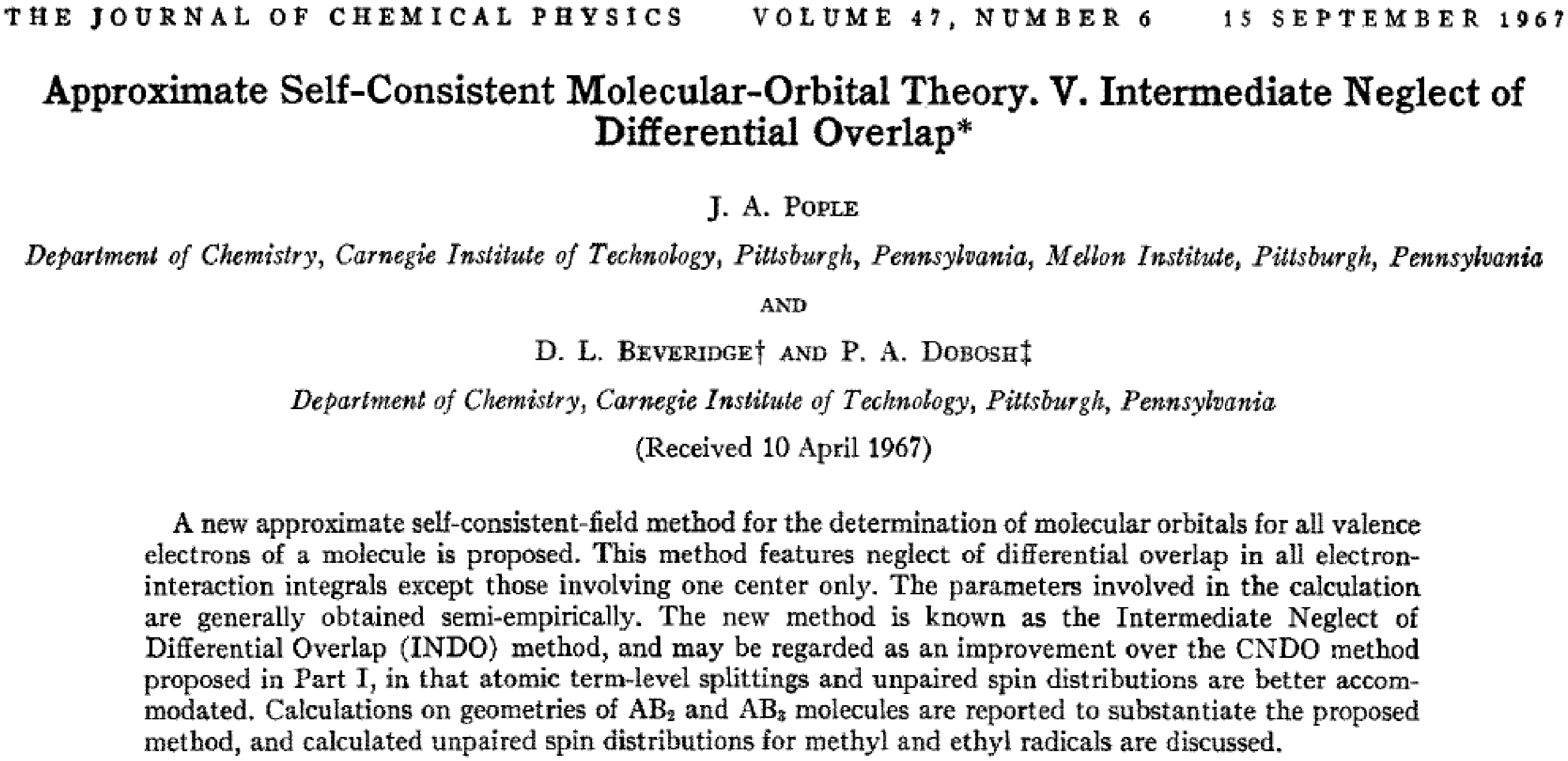
Pople, and Segal, *J. Chem. Phys.* **43**: S136-S149 (1966)

Intermediate Neglect of Differential Overlap (INDO)

Pople, Beveridge and Dobosh, *J. Chem. Phys.* **47**: 2026-2033 (1967)

$$F_{\mu\mu} = U_{\mu\mu} + \sum_{\lambda \text{ on A}} \sum_{\sigma \text{ on A}} \left[P_{\lambda\sigma}(\mu\mu|\lambda\sigma) - P^{\alpha}_{\lambda\sigma}(\mu\lambda|\mu\sigma) \right] + \sum_{B \neq A} \left(P_{BB} - Z_B \right) \gamma_{AB}$$

$$F_{\mu\nu} = U_{\mu\nu} + \sum_{\lambda \text{ on A}} \sum_{\sigma \text{ on A}} \left[P_{\lambda\sigma}(\mu\nu|\lambda\sigma) - P^{\alpha}_{\lambda\sigma}(\mu\lambda|\nu\sigma) \right]$$



THE JOURNAL OF CHEMICAL PHYSICS VOLUME 47, NUMBER 6 15 SEPTEMBER 1967

Approximate Self-Consistent Molecular-Orbital Theory. V. Intermediate Neglect of Differential Overlap*

J. A. POPLE

Department of Chemistry, Carnegie Institute of Technology, Pittsburgh, Pennsylvania, Mellon Institute, Pittsburgh, Pennsylvania

AND

D. L. BEVERIDGE† AND P. A. DOBOSH‡

Department of Chemistry, Carnegie Institute of Technology, Pittsburgh, Pennsylvania

(Received 10 April 1967)

A new approximate self-consistent-field method for the determination of molecular orbitals for all valence electrons of a molecule is proposed. This method features neglect of differential overlap in all electron-interaction integrals except those involving one center only. The parameters involved in the calculation are generally obtained semi-empirically. The new method is known as the Intermediate Neglect of Differential Overlap (INDO) method, and may be regarded as an improvement over the CNDO method proposed in Part I, in that atomic term-level splittings and unpaired spin distributions are better accommodated. Calculations on geometries of AB_2 and AB_3 molecules are reported to substantiate the proposed method, and calculated unpaired spin distributions for methyl and ethyl radicals are discussed.

Pople, Beveridge and Dobosh, *J. Chem. Phys.* **47**: 2026-2033 (1967)

Modified INDO (MINDO/3)

Bingham, Dewar, Lo, *J. Am. Chem. Soc.* **97**: 1285-1293, 1294-1301, 1302-1306, 1307-1310 (1975)

$$F_{\mu\mu} = U_{\mu\mu} + \sum_{\lambda \text{ on A}} \sum_{\sigma \text{ on A}} \left[P_{\nu\nu}(\mu\mu|\nu\nu) - \frac{1}{2} P_{\nu\nu}(\mu\nu|\mu\nu) \right] + \sum_{B \neq A} (P_{BB} - Z_B)\gamma_{AB}$$

$$F_{\mu\nu} = -\frac{1}{2} P_{\mu\nu}(\mu\nu|\mu\nu); \quad \mu \text{ and } \nu \text{ both on A}$$

$$F_{\mu\nu} = H_{\mu\nu}^{core} - \frac{1}{2} P_{\mu\nu}(\mu\nu|\mu\nu) = H_{\mu\nu}^{core} - \frac{1}{2} P_{\mu\nu}\gamma_{AB}; \quad \mu \text{ on A and } \nu \text{ on B}$$

Neglect of Diatomic Differential Overlap (NDDO)

Pople, Santry and Segal, *J. Chem. Phys.* **43**: S129-S135 (1966)

Pople, Beveridge and Dobosh, *J. Chem. Phys.* **47**: 2026-2033 (1967)

$$F_{\mu\mu} = H_{\mu\mu}^{core} + \sum_{\lambda \text{ on A}} \sum_{\sigma \text{ on A}} \left[P_{\lambda\sigma}(\mu\mu|\lambda\sigma) - \frac{1}{2} P_{\lambda\sigma}(\mu\lambda|\mu\sigma) \right] + \sum_{B \neq A} \sum_{\lambda \text{ on B}} \sum_{\sigma \text{ on B}} P_{\lambda\sigma}(\mu\mu|\lambda\sigma)$$

$$F_{\mu\nu} = H_{\mu\nu}^{core} + \sum_{\lambda \text{ on A}} \sum_{\sigma \text{ on A}} \left[P_{\lambda\sigma}(\mu\nu|\lambda\sigma) - \frac{1}{2} P_{\lambda\sigma}(\mu\lambda|\nu\sigma) \right] + \sum_{B \neq A} \sum_{\lambda \text{ on B}} \sum_{\sigma \text{ on B}} P_{\lambda\sigma}(\mu\nu|\lambda\sigma); \quad \mu \text{ and } \nu \text{ both on A}$$

$$F_{\mu\nu} = H_{\mu\nu}^{core} - \frac{1}{2} \sum_{\lambda \text{ on B}} \sum_{\sigma \text{ on A}} P_{\lambda\sigma}(\mu\sigma|\lambda\sigma); \quad \mu \text{ on A and } \nu \text{ on B}$$

Modified Neglect of Diatomic Overlap (MNDO)

Dewar and Thiel, *J. Am. Chem. Soc.* **99**: 4899-4907, 4907-4917 (1977)

$$F_{\mu\mu} = H^{core}_{\mu\mu} + \sum_{\nu \text{ on A}} \left[P_{\lambda\lambda}(\mu\mu|\lambda\lambda) - \frac{1}{2} P_{\nu\nu}(\mu\nu|\mu\nu) \right]$$

$$+ \sum_{B\neq A} \sum_{\lambda \text{ on B}} \sum_{\sigma \text{ on B}} P_{\lambda\sigma}(\mu\mu|\lambda\sigma)$$

$$\text{where } H^{core}_{\mu\mu} = U_{\mu\mu} - \sum_{B\neq A} V_{\mu\mu B}$$

$$V_{\mu\mu B} = -Z_B(\mu_A\mu_A|s_Bs_B)$$

$$V_{\mu\nu B} = -Z_B(\mu_A\nu_A|s_Bs_B)$$

$$F_{\mu\nu} = H^{core}_{\mu\mu} + \frac{3}{2} P_{\lambda\lambda}(\mu\mu|\lambda\lambda) - \frac{1}{2} P_{\mu\nu}(\mu\mu|\nu\nu)$$

$$+ \sum_{B\neq A} \sum_{\lambda \text{ on B}} \sum_{\sigma \text{ on B}} P_{\lambda\sigma}(\mu\nu|\lambda\sigma)$$

$$\text{where } H^{core}_{\mu\mu} = -\sum_{B\neq A} V_{\mu\nu B}$$

$$F_{\mu\nu} = H^{core}_{\mu\mu} - \frac{1}{2} \sum_{\lambda \text{ on B}} \sum_{\sigma \text{ on A}} P_{\lambda\sigma}(\mu\sigma|\nu\lambda)$$

$$\text{where } H^{core}_{\mu\mu} = -\frac{1}{2} S_{\mu\nu}\left(\beta_\mu + \beta_\nu\right)$$

$$E_{AB} = Z_A Z_B (s_A s_A | s_B s_B)\left[1 + e^{-\alpha_A R_{AB}} + e^{-\alpha_B R_{AB}}\right]$$

$$E_{XH} = Z_X Z_H (s_X s_X | s_H s_H)\left[1 + R_{XH} e^{-\alpha_X R_{XH}} + e^{-\alpha_H R_{XH}}\right]$$

Ground States of Molecules. 38.[1] The MNDO Method. Approximations and Parameters

Michael J. S. Dewar* and Walter Thiel

Contribution from the Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712. *Received October 27, 1976*

Abstract: The basic approximations of the MNDO (modified neglect of diatomic overlap) method are described including a semiempirical model for the two-center repulsion integrals. Parametric functions for the various terms in the MNDO Fock matrix are then chosen which contain atomic parameters only (no bond parameters). Using a nonlinear least-squares iterative optimization technique, numerical values of the parameters are determined for the elements H, C, N, O. Finally, the main differences between the MNDO and MINDO methods are discussed.

Ground States of Molecules. 39. MNDO Results for Molecules Containing Hydrogen, Carbon, Nitrogen, and Oxygen

Michael J. S. Dewar* and Walter Thiel

Contribution from the Department of Chemistry, The University of Texas at Austin, Austin, Texas 78741. *Received October 27, 1976*

Abstract: Heats of formation, molecular geometries, ionization potentials, and dipole moments are calculated by the MNDO method for a large number of molecules. The MNDO results are compared with the corresponding MINDO/3 results on a statistical basis. For the properties investigated, the mean absolute errors in MNDO are uniformly smaller than those in MINDO/3 by a factor of about 2. Major improvements of MNDO over MINDO/3 are found for heats of formation of unsaturated systems and molecules with NN bonds, for bond angles, for higher ionization potentials, and for dipole moments of compounds with heteroatoms.

Austin Model 1 (AM1)
(Second Parameterization of MNDO)

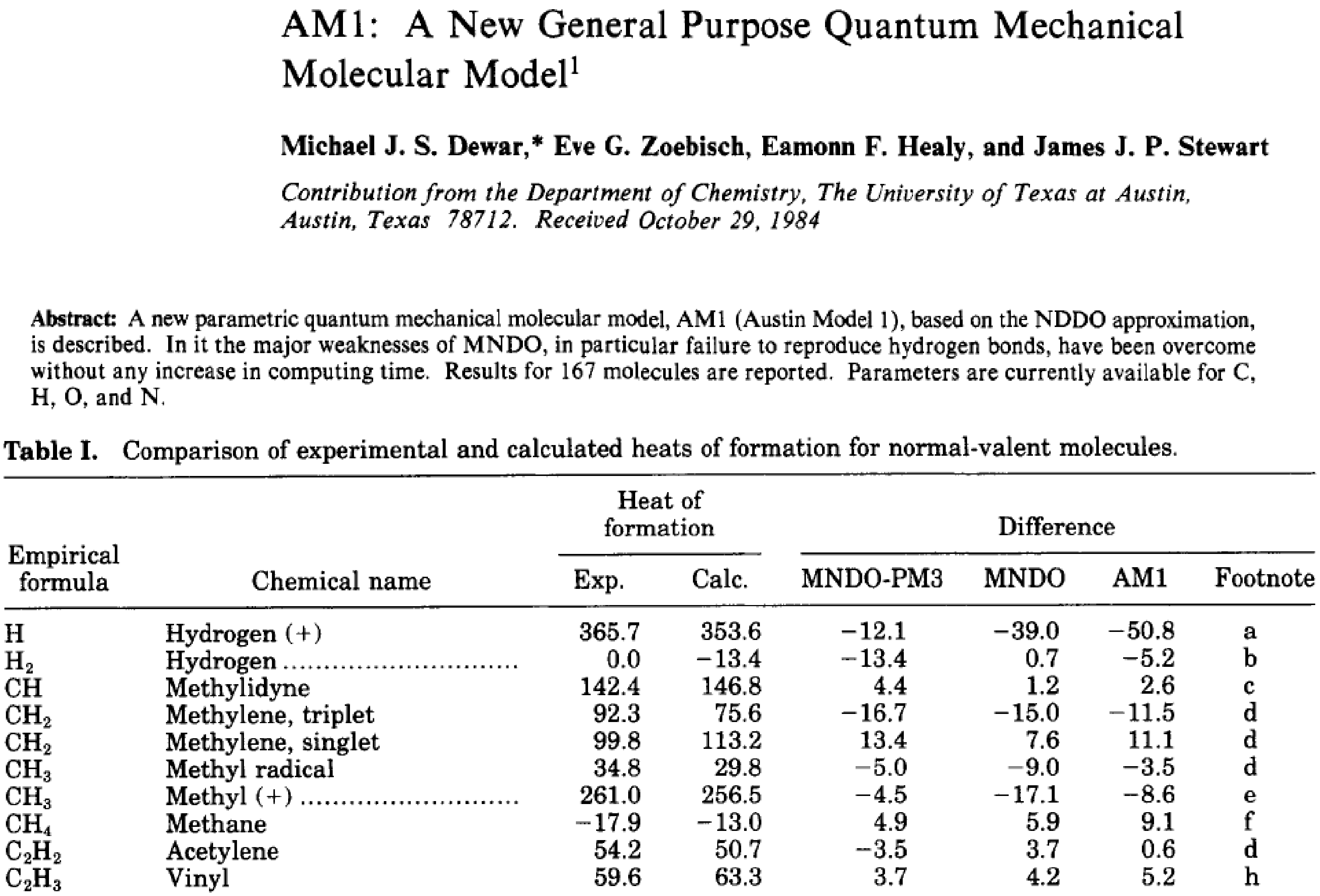
Dewar *et al.*, *J. Am. Chem. Soc.* **107**: 3902-3909 (1985)

Problems with MNDO: overestimate repulsions between atoms separated by distances approximately equal to the sum of their van der Waals radii

Remedy: modify the core-core term using attractive and repulsive Gaussian functions. The attractive Gaussians were designed to overcome the repulsion directly and were centered in the region where the repulsions were too large. The repulsive Gaussians were centered at smaller internuclear separations.

$$E_{AB} = E_{MNDO} + \frac{Z_A Z_B}{R_{AB}} \left\{ \sum_i K_{A_i} e^{-L_{A_i}\left(R_{AB} - M_{A_i}\right)^2} + \sum_j K_{B_j} e^{-L_{B_i}\left(R_{AB} - M_{B_j}\right)^2} \right\}$$

Times cited: 15273 (2018/3/15)



AM1: A New General Purpose Quantum Mechanical Molecular Model[1]

Michael J. S. Dewar,* Eve G. Zoebisch, Eamonn F. Healy, and James J. P. Stewart

Contribution from the Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712. *Received October 29, 1984*

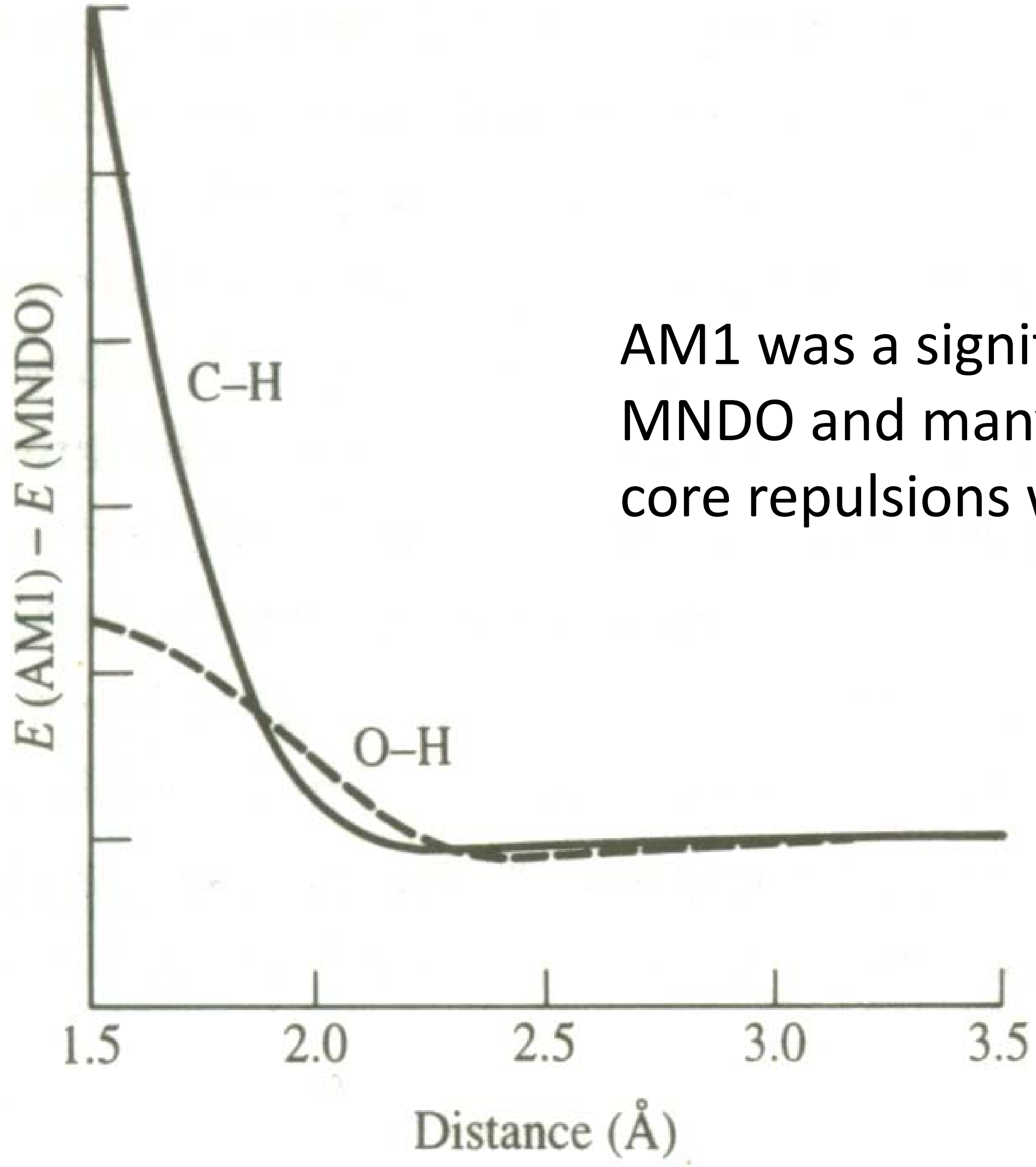
Abstract: A new parametric quantum mechanical molecular model, AM1 (Austin Model 1), based on the NDDO approximation, is described. In it the major weaknesses of MNDO, in particular failure to reproduce hydrogen bonds, have been overcome without any increase in computing time. Results for 167 molecules are reported. Parameters are currently available for C, H, O, and N.

Table I. Comparison of experimental and calculated heats of formation for normal-valent molecules.

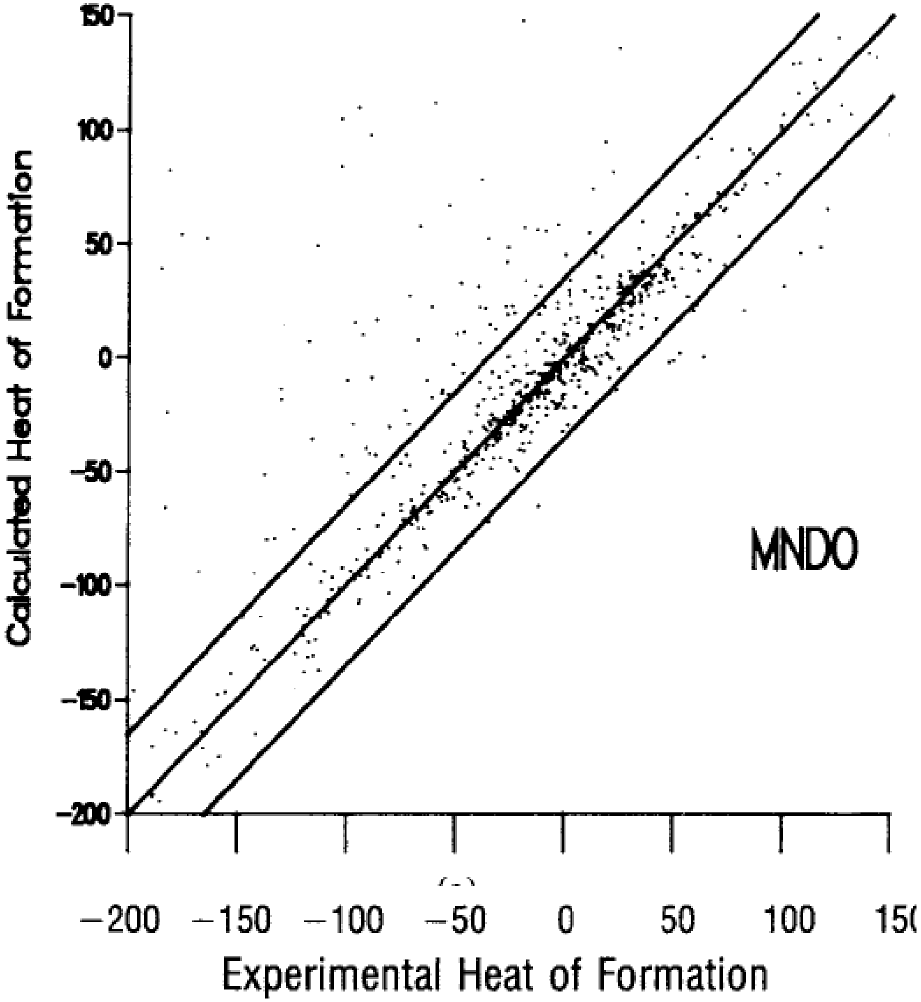
Empirical formula	Chemical name	Heat of formation Exp.	Heat of formation Calc.	Difference MNDO-PM3	Difference MNDO	Difference AM1	Footnote
H	Hydrogen (+)	365.7	353.6	−12.1	−39.0	−50.8	a
H_2	Hydrogen	0.0	−13.4	−13.4	0.7	−5.2	b
CH	Methylidyne	142.4	146.8	4.4	1.2	2.6	c
CH_2	Methylene, triplet	92.3	75.6	−16.7	−15.0	−11.5	d
CH_2	Methylene, singlet	99.8	113.2	13.4	7.6	11.1	d
CH_3	Methyl radical	34.8	29.8	−5.0	−9.0	−3.5	d
CH_3	Methyl (+)	261.0	256.5	−4.5	−17.1	−8.6	e
CH_4	Methane	−17.9	−13.0	4.9	5.9	9.1	f
C_2H_2	Acetylene	54.2	50.7	−3.5	3.7	0.6	d
C_2H_3	Vinyl	59.6	63.3	3.7	4.2	5.2	h

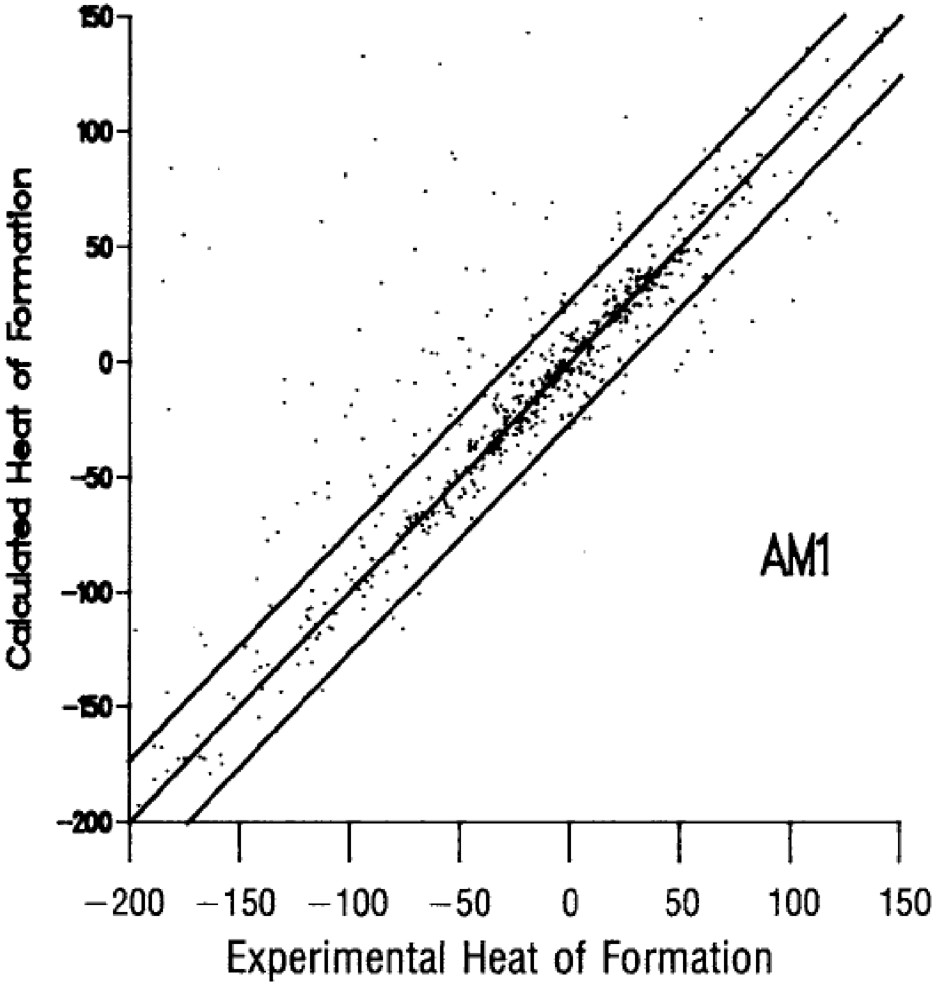
Dewar *et al.*, *J. Am. Chem. Soc.* **107**: 3902-3909 (1985)

Comparing MNDO and AM1



AM1 was a significant improvement over MNDO and many of the deficiencies with core repulsions were corrected.

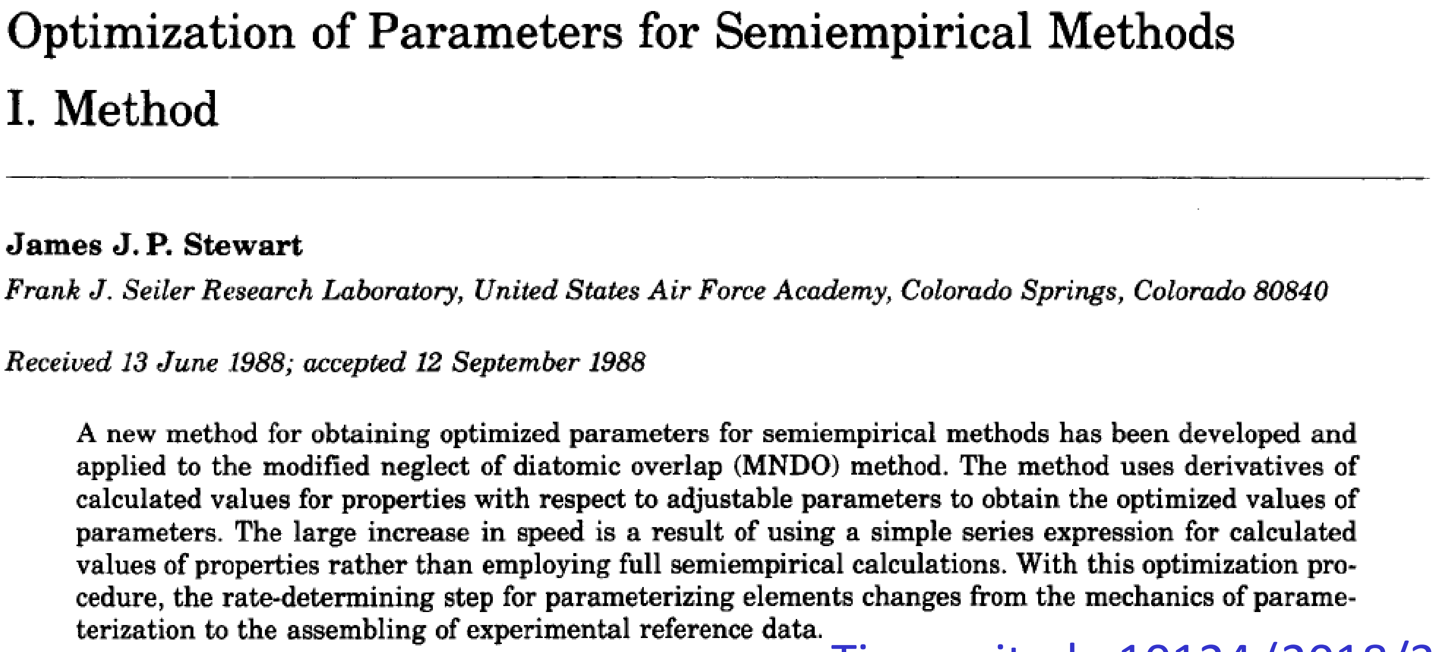

MNDO
Calculated Heat of Formation
150
100
50
0
-50
-100
-150
-200
Experimental Heat of Formation
−200
−150
−100
−50
0
50
100
150


AM1
Calculated Heat of Formation
150
100
50
0
-50
-100
-150
-200
Experimental Heat of Formation
−200
−150
−100
−50
0
50
100
150

Third Parameterization of MNDO (PM3)

Stewart, *J. Comput. Chem.* **10**: 209-220, 221-264 (1989)

- The inclusion of Gaussians in AM1 significantly increased the number of parameters per atom, from 7 in the MNDO to between 13 and 16 per atom in AM1, which made the parameterization process considerably more difficult. Many of parameters in AM1 were obtained by applying chemical knowledge and ‘intuition’.
- The PM3 Hamiltonian contains essentially the same elements as that for AM1 but the parameters for the PM3 were derived using an automatic parameterization procedure.
- PM3 may underestimate the rotational barrier of the amid bond.
- There is still considerable debate over the relative merits of the AM1 and PM3 approaches.
- J.J.P. Stewart was one of the coauthors of AM1.



Optimization of Parameters for Semiempirical Methods

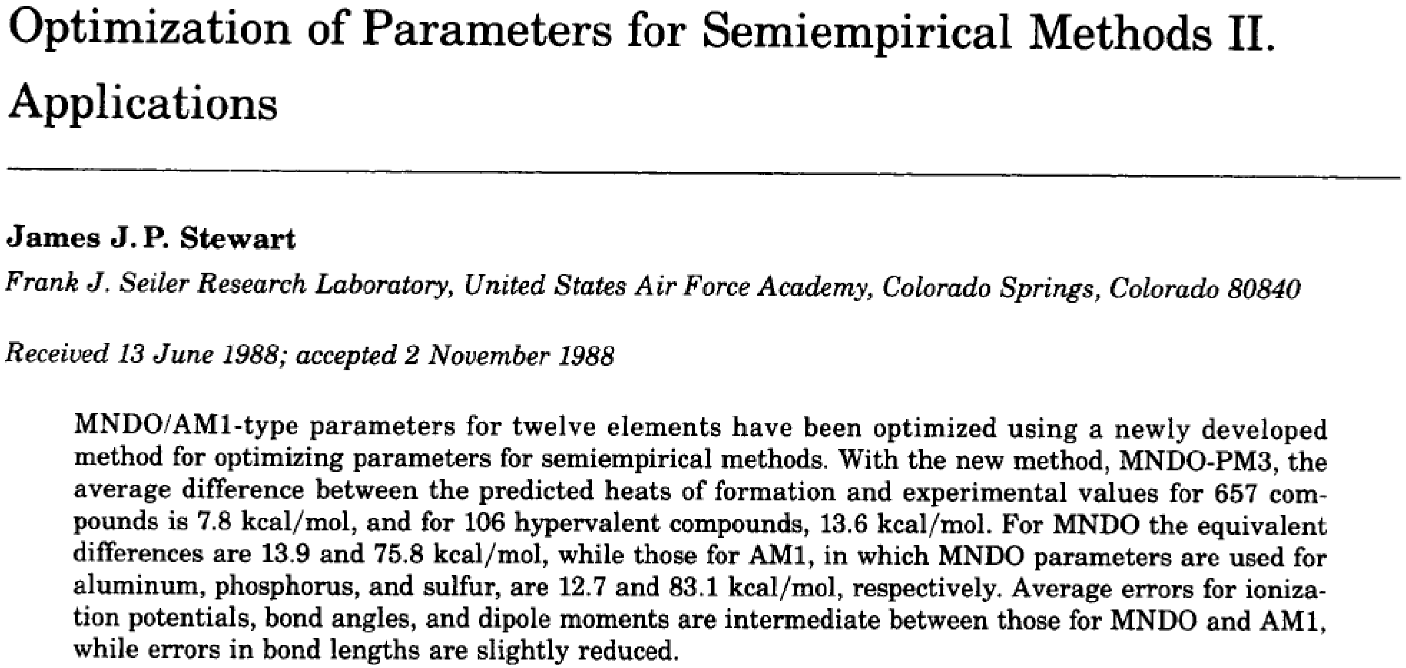
I. Method

James J. P. Stewart
Frank J. Seiler Research Laboratory, United States Air Force Academy, Colorado Springs, Colorado 80840

Received 13 June 1988; accepted 12 September 1988

A new method for obtaining optimized parameters for semiempirical methods has been developed and applied to the modified neglect of diatomic overlap (MNDO) method. The method uses derivatives of calculated values for properties with respect to adjustable parameters to obtain the optimized values of parameters. The large increase in speed is a result of using a simple series expression for calculated values of properties rather than employing full semiempirical calculations. With this optimization procedure, the rate-determining step for parameterizing elements changes from the mechanics of parameterization to the assembling of experimental reference data.

Times cited: 10124 (2018/3/15)



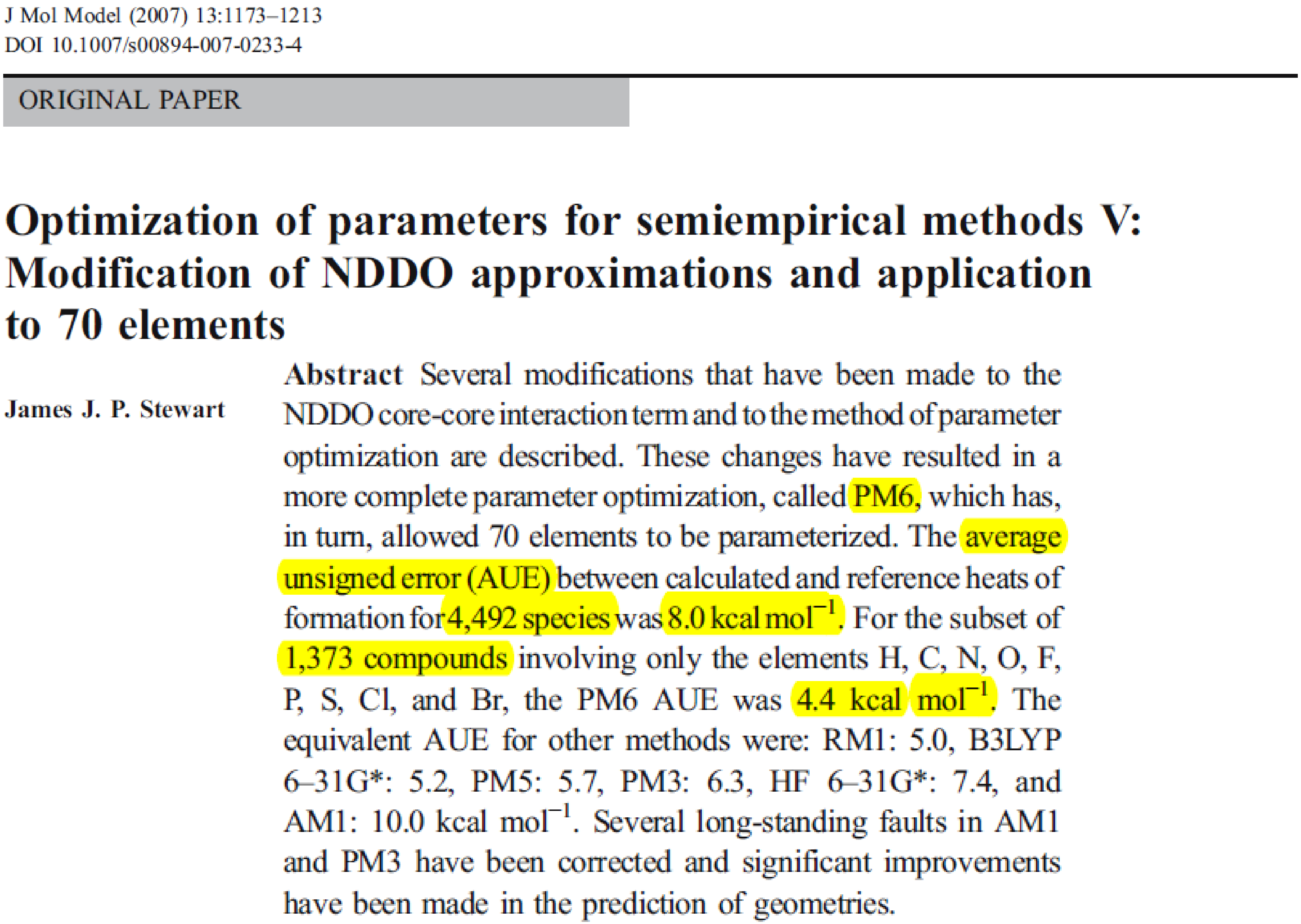
Optimization of Parameters for Semiempirical Methods II. Applications

James J. P. Stewart
Frank J. Seiler Research Laboratory, United States Air Force Academy, Colorado Springs, Colorado 80840

Received 13 June 1988; accepted 2 November 1988

MNDO/AM1-type parameters for twelve elements have been optimized using a newly developed method for optimizing parameters for semiempirical methods. With the new method, MNDO-PM3, the average difference between the predicted heats of formation and experimental values for 657 compounds is 7.8 kcal/mol, and for 106 hypervalent compounds, 13.6 kcal/mol. For MNDO the equivalent differences are 13.9 and 75.8 kcal/mol, while those for AM1, in which MNDO parameters are used for aluminum, phosphorus, and sulfur, are 12.7 and 83.1 kcal/mol, respectively. Average errors for ionization potentials, bond angles, and dipole moments are intermediate between those for MNDO and AM1, while errors in bond lengths are slightly reduced.

Times cited: 10120 (2018/3/15)



J Mol Model (2007) 13:1173–1213
DOI 10.1007/s00894-007-0233-4

ORIGINAL PAPER

Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements

James J. P. Stewart

Abstract Several modifications that have been made to the NDDO core-core interaction term and to the method of parameter optimization are described. These changes have resulted in a more complete parameter optimization, called PM6, which has, in turn, allowed 70 elements to be parameterized. The average unsigned error (AUE) between calculated and reference heats of formation for 4,492 species was 8.0 kcal mol^{-1}. For the subset of 1,373 compounds involving only the elements H, C, N, O, F, P, S, Cl, and Br, the PM6 AUE was 4.4 kcal mol^{-1}. The equivalent AUE for other methods were: RM1: 5.0, B3LYP 6–31G*: 5.2, PM5: 5.7, PM3: 6.3, HF 6–31G*: 7.4, and AM1: 10.0 kcal mol^{-1}. Several long-standing faults in AM1 and PM3 have been corrected and significant improvements have been made in the prediction of geometries.

Times cited: 2303 (2018/3/15)

Reference data

In contrast to earlier methods, in which reference data was assembled by painstakingly searching the original literature, the current work relies heavily on the large compendia of data that have been developed in recent years. The most important of these are the WebBook [20], for thermochemistry, and the Cambridge Structural Database [21] (CSD), for molecular geometries.

During the early stages of the current work, consistency checks were performed to ensure that erroneous data were not used. These checks revealed many cases in which the calculated heats of formation were inconsistent with the reference heats of formation reported in the NIST database. On further checking, many of these reference data were also found [22, 23] to be inconsistent with other data in the WebBook. In those cases where there was strong evidence of error in the reference data, the offending data were deleted, and the webbook updated [24].

For molecular geometries, gas phase reference data are preferred, but in many instances such data were unavailable, and recourse was made to condensed-phase data. Provided that care was taken to exclude those species whose geometries were likely to be significantly distorted by crystal forces, or which carried a large formal charge, condensed-phase data of the type found in the CSD were regarded as being suitable as reference data.

Because earlier methods used only a limited number of reference data, most of the cases where the method gave bad results were not discovered until after the method was published. In an attempt to minimize the occurrence of such unpleasant surprises, the set of reference data used was made as large as practical. To this end, where there was a dearth or even a complete absence of experimental reference data, recourse was made to high level calculations. Thus, for the Group VIII elements, there are relatively few stable compounds, and the main phenomena of interest involve rare gas atoms colliding with other atoms or molecules, so reference data representing the mechanics of rare gas atoms colliding with other atoms was generated from the results of ab-initio calculations. Additionally, there is an almost complete lack of thermochemical data for many types of complexes involving transition metals, so augmenting what little data there was with the results of ab-initio calculations was essential.

J Mol Model (2008) 14:499–535
DOI 10.1007/s00894-008-0299-7

OpenAccess

ORIGINAL PAPER

Application of the PM6 method to modeling the solid state

James J. P. Stewart

J Mol Model (2009) 15:765–805
DOI 10.1007/s00894-008-0420-y

ORIGINAL PAPER

Application of the PM6 method to modeling proteins

James J. P. Stewart

J Mol Model (2013) 19:1–32
DOI 10.1007/s00894-012-1667-x

ORIGINAL PAPER

Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters

James J. P. Stewart

Abstract Modern semiempirical methods are of sufficient accuracy when used in the modeling of molecules of the same type as used as reference data in the parameterization. Outside that subset, however, there is an abundance of evidence that these methods are of very limited utility. In an attempt to expand the range of applicability, a new method called PM7 has been developed. PM7 was parameterized using experimental and high-level ab initio reference data, augmented by a new type of reference data intended to better define the structure of parameter space. The resulting method was tested by modeling crystal structures and heats of formation of solids. Two changes were made to the set of approximations: a modification was made to improve the description of noncovalent interactions, and two minor errors in the NDDO formalism were rectified. Average unsigned errors (AUEs) in geometry and ΔH_f for PM7 were reduced relative to PM6; for simple gas-phase organic systems, the AUE in bond lengths decreased by about 5 % and the AUE in ΔH_f decreased by about 10 %; for organic solids, the AUE in ΔH_f dropped by 60 % and the reduction was 33.3 % for geometries. A two-step process (PM7-TS) for calculating the heights of activation barriers has been developed. Using PM7-TS, the AUE in the barrier heights for simple organic reactions was decreased from values of 12.6 kcal/mol^{-1} in PM6 and 10.8 kcal/mol^{-1} in PM7 to 3.8 kcal/mol^{-1}. The origins of the errors in NDDO methods have been examined, and were found to be attributable to inadequate and inaccurate reference data. This conclusion provides insight into how these methods can be improved.

Molecular Spectroscopy

- Spectroscopy is the measurement of absorption or emission intensity as a function of wavelength (or frequency).
- Spectroscopic studies were central to the development of quantum mechanics and included Max Planck's explanation of blackbody radiation, Albert Einstein's explanation of the photoelectric effect and Niels Bohr's explanation of atomic structure and spectra.
- Spectroscopy is used in physical and analytical chemistry because atoms and molecules have unique spectra.

Absorbance of Lights

Lambert's law:

$$-\frac{dI}{I} = b\,dx$$

$$\int \frac{dI}{I} = -b\int_0^l dx$$

the proportion of radiation absorbed by a substance is independent of the intensity of the incident radiation, i.e., each successive layer of the thickness *dx* of the medium absorbs an equal fraction *–dI/I* of the radiant intensity incident upon it. Thus, the relative loss of intensity is proportional to the thickness (*dx*).

$$\ln I = -bl + g$$

I=I

$$g = \ln I_0$$

$$\ln I = -bl + \ln I_0$$

$$\log_{10}\frac{I_0}{I} = \frac{bl}{2.303} = A$$

$$T \equiv \frac{I}{I_0}$$

T: transmittance

$$\log_{10}\frac{1}{T} = A$$

$$T\% = 100T = \frac{100I}{I_0}$$

A: *decadic absorbance*
or **absorbance**

$$\log_{10} T\% = \log_{10} 100 + \log_{10} T$$

$$\log_{10} T\% = 2 - A$$

The laws of Lambert and Beer

Beer studied the influence of the concentration of a substance in solution on the absorbance, and he found the same linear relationship between absorbance and concentration as Lambert had found between absorbance and thickness.

Beer's law:

$$\log_{10} \frac{I_0}{I} = A = \text{constant} \times c$$

Lambert-Beer law:

$$A = \log_{10} \frac{I_0}{I} = \varepsilon c l$$

Absorption coefficient

$$\varepsilon_l^c(\lambda)$$

Molar absorption coefficient

$$\varepsilon_{1cm}^{1M}(430\text{ nm})$$

Molecular Spectra

- Pure Rotational Spectra (microwave and far-infrared regions)
- Vibrational-Rotational Spectra (also called ro-vibrational, near-infrared region)
- Electronic Spectra (visible and ultraviolet regions)

Moment of Inertia Tensor and the Magnitude of Angular Momentum Vector

$$\mathbf{I}=\begin{pmatrix} \sum_i m_i\left(y_i^2+z_i^2\right) & -\sum_i m_i x_i y_i & -\sum_i m_i x_i z_i \\ -\sum_i m_i x_i y_i & \sum_i m_i\left(x_i^2+z_i^2\right) & -\sum_i m_i y_i z_i \\ -\sum_i m_i x_i z_i & -\sum_i m_i y_i z_i & \sum_i m_i\left(x_i^2+y_i^2\right) \end{pmatrix}$$

$$\left|\mathbf{J}\right| = \sqrt{J(J+1)}\hbar$$

$$E_J = J\left(J+1\right)\frac{\hbar^2}{2I}$$

$$J = 0,1,2,3,\ldots$$

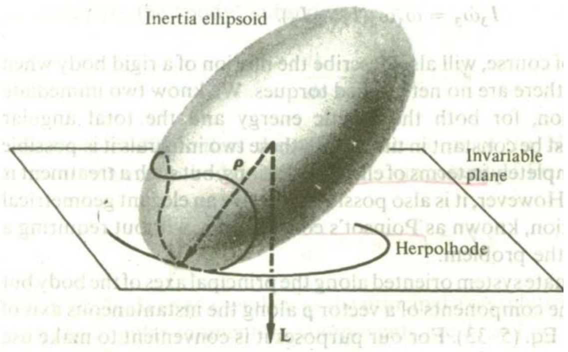


FIGURE 5–4
The motion of the inertia ellipsoid relative to the invariable plane.

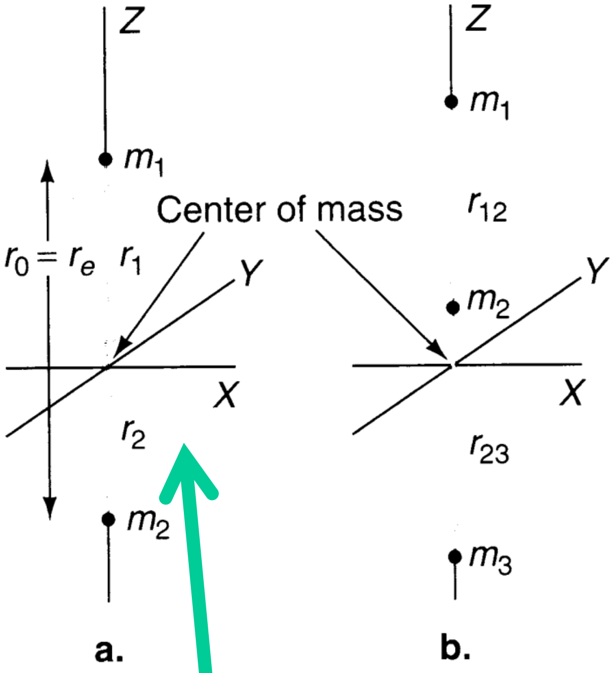


FIGURE 13.15
(a) A diatomic molecule. (b) A linear triatomic molecule. The center of mass depends on the relative masses.

$$I = m_1 r_1^2 + m_2 r_2^2$$

Pure Rotational Spectra

$$\Delta E = E_{J+1} - E_J = [(J+1)(J+2) - J(J+1)]\frac{\hbar^2}{2I}$$

$$= 2(J+1)\frac{\hbar^2}{2I} = 2(J+1)\frac{h^2}{8\pi^2 I}$$

frequencies

$$\nu_j = \frac{\Delta E}{h} = 2(J+1)\frac{h}{8\pi^2 I}$$

$$\tilde{\nu}_j = \nu_j / c$$

$$\longrightarrow$$

wavenumbers

$$\tilde{\nu}_j = 2(J+1)\frac{h}{8\pi^2 I\, c}$$

$$\tilde{\nu}_j = 2(J+1)\tilde{B}$$

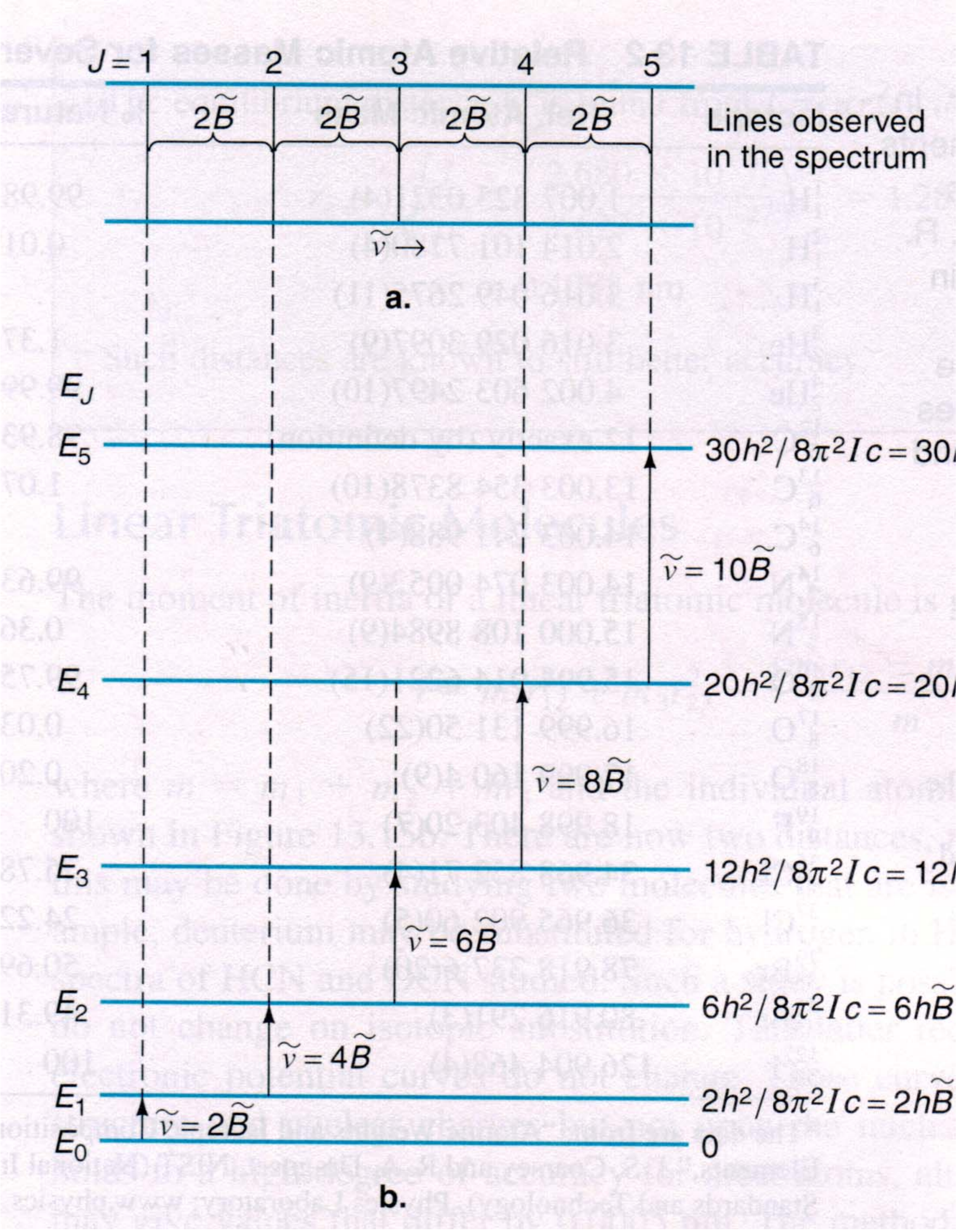
$$\tilde{B} = \frac{h}{8\pi^2 I\, c}$$

rotational constant

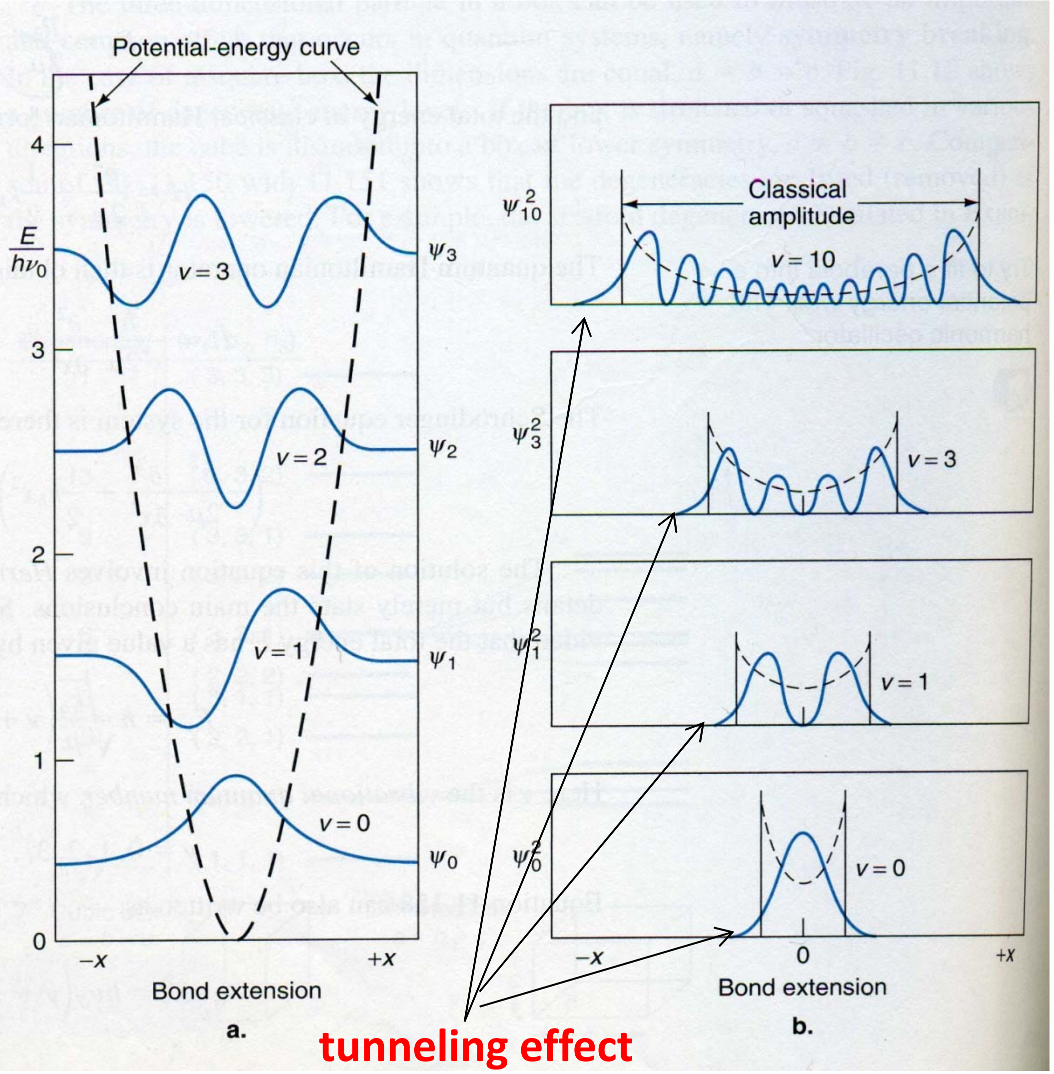
$$\tilde{\nu}_j = 2(J+1)\tilde{B}$$

$$E_J = J\left(J+1\right)\frac{\hbar^2}{2I}$$

$$J = 0, 1, 2, 3, \ldots$$



Energy Levels of a Harmonic Oscillator



tunneling effect

$$V(x) = \frac{1}{2}kx^2$$

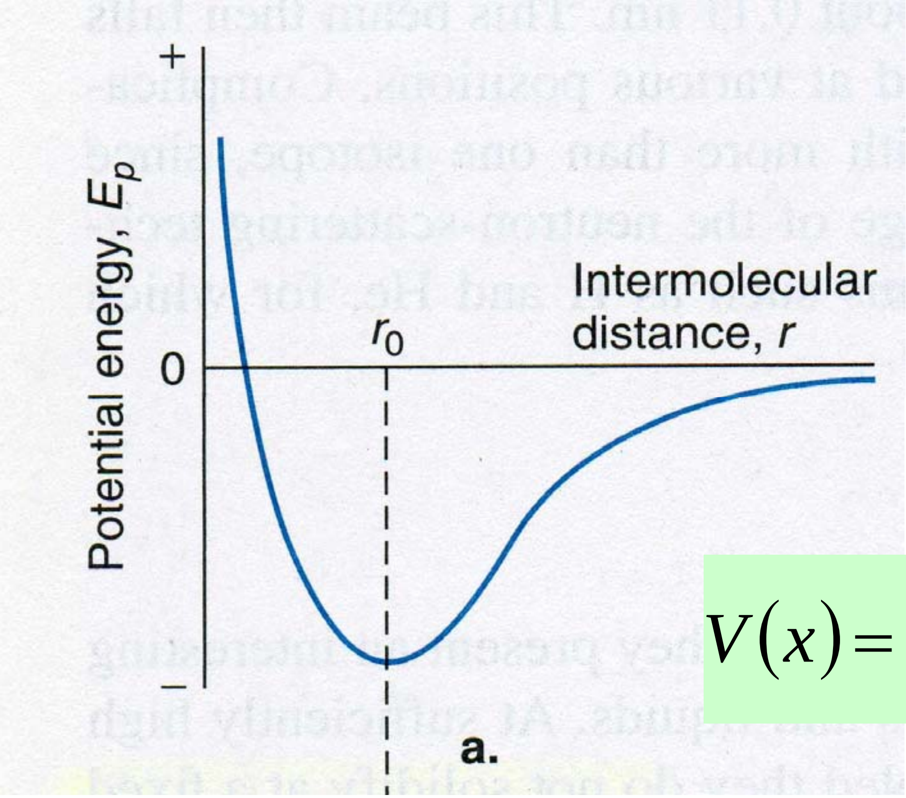
$$E_v = \hbar\sqrt{\frac{k}{\mu}}\left(v + \frac{1}{2}\right)$$

$$v = 0,1,2,3,\ldots$$

$$E_0 = \frac{1}{2}\hbar\sqrt{\frac{k}{\mu}} = \frac{\hbar\nu_0}{2}$$

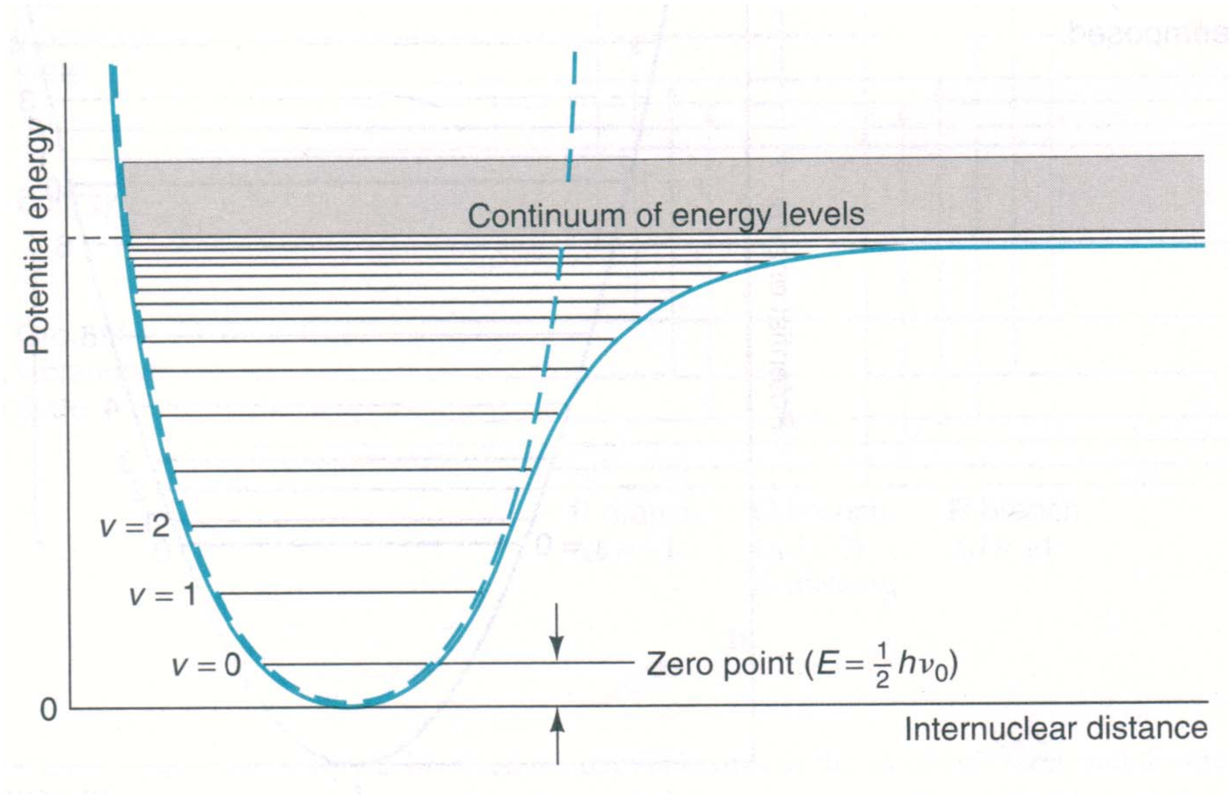
Zero point energy

Potential energy between two atoms and its Taylor expansion



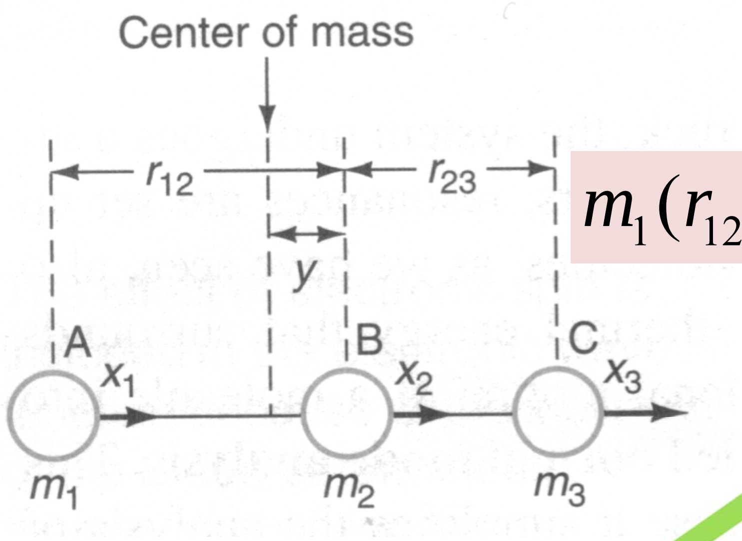
$$V(x)=V(x_0)+\frac{1}{2}(x-x_0)^2V''(x_0)+\ldots$$

Effect of deviation from Parabola



As the potential energy curve deviates from the parabola, the levels become closer together.

Coupled Vibrational Motion



$$m_1(r_{12}-y)=m_2y+m_3(r_{23}+y)$$

$$m_1(r_{12}-y-x_1)=m_2(y+x_2)+m_3(r_{23}+y+x_3)$$

$$m_1x_1+m_2x_2+m_3x_3=0$$

$$k_{12}(x_2-x_1)=m_1\ddot{x}_1$$

$$k_{23}(x_3-x_2)=-m_3\ddot{x}_3$$

$$m_1m_2\ddot{x}_1+k_{12}[(m_1+m_2)x_1+m_3x_3]=0$$

$$m_2m_3\ddot{x}_3+k_{23}[m_1x_1+(m_2+m_3)x_3]=0$$

Coupled Differential Equations

If we look for solutions of the following form:

$$x_1 = A_1 \cos 2\pi\nu t \qquad \ddot{x}_1 = -4\pi^2\nu^2 x_1 = -\lambda x_1$$

$$x_3 = A_3 \cos 2\pi\nu t \qquad \ddot{x}_3 = -4\pi^2\nu^2 x_3 = -\lambda x_3$$

where **the frequency is the same**.

$$[-\lambda m_1 m_2 + k_{12}(m_1 + m_2)]x_1 + k_{12} m_3 x_3 = 0$$

$$m_1 k_{23} x_1 + [-\lambda m_2 m_3 + k_{23}(m_2 + m_3)]x_3 = 0$$

$$\begin{vmatrix} -\lambda m_1 m_2 + k_{12}(m_1 + m_2) & k_{12} m_3 \\ m_1 k_{23} & -\lambda m_2 m_3 + k_{23}(m_2 + m_3) \end{vmatrix} = 0$$

Normal Modes of Vibration

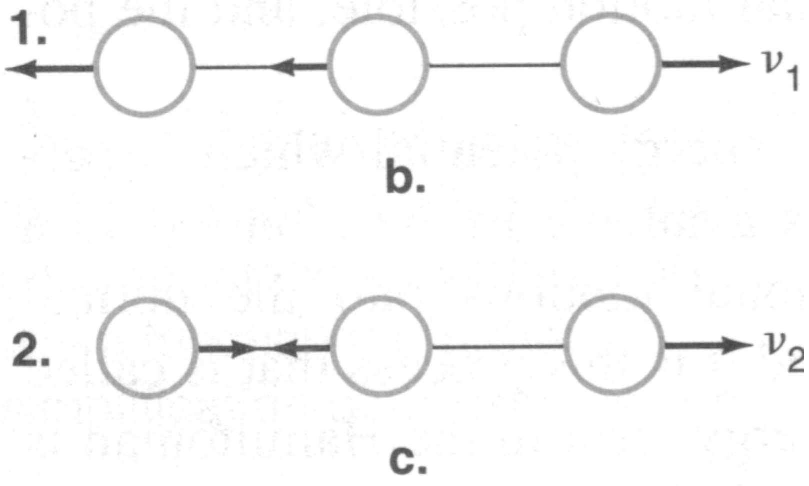
$$\lambda^2 - \underbrace{\boxed{\left(\frac{m_1+m_2}{m_1 m_2}k_{12} + \frac{m_2+m_3}{m_2 m_3}k_{23}\right)}}_{b}\lambda + \underbrace{\boxed{\frac{k_{12}k_{23}(m_1+m_2+m_3)}{m_1 m_2 m_3}}}_{c} = 0$$

$$\lambda^2 - b\lambda + c = 0$$

$$\lambda_1 = \frac{b - \sqrt{b^2 - 4c}}{2} \qquad \lambda_2 = \frac{b + \sqrt{b^2 - 4c}}{2}$$

$$\nu_1 = \frac{\sqrt{\lambda_1}}{2\pi} \quad \leq \quad \nu_2 = \frac{\sqrt{\lambda_2}}{2\pi}$$

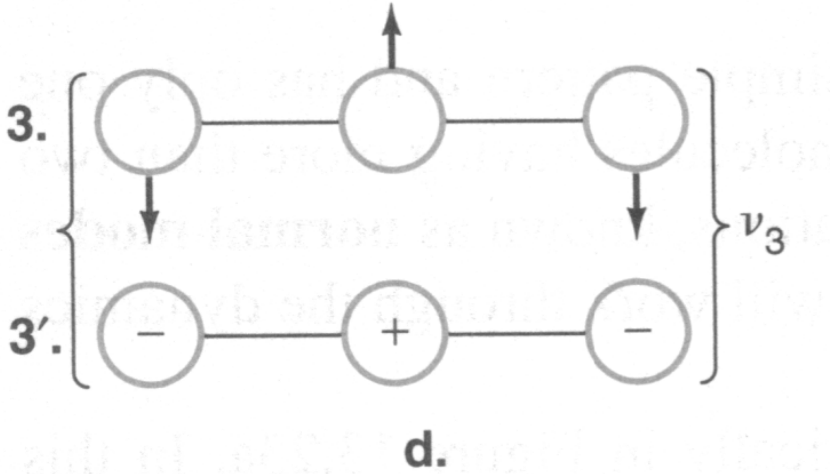
Symmetric stretching (breathing) vibration



Anti-symmetric stretching vibration

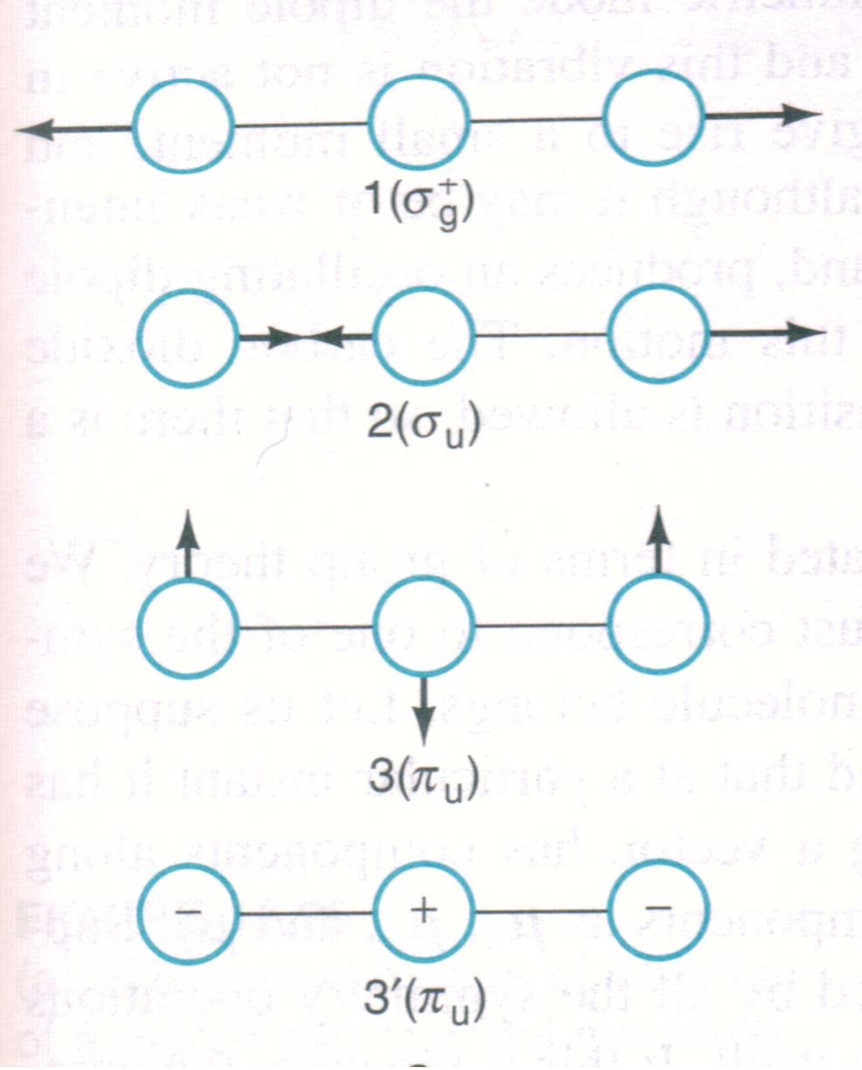
Bending Vibration

Similar to the previous treatment for stretching vibrations, we can have two normal modes for bending vibrations.



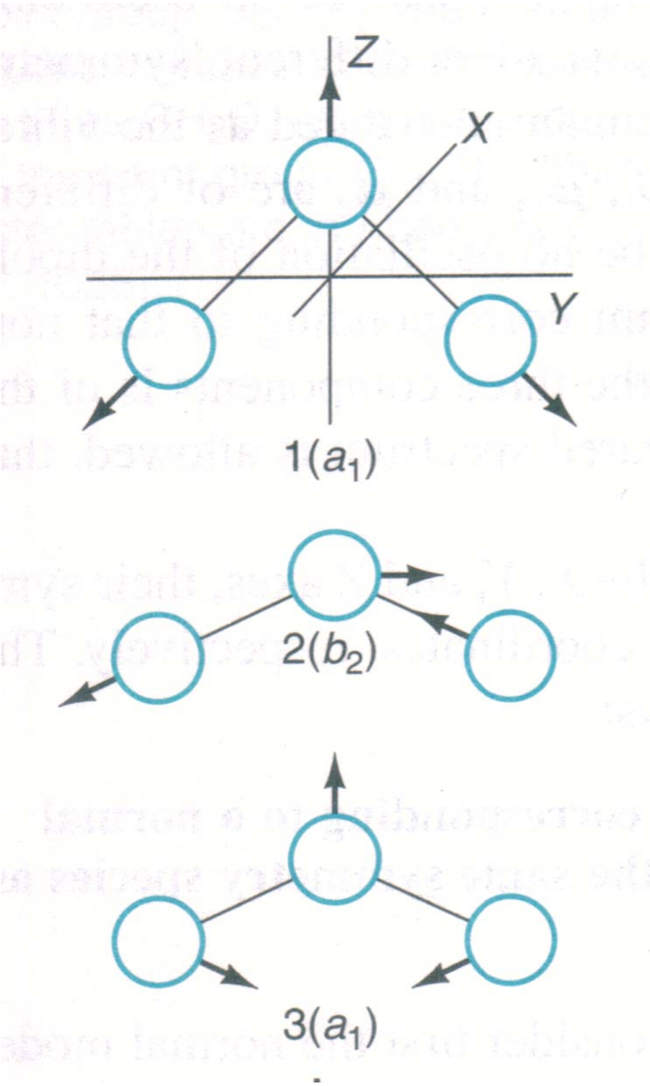
A bending motion in any plane cab be treated as a linear combination of the bending motions in these two planes.

Normal modes of vibration of CO_2



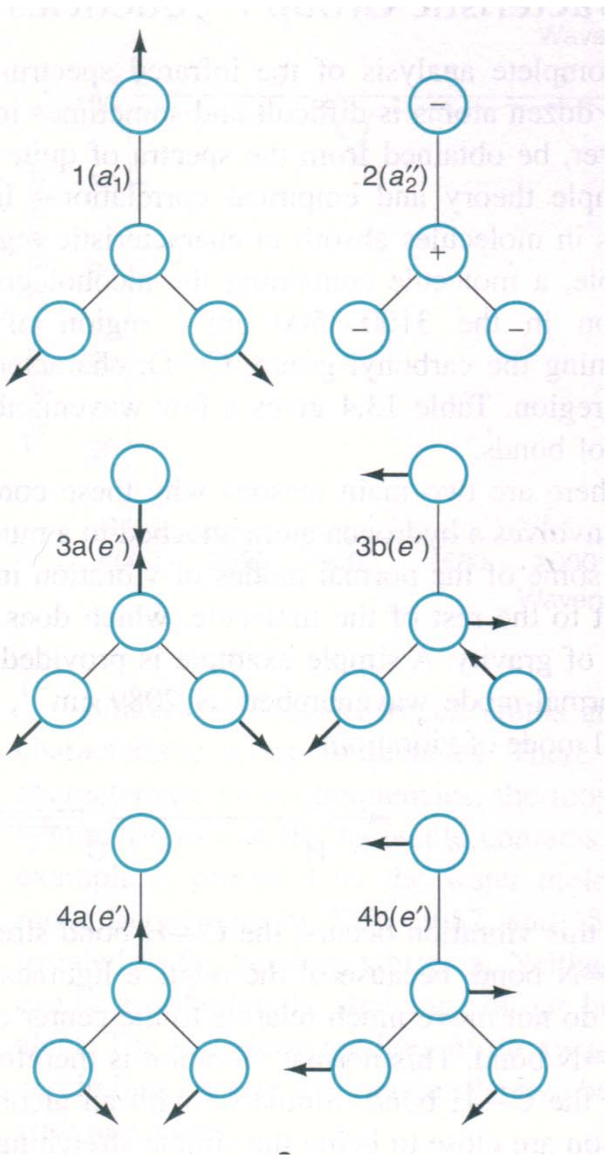
D_{h}

Normal modes of vibration of H_2O



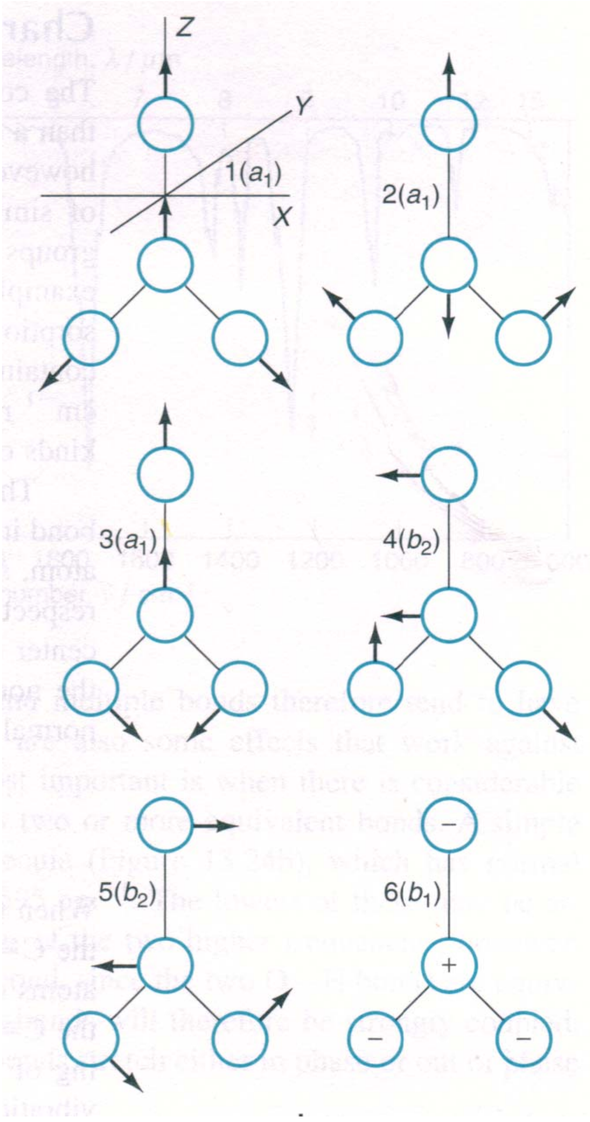
C_{2v}

Normal modes of vibration of BF_3



D_{3h}

Normal modes of vibration of CH_2O



C_{2v}

Infrared Spectra of Complex Molecules

A molecular can give a vibrational-rotational spectrum only if its normal mode vibrations can give rise to **oscillating dipole moments.**

There will be a vibrational-rotational spectrum corresponding to a normal mode of a vibration only if that mode belongs to the same symmetry species as one or more of three coordinates x, y, and z.

Infrared versus Ramen Spectra

- For a transition to be **Raman active** there must be a **change in the polarizability** of the molecule during the vibration. This means that the **electron cloud of the molecule** must undergo positional change.
- On the other hand, for an IR detectable transition, the molecule must undergo **dipole moment change** during vibration. So, when a molecule is symmetrical, e.g. O_2, we cannot observe any IR absorption lines, since the molecule cannot change its dipole moment.
- It has been observed that **molecules with a strong dipole moment are typically hard to polarize**.

Symmetrical stretch

Asymmetrical stretch

In plane bending

Out of plane bending

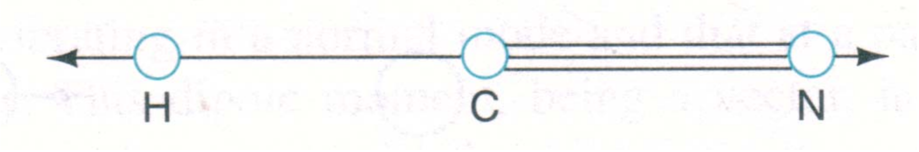
No change in dipole moment
therefore IR inactive
There is change in polarisability
therefore Raman active

There is change in
dipole moment
therefore IR active
but Raman inactive

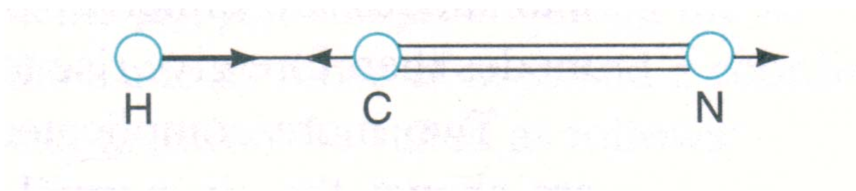
The deformation vibrations of CO_2 are degenerate and appear at the same region (666 cm^{-1}) in the IR spectrum of CO_2. There is no change in polarisability therefore these vibrations are Raman inactive.

	Raman	IR
1.	It is due to the **scattering of light** by the vibrating molecules.	It is the result of **absorption of light** by vibrating molecules.
2.	The vibration is Raman active if it causes a **change in polarizability**.	Vibration is IR active if there is **change in dipole moment**.
3.	The molecule needs not possess a permanent dipole moment.	The vibration concerned should have a change in dipole moment due to that vibration.
4.	Water **can** be used as a solvent.	Water **cannot** be used due to its intense absorption of IR.
5.	Sample preparation is **not very elaborate**, it can be in any state.	Sample preparation is **elaborate** Gaseous samples can rarely be used.
6.	Gives an indication of **covalent character** in the molecule.	Gives an indication of **ionic character** in the molecule.
7.	Cost of instrumentation is very high	Comparatively inexpensive.

Characteristic Group Frequencies



symmetric vibration, 2089 cm^{-1}



antisymmetric vibration, 3312 cm^{-1}

Factors favoring determination of characteristic group frequencies: **1.** a bond involving a hydrogen atom (or a light atom); **2:** the presence of multiple bonds

Factors disfavoring determination of characteristic group frequencies: **symmetry**

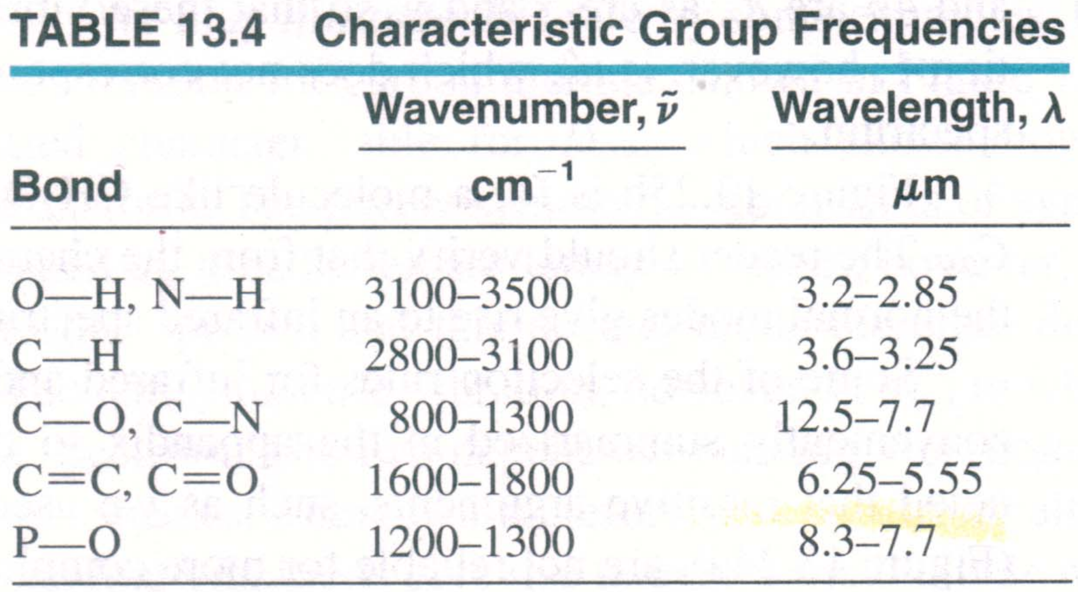


TABLE 13.4 Characteristic Group Frequencies

Bond	Wavenumber, $\tilde{\nu}$ cm^{-1}	Wavelength, λ μm
O—H, N—H	3100–3500	3.2–2.85
C—H	2800–3100	3.6–3.25
C—O, C—N	800–1300	12.5–7.7
C=C, C=O	1600–1800	6.25–5.55
P—O	1200–1300	8.3–7.7

Infrared spectrum of methyl acetate in the liquid phase

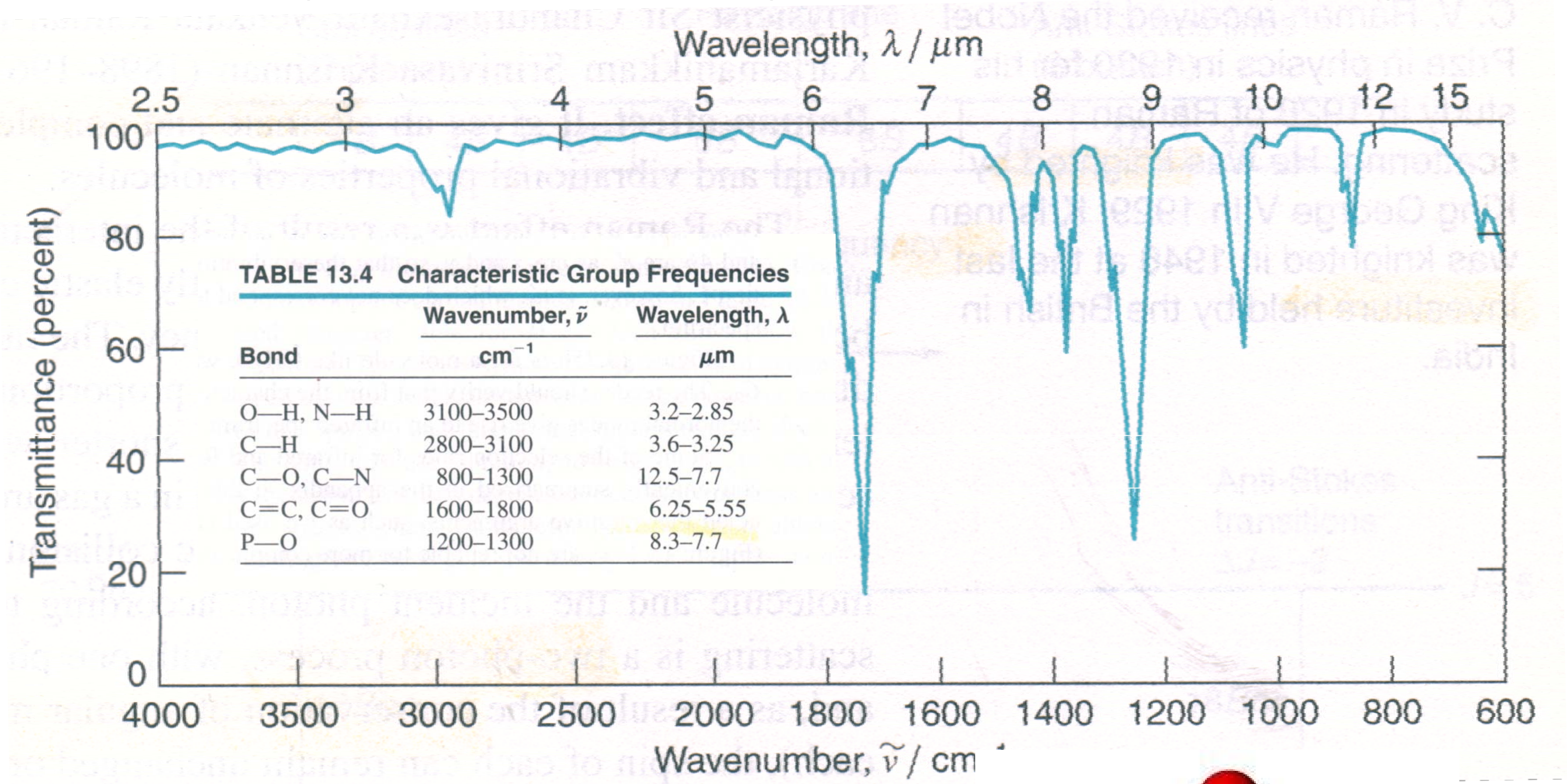
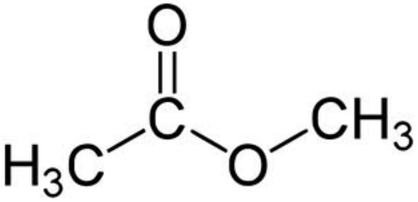
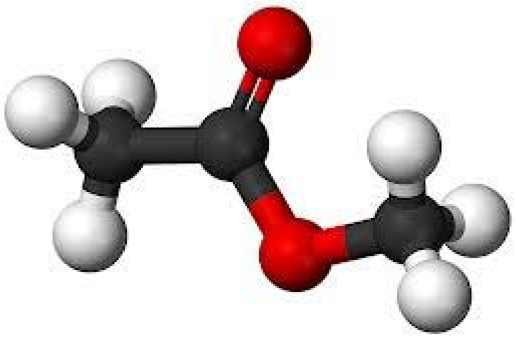


TABLE 13.4 Characteristic Group Frequencies

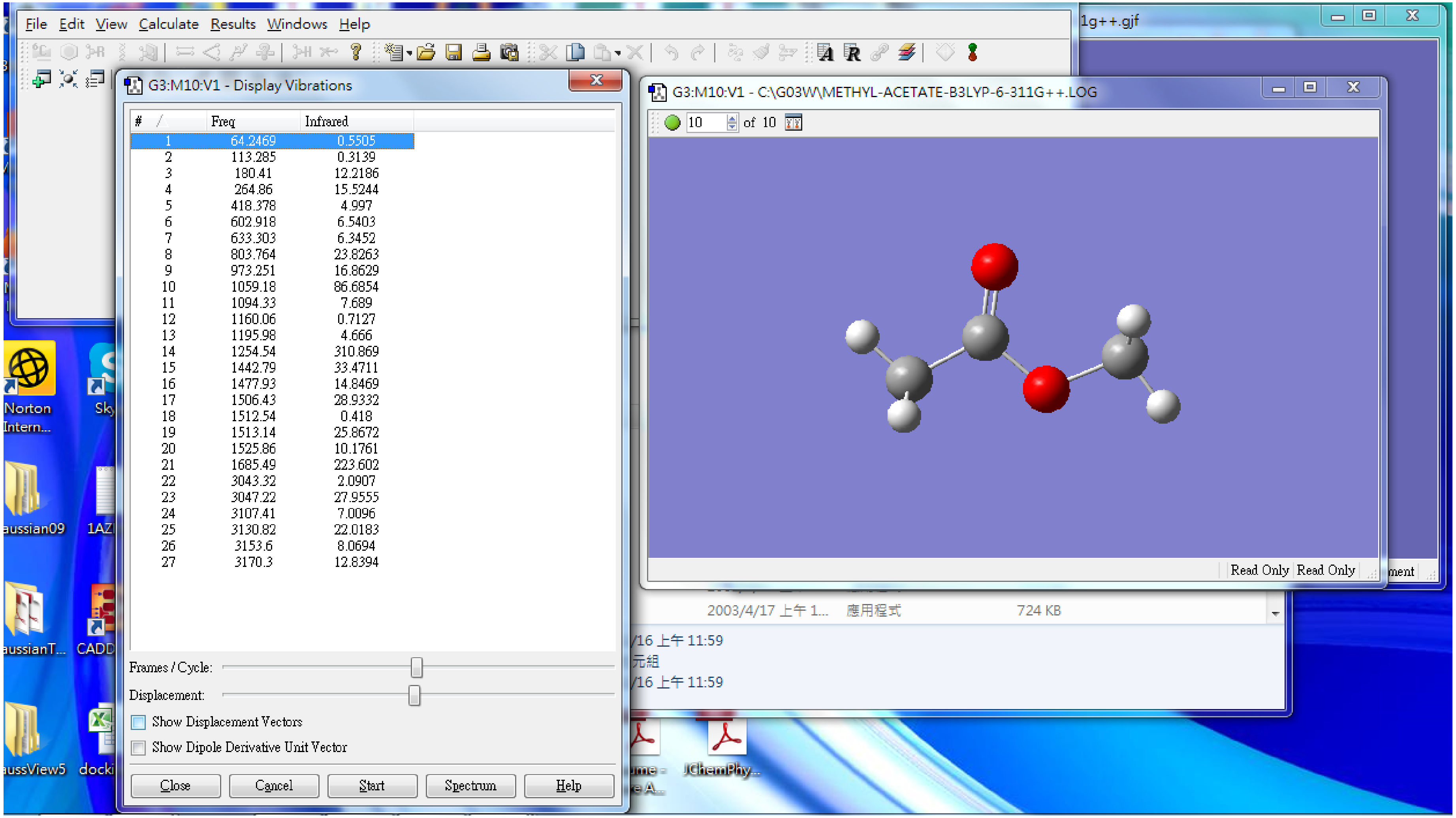
Bond	Wavenumber, $\tilde{\nu}$ cm^{-1}	Wavelength, λ μm
O—H, N—H	3100–3500	3.2–2.85
C—H	2800–3100	3.6–3.25
C—O, C—N	800–1300	12.5–7.7
C=C, C=O	1600–1800	6.25–5.55
P—O	1200–1300	8.3–7.7

The IR spectrum of a substance is also referred to as a fingerprint spectrum.





A screenshot of GaussView5 of methyl acetate IR calculation

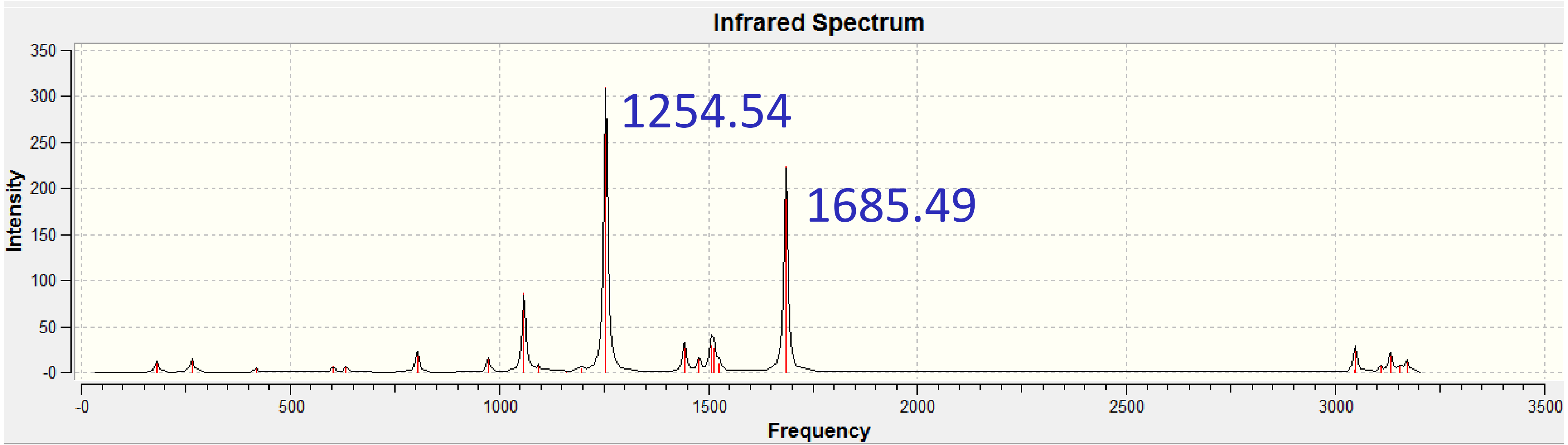


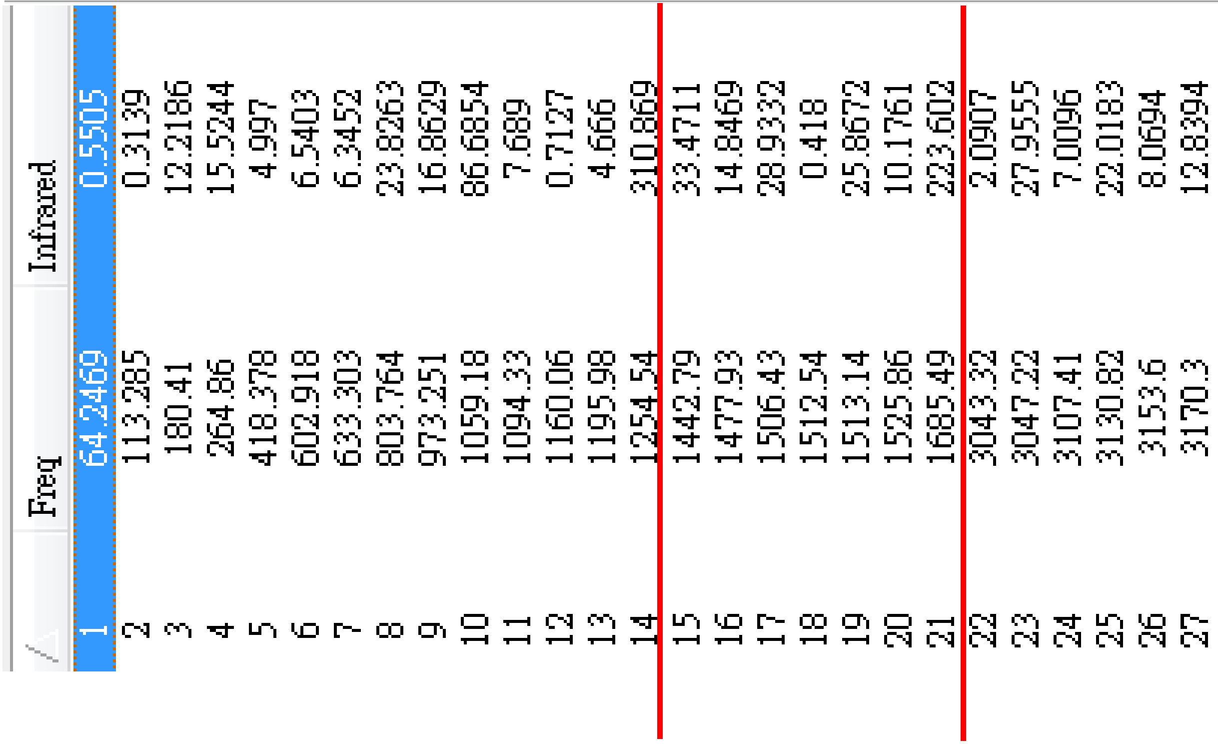
G3:M10:V1 - Display Vibrations

#	Freq	Infrared
1	64.2469	0.5505
2	113.285	0.3139
3	180.41	12.2186
4	264.86	15.5244
5	418.378	4.997
6	602.918	6.5403
7	633.303	6.3452
8	803.764	23.8263
9	973.251	16.8629
10	1059.18	86.6854
11	1094.33	7.689
12	1160.06	0.7127
13	1195.98	4.666
14	1254.54	310.869
15	1442.79	33.4711
16	1477.93	14.8469
17	1506.43	28.9332
18	1512.54	0.418
19	1513.14	25.8672
20	1525.86	10.1761
21	1685.49	223.602
22	3043.32	2.0907
23	3047.22	27.9555
24	3107.41	7.0096
25	3130.82	22.0183
26	3153.6	8.0694
27	3170.3	12.8394

G3:M10:V1 - C:\G03W\METHYL-ACETATE-B3LYP-6-311G++.LOG

Calculated infrared spectrum of methyl acetate

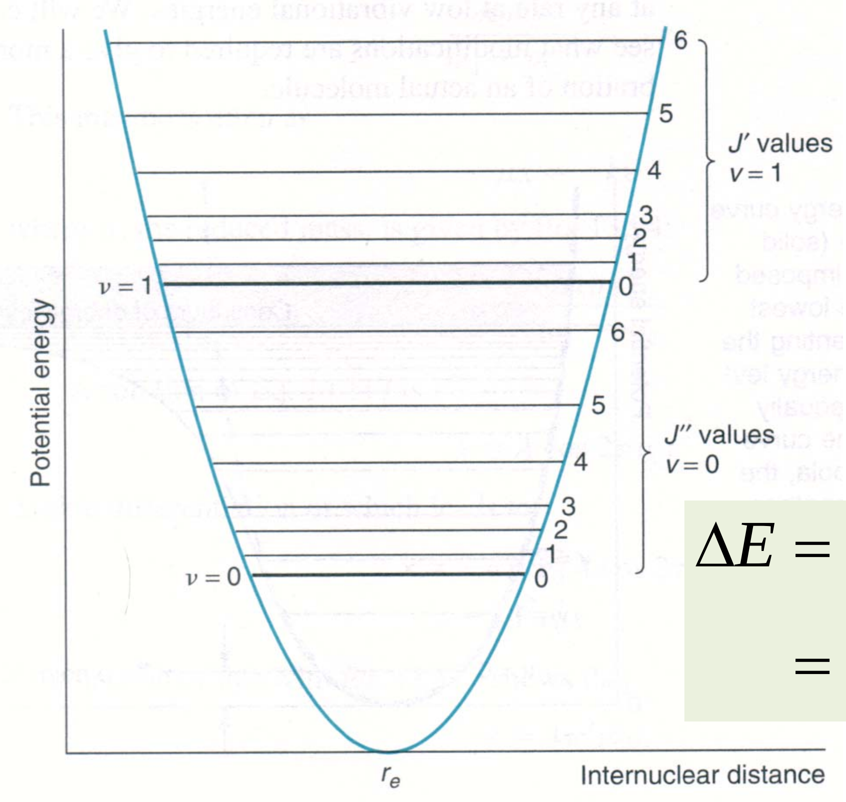




	Freq	Infrared
1	64.2469	0.5505
2	113.285	0.3139
3	180.41	12.2186
4	264.86	15.5244
5	418.378	4.997
6	602.918	6.5403
7	633.303	6.3452
8	803.764	23.8263
9	973.251	16.8629
10	1059.18	86.6854
11	1094.33	7.689
12	1160.06	0.7127
13	1195.98	4.666
14	1254.54	310.869
15	1442.79	33.4711
16	1477.93	14.8469
17	1506.43	28.9332
18	1512.54	0.418
19	1513.14	25.8672
20	1525.86	10.1761
21	1685.49	223.602
22	3043.32	2.0907
23	3047.22	27.9555
24	3107.41	7.0096
25	3130.82	22.0183
26	3153.6	8.0694
27	3170.3	12.8394

Combining rotational and vibrational spectra

Two lowest vibrational states (v=0 and v=1) with rotational states superimposed.



$$E_{\mathrm{v,J}} = \left(\mathrm{v} + \frac{1}{2}\right) h\nu_0 + \tilde{B} J(J+1) hc$$

$$\mathrm{v} = 0, 1, 2, 3, \ldots$$

$$J = 0, 1, 2, 3, \ldots$$

$$\begin{aligned} \Delta E &= E_{\mathrm{v}', J'} - E_{\mathrm{v}'', J''} \\ &= \left(\mathrm{v'\text{-}v''}\right) h\nu_0 + \tilde{B}[J'(J'+1) - J''(J''+1)]hc \end{aligned}$$

Vibrational-Rotational Spectra

Spectroscopic **term**

$$T = \frac{E}{hc}$$

$$\Delta T = \left(\mathrm{v}'\text{-}\mathrm{v}''\right)\tilde{\nu}_0 + \tilde{B}[J'(J'+1) - J''(J''+1)]$$

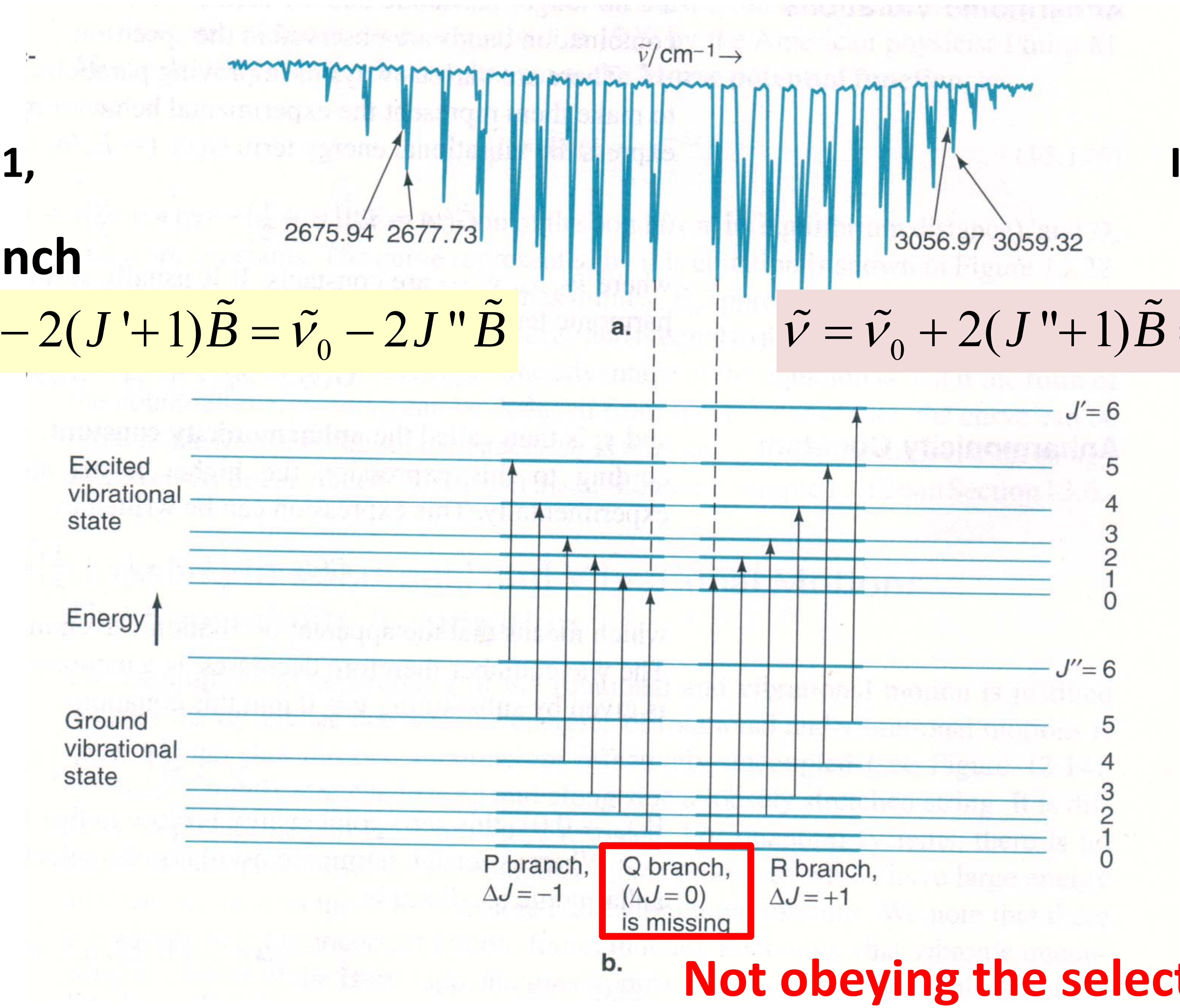
Selection rules for strict harmonic motion:

$$\Delta \mathrm{v} = \pm 1 \quad \text{and} \quad \Delta J = \pm 1$$

If v’-v”=1,

$$\Delta T = \tilde{\nu}_0 + \tilde{B}[J'(J'+1) - J''(J''+1)]$$

Vibrational-rotational spectrum of HCl



If J’=J”-1,

P branch

$$\tilde{\nu} = \tilde{\nu}_0 - 2(J'+1)\tilde{B} = \tilde{\nu}_0 - 2J''\tilde{B}$$

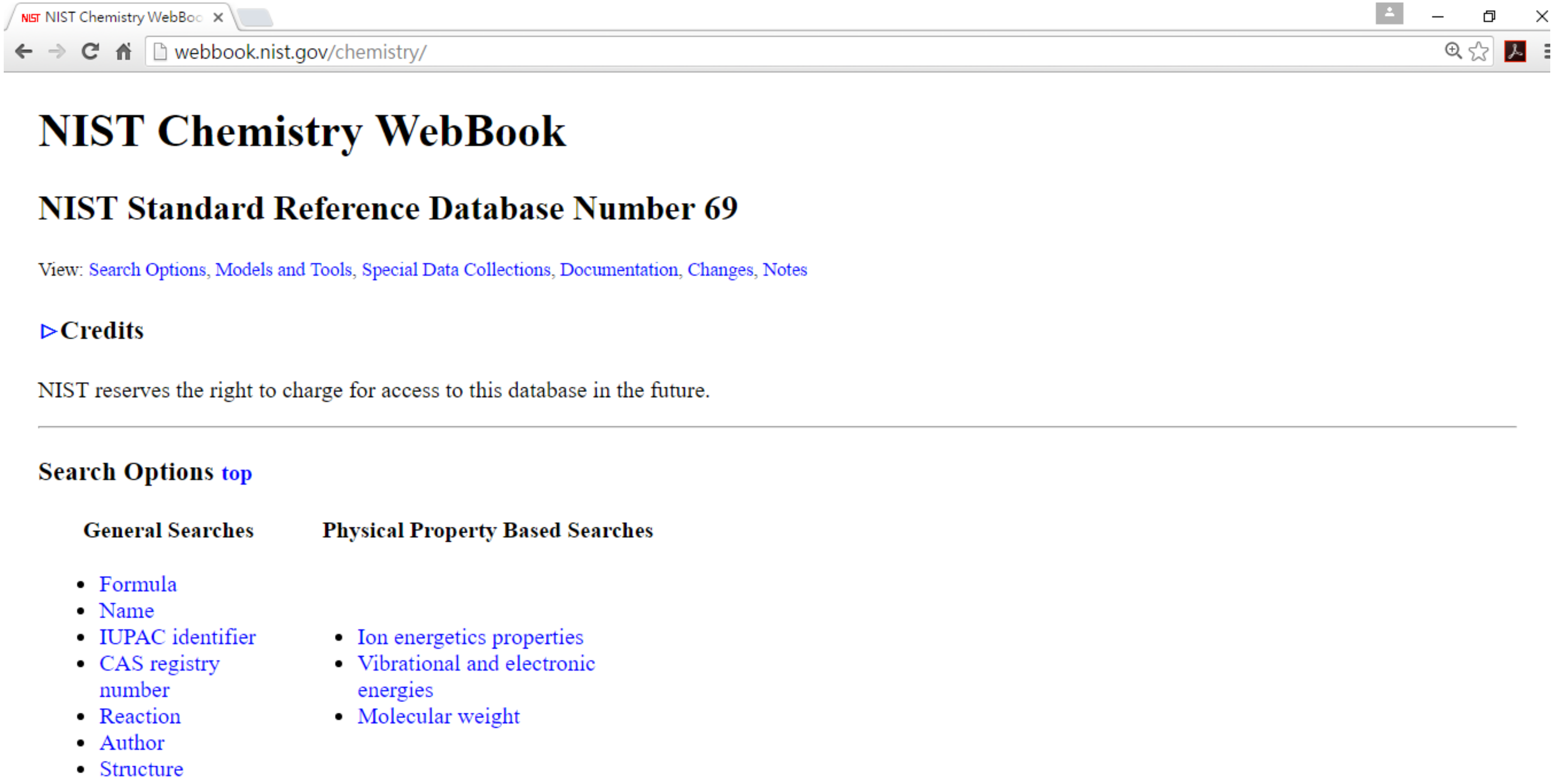
If J’=J”+1,

R branch

$$\tilde{\nu} = \tilde{\nu}_0 + 2(J''+1)\tilde{B} = \tilde{\nu}_0 + 2J'\tilde{B}$$

Not obeying the selection rules

A resource for structures, energies, reactions of chemicals



NIST Chemistry WebBook

NIST Standard Reference Database Number 69

View: Search Options, Models and Tools, Special Data Collections, Documentation, Changes, Notes

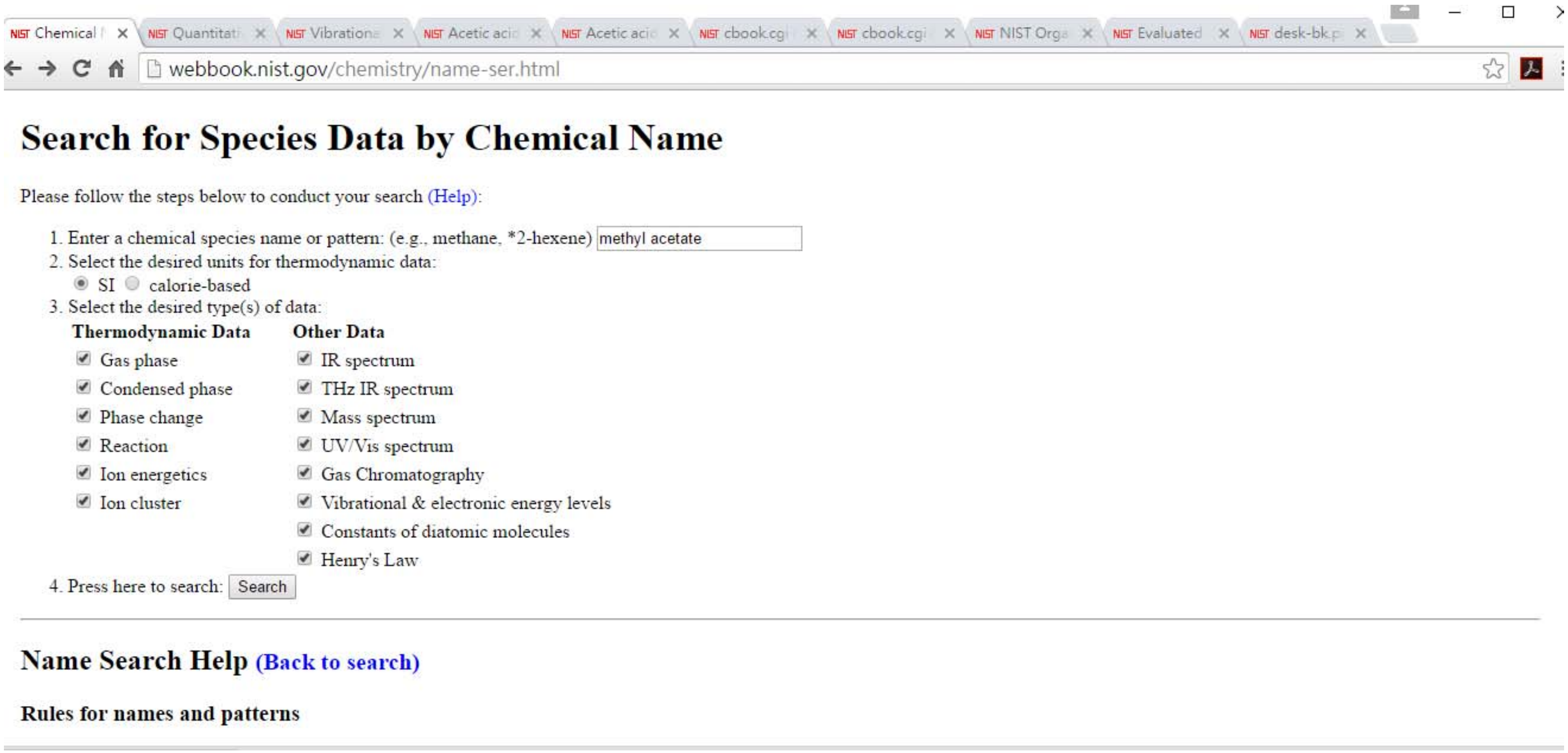
▷Credits

NIST reserves the right to charge for access to this database in the future.

Search Options top

General Searches	Physical Property Based Searches
Formula Name IUPAC identifier CAS registry number Reaction Author Structure	Ion energetics properties Vibrational and electronic energies Molecular weight

http://webbook.nist.gov/chemistry/



Search for Species Data by Chemical Name

Please follow the steps below to conduct your search (Help):

1. Enter a chemical species name or pattern: (e.g., methane, *2-hexene) methyl acetate
2. Select the desired units for thermodynamic data:
   ● SI ○ calorie-based
3. Select the desired type(s) of data:

Thermodynamic Data	Other Data
☑ Gas phase	☑ IR spectrum
☑ Condensed phase	☑ THz IR spectrum
☑ Phase change	☑ Mass spectrum
☑ Reaction	☑ UV/Vis spectrum
☑ Ion energetics	☑ Gas Chromatography
☑ Ion cluster	☑ Vibrational & electronic energy levels
	☑ Constants of diatomic molecules
	☑ Henry's Law

4. Press here to search: Search

Name Search Help (Back to search)

Rules for names and patterns

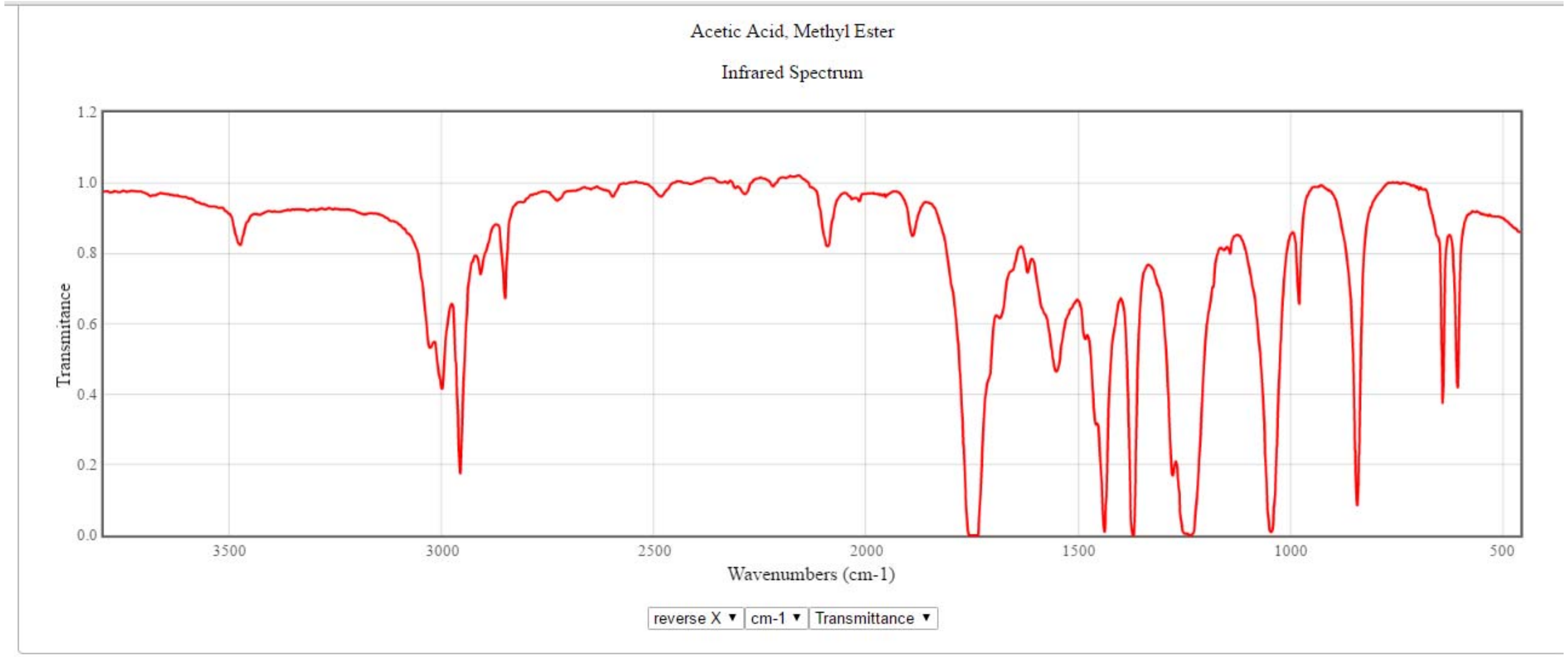
http://webbook.nist.gov/chemistry/name-ser.html

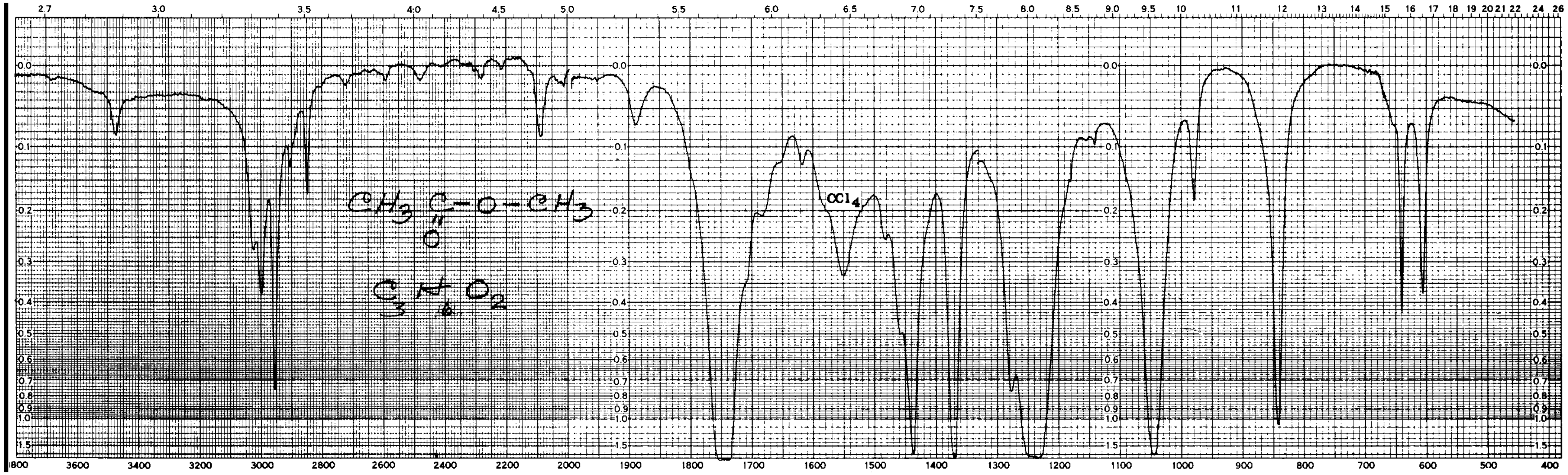
Acetic acid, methyl ester

- **Formula:** $C_3H_6O_2$
- **Molecular weight:** 74.0785
- **IUPAC Standard InChI:**
  - `InChI=1S/C3H6O2/c1-3(4)5-2/h1-2H3`
  - Download the identifier in a file.
- **IUPAC Standard InChIKey:** `KXKVLQRXCPHEJC-UHFFFAOYSA-N`
- **CAS Registry Number:** 79-20-9
- **Chemical structure:**

  This structure is also available as a 2d Mol file or as a computed 3d SD file
  The 3d structure may be viewed using Java or Javascript.
- **Other names:** Methyl acetate; Devoton; Tereton; CH3COOCH3; Methyl ethanoate; Acetate de methyle; Methyl acetic ester; Methylacetaat; Methylacetat; Methyle (acetate de); Methylester kiseliny octove; Metile (acetato di); Ethyl ester of monoacetic acid; UN 1231; Methyl ester of acetic acid; NSC 405071
- **Permanent link** for this species. Use this link for bookmarking this species for future reference.
- **Information on this page:**
  - Gas phase thermochemistry data
  - Condensed phase thermochemistry data
  - Phase change data
  - Reaction thermochemistry data

http://webbook.nist.gov/cgi/cbook.cgi?Name=methyl+acetate&Units=SI&cTG=on&cIR=on&cTC=on&cTZ=on&cTP=on&cMS=on&cTR=on&cUV=on&cIE=on&cGC=on&cIC=on&cES=on&cDI=on&cSO=on#IR-Spec

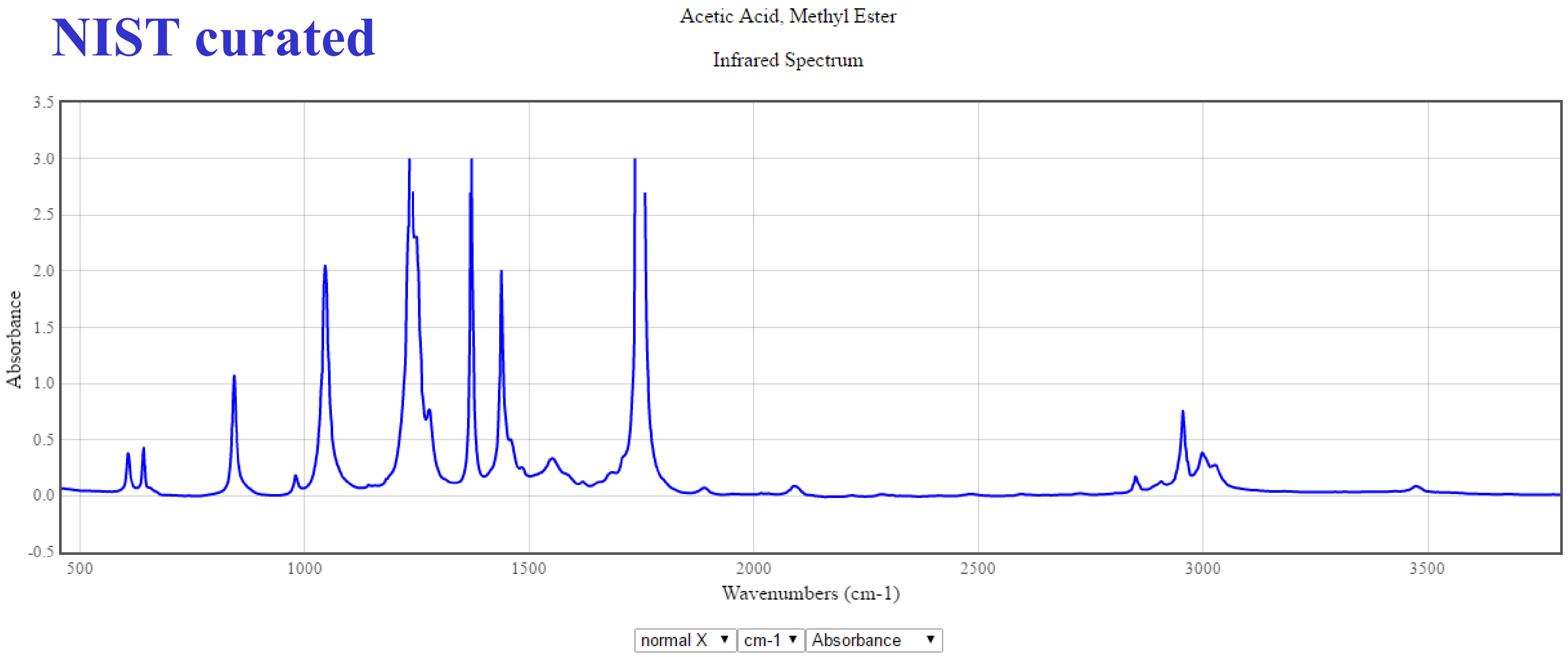




$CH_3\underset{O}{\underset{\|}{C}}-O-CH_3$

$C_3H_6O_2$

NIST curated



Calculated