Contents

1 Introduction

1.1 What do we care about?

Things chemistry/materials-related:

- What are the properties of atoms?
- What molecules do they make? What other substances do they make?
- What are the shapes of those molecules? Structures of those solids? Properties of them?
- How do those substances react with each other?
- What are the energies of those reactions?
- What are the rates of those reactions?
- What is the strongest substance?
- How do we make a substance to do. . . .?
- add your own questions. . . .

Things that relate to the *chemical* properties of substances.

1.2 How are we going to figure these out? With only a computer?

1926 Erwin Schrödinger equation: $\hat{H}\Psi = E\Psi$

1929 Paul Dirac, British physicist

> The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus *completely known*, and the difficulty lies

only in the fact that application of these laws leads to equations that are *too complex to be solved*.

It therefore becomes desirable that *approximate practical methods* of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

1930's-1950's Elaboration, analytical applications

1950's Computers start to appear for technical applications

1960's-1970's Numerical solutions of Schrödinger equation for atoms/molecules—expert users

1980's "Supercomputer" era—routine use of computational chemistry software becomes possible

Figure 1: Ohio State Cray Y-MP supercomputer, ca. 1989. World's fastest computer at the time. 333 MFlop top speed, 512 Mb RAM

1990's "Chemical accuracy" era—very precise solutions routinely available, for a cost! See Schneider, *J. Phys. Chem.*, **1994**.

1990's Density functional theory (DFT) allows applications to solids/surfaces/liquids to become common. See Hass, *Science*, **1998**

1990's Visualization moves to the desktop

2000's Computational "screening," materials discovery, materials genome. See Greeley, *J. Phys. Chem. C*, **2009**. Also see Gurkan, *J. Phys. Chem. Lett.*, **2010**.

Today Computational chemistry widely integrated into all aspects of chemical, materials, biological research

`Computational chemistry` is now so vast it is impossible to cover everything completely. We limit ourselves to quantum-mechanics-based calculations.

1.3 Our goals

1. Understand when it is appropriate to use quantum-mechanics methods.
2. Be able to state the basic theoretical, mathematical, and numerical concepts behind quantum mechanical "wavefunction theory" (WFT) and "density functional theory," (DFT) calculations.

3. Understand the terminology and practical issues associated with doing quantum chemical simulations.
4. Get hands-on experience with these concepts using popular computational tools of today, including GAMESS for molecular systems and Vasp for condensed phase systems.
5. Learn how to set up, execute, and analyze results in a modern, Python notebook environment.
6. Learn how to apply the results of quantum chemical simulations to calculate things you care about.
7. Demonstrate an ability to formulate a problem and apply QM methods to it.
8. Develop the skills to understand a literature paper in the area.

1.4 Reading resources

- These notes
- Chris Cramer, *Essentials of Computational Chemistry*, Wiley, 2004
- Martin,
- Sholl and Steckel, *Density Functional Theory: A Practical Introduction*, Wiley, 2009
- Kitchin book, `http://kitchingroup.cheme.cmu.edu/dft-book/`

1.5 Software tools

1.5.1 Notebooks

- org-mode
- jupyter/ipython

1.5.2 Molecular methods

- Avogadro environment `http://avogadro.cc/wiki/Main_Page`
- GAMESS code `http://www.msg.ameslab.gov/GAMESS/GAMESS.html`

1.5.3 Supercell methods

- ASE environment `https://wiki.fysik.dtu.dk/ase/`
- Vasp code `http://www.vasp.at/`

1.5.4 Great for getting started

- Webmo `http://www.webmo.net/`

2 Refresher on Quantum Mechanics

2.1 Why quantum mechanics?

Want to describe "mechanics" (equations of motion) of atomic-scale things, like electrons in atoms and molecules

Why? These ultimately determine the energy, the shape, and all the properties of matter.

de Broglie wavelength (1924)

$$\lambda = h/p = h/mv \tag{1}$$

$$h = 6.626 \times 10^{-34}\,\mathrm{J\,s}(\text{Planck's constant}) \tag{2}$$

	Car	Electron
mass m	$1000\,\mathrm{kg}$	$9.1 \times 10^{-31}\,\mathrm{kg}$
velocity v	$100\,\mathrm{km/hr}$	$0.01\,\mathrm{c}$
	typical value on the highway	typical value in an atom
momentum p	$2.8 \times 10^{-4}\,\mathrm{kg\,m/s}$	$2.7 \times 10^{-24}\,\mathrm{kg\,m/s}$
wavelength λ	$2.4 \times 10^{-38}\,\mathrm{m}$	$2.4 \times 10^{-10}\,\mathrm{m}$
	too small to detect. Classical!	Comparable to size of an atom. *Must* treat with QM!

How to describe wave properties of an electron? Schrödinger equation (1926)

Kinetic energy + Potential energy = Total Energy

Expressed as differential equation (Single particle, non-relativistic):

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r},t) + V(\mathbf{r},t)\Psi(\mathbf{r},t) = -i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) \tag{3}$$

If the potential V is time-invariant, can use separation of variables to show that the steady-state, time-independent solutions are characterized by an energy E and described by:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{4}$$

$$\Psi(\mathbf{r},t) = \psi(\mathbf{r})e^{-iEt/\hbar} \tag{5}$$

2.2 Postulates of non-relativistic quantum mechanics

2.3 Notes on constants and units

Resource on physical constants: http://physics.nist.gov/cuu/Constants/ Resource for unit conversions: http://www.digitaldutch.com/unitconverter/

Unit converter available in Calc mode of Gnu emacs **highly recommended**

Energy units 1 eV $= 1.602\,18 \times 10^{-19}\,\mathrm{J} = 96.485\,\mathrm{kJ/mol} = 8065.5\,\{\mathrm{cm}^{-1}\} = 11\,064\,\mathrm{K\,k_B}$

Table 1: Postulates of Non-relativistic Quantum Mechanics

Postulate 1: The physical state of a system is completely described by its wavefunction Ψ. In general, Ψ is a complex function of the spatial coordinates and time. Ψ is required to be:

I. Single-valued
II. continuous and twice differentiable
III. square-integrable ($\int \Psi^*\Psi d\tau$ is defined over all finite domains)
IV. For bound systems, Ψ can always be normalized such that $\int \Psi^*\Psi d\tau = 1$

Postulate 2: To every physical observable quantity M there corresponds a Hermitian operator $\hat{M}$. **The only observable values of M are the eignevalues of $\hat{M}$.**

Physical quantity	Operator	Expression
Position x, y, z	$\hat{x}, \hat{y}, \hat{z}$	$x\cdot, y\cdot, z\cdot$
Linear momentum $p_x, \ldots$	$\hat{p}_x, \ldots$	$-i\hbar\frac{\partial}{\partial x}, \ldots$
Angular momentum $l_x, \ldots$	$\hat{p}_x, \ldots$	$-i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \ldots$
Kinetic energy T	$\hat{T}$	$-\frac{\hbar^2}{2m}\nabla^2$
Potential energy V	$\hat{V}$	$V(\mathbf{r}, t)$
Total energy E	$\hat{H}$	$-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}, t)$

Postulate 3: If a particular observable M is measured many times on many identical systems is a state Ψ, the average resuts with be the expectation value of the operator $\hat{M}$:

$$\langle M \rangle = \int \Psi^*(\hat{M}\Psi)d\tau$$

Postulate 4: The energy-invariant states of a system are solutions of the equation

$$\begin{aligned} \hat{H}\Psi(\mathbf{r}, t) &= i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r}, t) \\ \hat{H} &= \hat{T} + \hat{V} \end{aligned}$$

The time-independent, stationary states of the system are solutions to the equation

$$\hat{H}\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$$

Postulate 5: (The **uncertainty principle**.) Operators that do not commute ($\hat{A}(\hat{B}\Psi) \neq \hat{B}(\hat{A}\Psi)$) are called *conjugate*. Conjugate observables cannot be determined simultaneously to arbitrary accuracy. For example, the standard deviation in the measured positions and momenta of particles all described by the same Ψ must satisfy $\Delta x \Delta p_x \geq \hbar/2$.

Table 2: Atomic units common for quantum mechanical calculations (see http://en.wikipedia.org/wiki/Atomic_units)

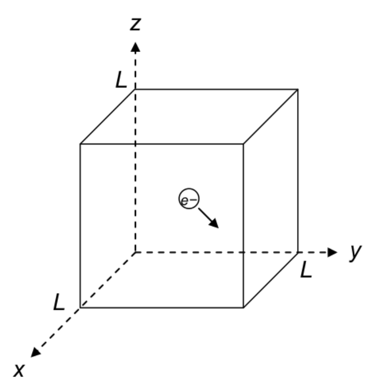
	Atomic unit	SI unit	Common unit
Charge	$e = 1$	1.6021×10^{-19} C	
Length	$a_0 = 1$ (bohr)	$5.291\,77 \times 10^{-11}$ m	$0.529\,177$ Å
Mass	$m_e = 1$	$9.109\,38 \times 10^{-31}$ kg	
Angular momentum	$\hbar = 1$	$1.054\,572 \times 10^{-34}$ J s	
Energy	$E_h = 1$ (hartree)	$4.359\,744 \times 10^{-18}$ J	27.2114 eV
Electrostatic force	$1/(4\pi\epsilon_0) = 1$	$8.987\,552 \times 10^{-9}$ N m^2/C^2	
Boltzmann constant		$1.380\,65 \times 10^{-23}$ J/K	$8.617\,33 \times 10^{-5}$ eV/K

2.4 Example: Energy states of a particle in a box

System defined by potential experienced by particle:

$V(\mathbf{r}) = 0, \qquad 0 < x, y, z < L$

$V(\mathbf{r}) = \infty, \qquad x, y, z \le 0,\ x, y, z \ge L$



3D box → 3 degrees of freedom/coordinates

Schrödinger equation

$$-\frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi(x, y, z) = E\psi(x, y, z) \tag{6}$$

$$\psi(x, y, z) = 0, \quad x, y, z \le 0,\ x, y, z \ge L \tag{7}$$

A second-order, linear, partial differential equation. Boundary value problem. Solve by separation of variables. Postulate $\psi(x, y, z) = X(x)Y(y)Z(z)$. Substituting and rearrange to get

$$-\frac{\hbar^2}{2m_e}\left(\frac{1}{X(x)}\frac{\partial^2 X(x)}{\partial x^2} + \frac{1}{Y(y)}\frac{\partial^2 Y(y)}{\partial y^2} + \frac{1}{Z(z)}\frac{\partial^2 Z(z)}{\partial z^2}\right) = E \qquad 0 < x, y, z < L \tag{8}$$

ftn x + ftn y + ftn z = constant → each term must be constant.

Equation for each dimension

$$-\frac{\hbar^2}{2m_e}\frac{\partial^2 X(x)}{\partial x^2} = E_x X(x), \qquad X(0) = X(L) = 0 \tag{9}$$

Seek function that twice differentiated returns itself and satisfies boundary conditions.

$$X(x) = \sin\frac{n_x \pi x}{L}, \qquad n_x = 1, 2, 3, \ldots \tag{10}$$

$$E_{n_x} = \frac{n_x^2 \pi^2 \hbar^2}{2m_e L^2} \tag{11}$$

Solutions called *eigenfunctions* (or *wavefunctions*) and *eigenvalues*. Characterized by *quantum numbers*, one for each degree of freedom. These (and all QM) solutions have certain special properties, including that they are orthonormal and form a complete set.

Normalization

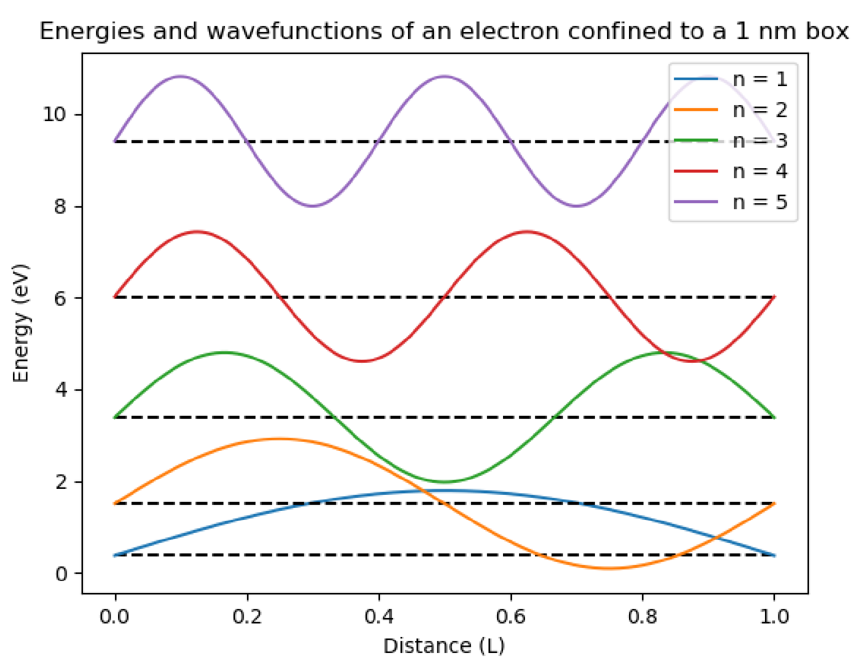
Seek a constant such that the inner eigenfunction product is unity.

$$C^2\int_0^L \sin^2\frac{n_x\pi x}{L}dx = C^2 L/2 = 1 \rightarrow C = \pm\sqrt{\frac{2}{L}} \tag{12}$$

$$X(x) = \pm\sqrt{\frac{2}{L}}\sin\frac{n_x\pi x}{L}, \qquad n_x = 1, 2, 3, \ldots \tag{13}$$

Orthonormal

$$\langle X_{n_x}|X_{n_x'}\rangle = \delta_{n_x,n_x'} \qquad \text{Dirac notation} \tag{14}$$



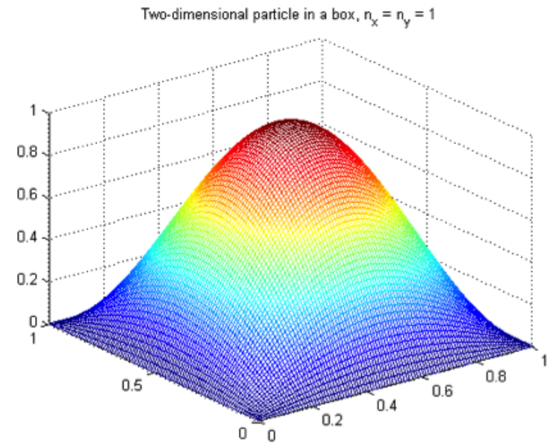
- Energy increase with number of *nodes*.

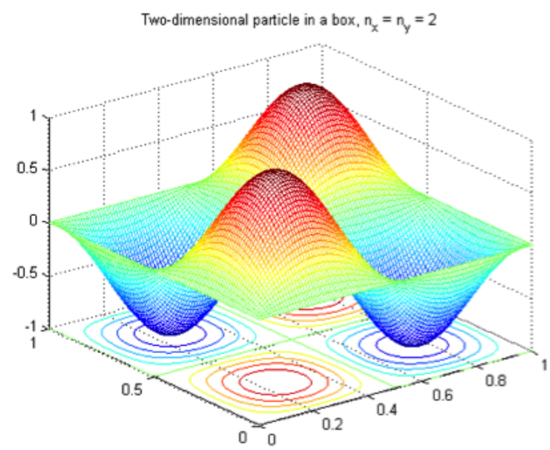
- $E \propto n^2, \Delta E \propto n, \Delta E/E \propto 1/n$. *Relative* spacing decreases with n.
- Is this real? See Ho, *J. Phys. Chem. B* **2005**, *109*, 20657.

Three-dimensional solutions

$$\psi(x, y, z) = X(x)Y(y)Z(z) = \left(\frac{2}{L}\right)^{3/2} \sin\frac{n_x \pi x}{L} \sin\frac{n_y \pi y}{L} \sin\frac{n_z \pi z}{L}, \qquad n_x, n_y, n_z = 1, 2, 3, \ldots \tag{15}$$

$$E = E_x + E_y + E_z = \frac{(n_x^2 + n_y^2 + n_z^2)\pi^2\hbar^2}{2mL^2} \tag{16}$$





Properties of solutions:

- Symmetry of system introduces degeneracy in solutions
- Energy depends on volume → pressure!

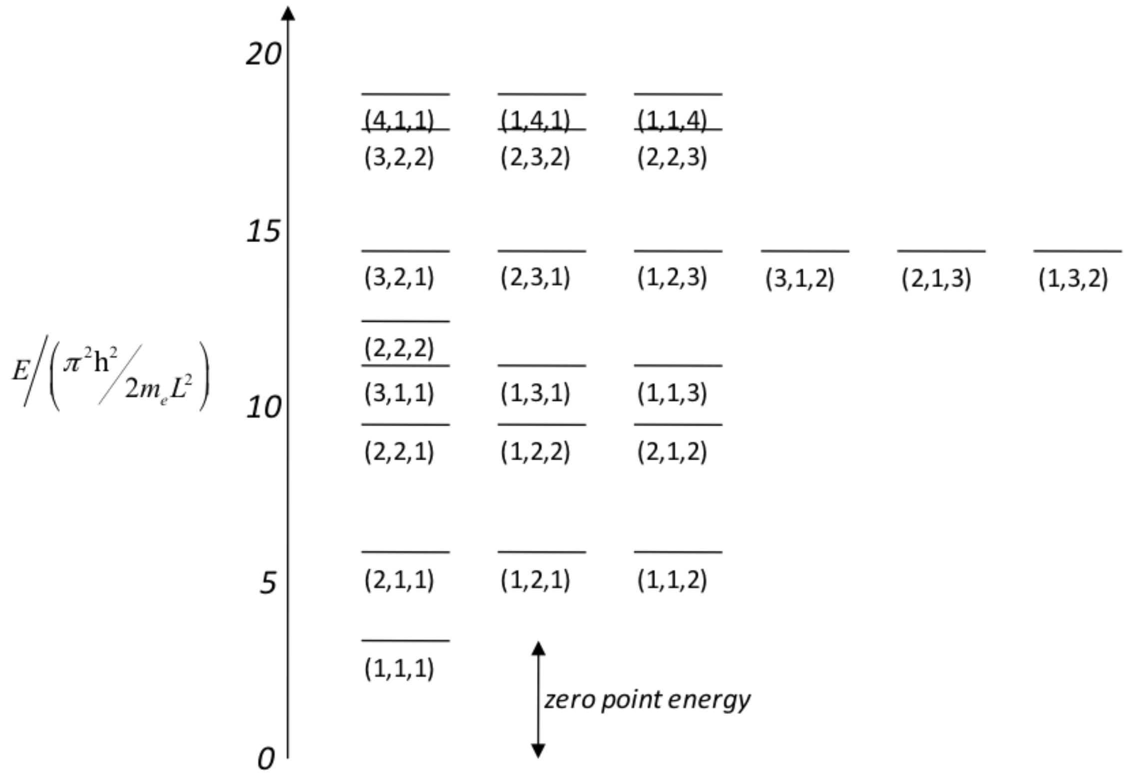


Figure 2: Energy sates of 3D Particle in a box

3 Hydrogen atom

3.1 Hydrogen atom solutions

3.2 Variational principle

3.3 Basis functions

3.4 Secular equations

3.5 Spin

4 (Two is too) many electrons

4.1 What's the problem?

Helium: next (after hydrogen) simplest atom

In a sense, we "know" the answer… $1s^2$. But is this same $1s$ as H? No! Different nuclear charge, interactions between the two electrons. This is an approximation and a very convenient shorthand!

4.1.1 Schrödinger equation for He

Wavefunction $\Psi(\mathbf{r}_1, \mathbf{r}_2)$, atom energy E.

Define 1-electron operator for each electron, in atomic units. Include kinetic energy of electron and its attraction to nucleus of charge $Z = 2$:

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \frac{Z}{|\mathbf{r}_i|}$$

Looks similar to hydrogen atom.

BUT, electrons also repel. Total Schrödinger equation for He:

$$\left\{\hat{h}_1 + \hat{h}_2 + \frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|}\right\}\Psi(\mathbf{r}_1, \mathbf{r}_2) = E\Psi(\mathbf{r}_1, \mathbf{r}_2)$$

Last term accounts for electron-electron electrostatic repulsion. Makes problem non-separable and really hard to solve. (How many solutions are there?)

Generalize to n-electron atom, in atomic units:

$$\hat{H} = \sum_i \hat{h}_i + \sum_{j>i+1} \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \tag{17}$$

$$\hat{H}\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n) = E\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n) \tag{18}$$

First summation over all electrons, second gets all electron pairs.

Solutions are many-dimensional functions of the coordinates of all the electrons. Cannot solve this analytically, although approaches exist (eg quantum Monte Carlo) that can in principle get very close. Thankfully, though, we can make approximations that work out really well. We'll look at three historically important ones.

4.2 The Hartree atom

Simplest approach is to approximate Ψ. Douglas Hartree (1897-1958) writes:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) \approx \psi_1(\mathbf{r}_1) \cdot \psi_n(\mathbf{r}_n)$$

So-called Hartree product. Can't be right. It gives the probability of two electrons being in the same place as some number > 0! Neglects *electron correlation*. How to apply?

1. Apply variational principle: What's the best possible set of ψ_i? We'll say best are the set that give the lowest expectation value of energy.

$$\langle E \rangle = \langle \Psi | \hat{H} | \Psi \rangle \tag{19}$$

$$\frac{\delta \langle E \rangle}{\delta \psi_i} = 0, \forall i \tag{20}$$

2. Lagrange multipliers to impose orthonormality constraint on ψ_i:

$$\langle\psi_i|\psi_j\rangle = \delta_{ij} \tag{21}$$

$$L = \langle E\rangle - \sum_{i,j}\epsilon_{ij}\left(\langle\psi_i|\psi_j\rangle - \delta_{ij}\right) \tag{22}$$

$$\delta L = 0 \tag{23}$$

3. Coupled, one-electron Hartree eigenvalue equations for energy-optimal ψ_i:

$$\left\{\hat{h}_i + \hat{v}_i^{\text{Hartree}}\right\}\psi_i(\mathbf{r}_1) = \epsilon_i\psi_i(\mathbf{r}_1) \tag{24}$$

$$\hat{v}_i^{\text{Hartree}}(\mathbf{r}_i) = \sum_{j\neq i}\int|\psi_j(\mathbf{r}_2)|^2\frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|}d\mathbf{r}_2 \tag{25}$$

Have to solve this for all n electrons of an atom/molecule. "Hartree potential" represents Coulomb repulsion between electron i and all other electrons, *averaged* over position of those electrons. Always positive. This is a *mean field* approximation. Note appearance of "one electron" energies, ϵ_i, kinetic energy plus repulsion of electron with all others. Total energy is sum of these ϵ_i corrected to avoid overcounting repulsions:

$$\langle E\rangle = \sum_i\epsilon_i - \frac{1}{2}\sum_i\langle\psi_i|\hat{v}_i^{\text{Hartree}}|\psi_i\rangle$$

Presents an obvious difficulty. If we don't know ψ_j ahead of time, how can we even construct Hartree equations, let alone solve them? Hartree offered a numerical solution, in the 1930's, called the ***self-consistent field* (SCF) approach**:

1. Guess an initial set of ψ_i, one for each electron (he did this on a grid, and jumping ahead a bit, allowed each ψ_i to represent two electrons)
2. Construct Hartree potential for each ψ_i
3. Solve the n differential equations for n new ψ_i
4. Compare new to old ψ_i
5. If the same within a desired tolerance, you are done!
6. If not, return to step 2, using new ψ_i, and repeat.

Hartree's father did this by hand for all the atoms of the periodic table, tabulating wavefunctions and energies for all the electrons in each. For instance, for He, he'd solve one equation, self-consistently, to get one ψ_1, and then combine to get $\Psi(1,2) = \psi_1(1)\alpha(1)\psi_1(2)\beta_2$. Tedious! Qualitatively great, quantitatively not so hot. Mean-field approximation just not so hot.

Nonetheless, basic idea of representing many-body wavefunction in terms of "orbitals," of setting up orbital equations, and solving using a self-consistent procedure, remain today at the heart of virtually all electronic structure calculations. Hurrah Hartree!

4.3 The Pauli principle

One big conceptual short-coming of the Hartree model is that it treats the electrons as if they were distinguishable. QM says electrons are indistinguishable. Furthermore, they have a quantized angular momentum, called a spin, that is either up or down, making them fermions.

> **Pauli principle**: The wavefunction of a multi-particle fermion system must be anti-symmetric to coordinate exchange.
>
> $$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = -\Psi(\boldsymbol{x}_2, \boldsymbol{x}_1)$$

Here the coordinate $\mathbf{x}$ includes both the position and the spin (up or down, α or β) of the electron.

Sorry Hartree. Can fix for He by writing

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \psi_1(\boldsymbol{r}_1)\psi_i(\boldsymbol{r}_2)\left(\alpha(1)\beta_2 - \beta(1)\alpha(2)\right)$$

4.4 Slater determinants and exchange

4.5 Hartree-Fock equation

4.5.1 Basis of wavefunction theory (WFT)

4.6 Hartree-Fock-Slater

4.6.1 Basis of density functional theory (DFT)

5 Practical electronic structure

5.1 Born-Oppenheimer approximation

In principle all nuclei and electrons should be described quantum mechanically. For H_2, for instance, true wavefunction would be a function of the positions of nuclei and electrons, $\Upsilon(\mathbf{r}_1, \mathbf{r}_2, \mathbf{R}_1, \mathbf{R}_2)$.

Nuclei much heavier than electrons and move much more slowly. Assume nuclei are fixed in space ("clamped") and electrons move in static field of those electrons. Equivalent to assuming that nuclear kinetic energy is decoupled from electron dynamics. Schrödinger equation becomes parameteric in nuclear positions; solutions $E(\mathbf{R}_1, \mathbf{R}_2)$ define a potential energy surface (PES).

5.2 Model chemistry

Essentially always start with

$$\left\{\hat{h} + v_{\text{Coulomb}}[\rho] + v_{\text{exchange}}[\psi_i] + v_{\text{correlation}}[\psi_i]\right\}\psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \tag{26}$$

Standard models of today all treat the one-electron and Coulomb pieces exactly and treat the electron-electron interactions at various levels of approximation.

	v_{exchange}	$v_{\text{correlation}}$	
Wave function theory (WFT)			
Hartree	self-interaction	neglect	historic
Hartree-Fock	exact	neglect	superceded
MPn, CC	exact	perturbative	state-of-the-art
CI	exact	variational	specialized
Density functional theory (DFT)			
Hartree-Fock-Slater	$[\rho^{4/3}]$	neglect	historic
Local density approximation (LDA)	$[\rho]$	$[\rho]$	general purpose solids
Generalized gradient approximation (GGA)	$[\rho, \nabla\rho]$	$[\rho, \nabla\rho]$	general purpose
Hybrid	$\approx$ exact	$[\rho, \nabla\rho]$	general purpose molecules

The choice of the electronic structure model is the most fundamental approximation in applying these methods. Determined from experience and need.

Specification in GAMESS is a bit arcane. Default is Hartree-Fock. To specify DFT model, use

```
$CONTRL DFTTYP =   Slater (HFS), SVWN (LDA), PBE (GGA), B3LYP (Hybrid)
```

5.3 Open-shell systems

Model has to be generalized somewhat to deal with systems with unpaired electrons.

5.4 Bring back the basis sets

The one-electron HF or HFS equations give us defining equations for the energy-optimal orbitals, but they aren't convenient to solve for anything more complicated than an atom. What to do? Reintroduce idea of a basis set.

5.5 Semi-empirical methods

5.6 Examples

5.6.1 H2

5.6.2 HF

5.7 Symmetry

5.8 Population analysis

5.9 Molecular orbital (MO) diagrams

5.10 Gaussian basis sets

Gaussian functions ($e^{-\zeta|\mathbf{r}|^2}$) are the most popular choice for atom-centered basis sets. They do not efficiently represent molecular wavefunctions, but one- and two-electron integrals in WFT can be solved analytically over Gaussians.

Other choices, like Slater functions ($e^{-\zeta|\mathbf{r}|}$) are possible but require numerical quadrature.

Gaussian basis sets have to be created for any given atom and must be used consistently within a set of calculations.

- Primitive is a single Gaussian function, possibly multipled by a polynomial to look like s, p, ... Defined by an exponent ζ that determines how extensive (small ζ) or compact the function is.
- Contraction is a pre-set linear combination of several primitive Gaussians.
- Basis set is a predefined set of exponents and contraction coefficients appropriate for some specific atom.

Common notation

- Minimal basis contains one contracted function for every atomic orbital

5.11 Electron cores

5.12 Performance details of SCF methods

Basis is often *orthonormalized* to eliminate overlap from H-F-R equation; allows equations to be solved by matrix diagonalization.

Initial density matrix $\mathbf{P}$ are obtained by solving an approximate Hamiltonian (like extended Hückel). Always beware! Initial guess can influence final converged state.

Because the number of 2-electron integrals grows as N^4, they are sometimes calculated as needed "on-the-fly", so-called direct SCF.

The SCF procedure is an optimization problem: find set of coefficients that minimizes the total energy. As discussed above, success depends on a reasonable initial guess for density matrix and judicious updating. Various strategies can be used to speed and stabilize convergence, like damped mixing of previous cycles.

Second-order SCF is a convergence acceleration method that requires calculation or estimation of the first- and second-derivatives of the energy with respect to the orbital coefficients. See e.g. Chaban et al., *Theor. Chem. Accts.* **1997**, *97*, 88-95.

Pulay's "direct inversion in the iterative subspace," or "DIIS," is a popular and powerful acceleration procedure that extrapolates from several previous Fock matrices to predict optimal next Fock to diagonalize.

Controlled in GAMESS using the `$SCF` group.

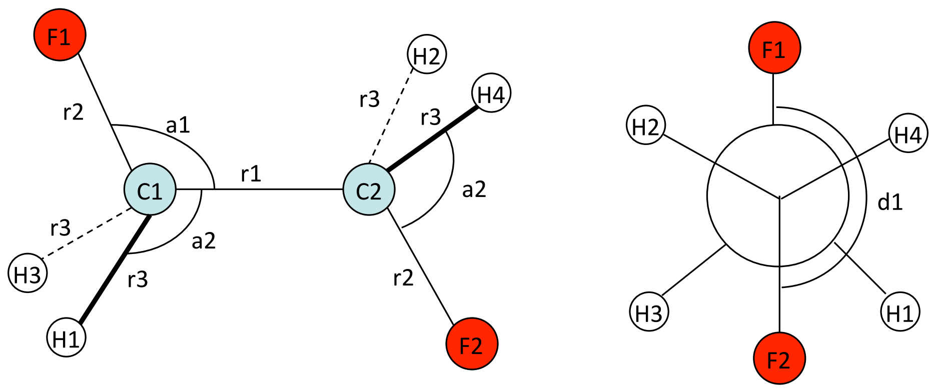
```
$SCF DIRSCF=   .T./.F. controls direct scf
     SOSCF=    .T./.F. second-order scf
     DIIS=     .F./.T. direct inversion in the iterative subspace
     DAMP=     .T./.F. damping, on for initial iterations
$END
```


6 Potential energy surfaces

The potential energy surface ("PES") is the sum of the repulsive energy of the nuclei and the kinetic and potential energies of all the electrons:

$$E_{\text{PES}}(\mathbf{R}_\alpha, \mathbf{R}_\beta, \ldots) = E_{\text{elec}} + \sum_{\alpha=1}^{N} \sum_{\beta=\alpha+1}^{N} \frac{Z_\alpha Z_\beta e^2}{R_{\alpha\beta}} \tag{27}$$

6.1 Specifying atomic positions



6.1.1 Cartesian

Computationally straightforward but don't correspond with our physical notion of bonds, bends, etc. Easiest to get out of a piece of software. A molecule has $3N - 6$ internal degrees of freedom ($3N - 5$ if linear), but Cartesians specify $3N$. The extra values correspond to the location of the center of mass and molecular oriendation. Codes will typically center and reorient the Cartesions.

In Gamess, would specify Cartesian coordinates for FCH_2CH_2F like this:

```
 $CONTRL COORD=CART $END
 $DATA
FCH2CH2F drag calculation
C1
C     6.0    -3.76764     0.33879     0.03727
C     6.0    -2.35246     0.34495     0.03689
F     9.0    -4.72277     0.58147    -1.18012
F     9.0    -1.59909    -0.68487    -0.83662
H     1.0    -4.04387     1.08375     0.75395
H     1.0    -3.92958    -0.71060     0.16941
H     1.0    -2.03786     0.18875     1.04760
H     1.0    -2.09983     1.28759    -0.40187
 $END
```

6.1.2 Internal coordinates

These provide a more intuitive representation and can be convenient when building molecules by hand. In codes like `GAMESS`, most commonly defined using "z-matrix" notation. Specify each atom in terms of its distance, angle, and dihedral angle with three previous atoms.

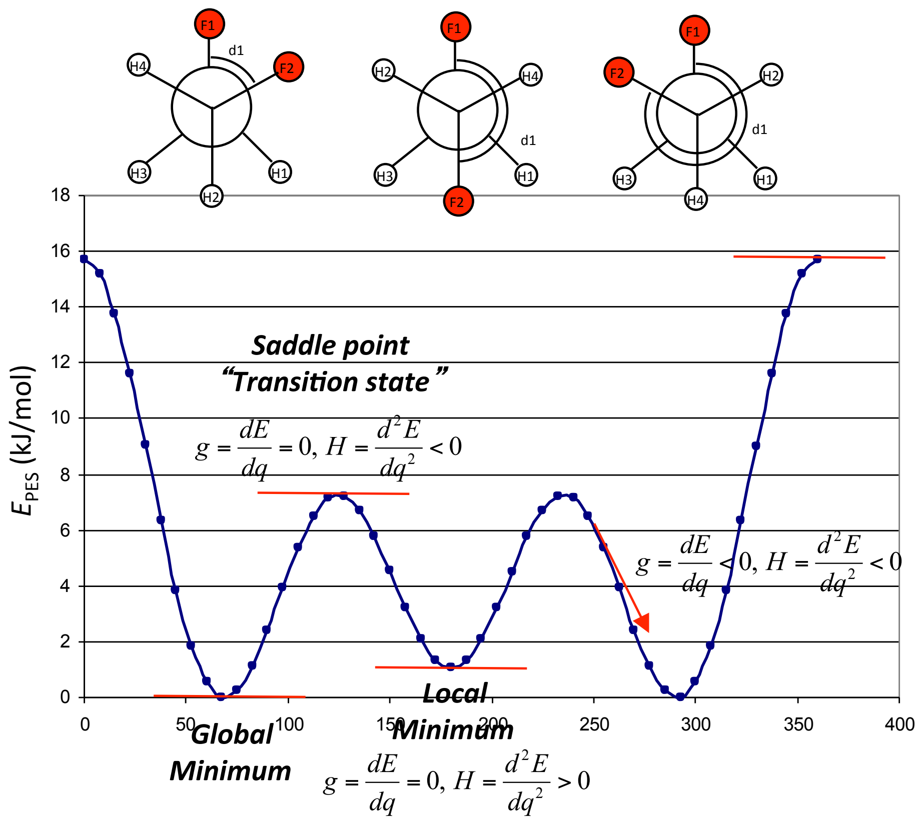
In Gamess, would specify z-matrix for FCH_2CH_2F like this:

```
$CONTRL SCFTYP=RHF RUNTYP=ENERGY COORD=ZMT $END
$DATA
FCH2CH2F drag calculation
C1
C
C   1   r1
F   2   r2   1   A1
H   2   r3   1   A2   3   D1
H   2   r4   1   A3   3   D2
F   1   r2   2   A1   3   D3
H   1   r3   2   A2   6   D1
H   1   r4   2   A3   6   D2

r1=1.5386
r2=1.39462
r3=1.11456
r4=1.12
A1=109.54214
A2=111.
A3=110.
D1=120.
D2=-120.5
D3=50.
$END
```

Particularly convenient when you'd like to "scan" over the value of some coordinate. Variable can be applied to more than one independent coordinate, if the molecule has symmetry.

6.2 Features of potential energy surfaces



6.2.1 Gradients

6.2.2 Hessians

6.2.3 Minima

- Local
- Global

6.2.4 Saddle points

- First-order
- Higher order

6.2.5 Minimum energy paths

6.3 Energy gradients and second derivatives

6.4 Optimization algorithms

6.4.1 Steepest descent

6.4.2 Conjugate gradient

6.4.3 Quasi-Newton Raphson

6.4.4 Direct inversion in the iterative subspace (DIIS)

6.4.5 Trudge

6.4.6 Genetic algorithm

6.4.7 Convergence criteria

6.5 Geometry optimization algorithms

6.6 Efficient coordinate systems

6.7 Performance of models

6.8 Vibrational frequencies

6.9 Transition states

6.10 Intrinsic reaction coordinates

6.11 Molecular dynamics

7 GAMESS Cheat Sheet

7.1 Specifying electronic configuration

7.1.1 Closed shell (default)

```
$CONTRL RHFTYP=RHF MULT=1 $END
```

7.1.2 Open-shell (n = spin-multiplicity = # unpaired electrons + 1)

```
$CONTRL RHFTYP=UHF MULT=n $END
```

7.2 Specifying electronic structure method

7.2.1 Hartree-Fock (default)

7.2.2 DFT

```
$CONTRL DFTTYP=xxx $END
```

`Gamess` supports many DFT functionals. See the $DFT section of the manual for a full set. Common DFT methods include:

method	xxx
Slater	Slater
Local density approx	SVN
Generalized gradient approx	PBE
Hybrid DFT	B3LYP
"Minnesota" optimized	M06
	M11

7.2.3 Beyond Hartree-Fock

Many methods available. See manual for full description. Most common is second-order perturbation theory, "MP2," :

```
$CONTRL MPLEVL=2 $END
```

If you want a very high quality number, have a big computer and time to wait, try "coupled cluster,":

```
$CONTRL CCTYP=CCSD(T) $END
```

7.3 Specifying basis sets

Gamess uses atom-centered basis functions. A "basis set" is a set of such functions for many atoms, all constructed (hopefully) at a consistent level of accuracy. Many such basis sets exist and are coded into Gamess, and for the daring new basis sets can be input by hand. Choice of basis set is always a compromise between accuracy and computational cost. In general should always check sensitivity of property of interest to basis set.

Specified in $BASIS group. Some common choices, in increasing level of sophistication:

Name	Type	Flags
	Pople type	**The most venerable and widely used**
STO-3G	Minimal	GBASIS=STO NGAUSS = 3
3-21G	Split valence	GBASIS=N21 NGAUSS=3
6-31G(d)	Split valence polarized	GBASIS=N31 NGAUSS =6 NDFUNC=1
6-311+G(d,p)	Triple-split valence	GBASIS=N311 NGAUSS=6 NDFUNC=1 NPFUNC=1
	polarized and augmented	DIFFSP=1
	Polarization-consistent	**Good for DFT**
PC0	Minimal	GBASIS=PC0
PC1	Split valence polarized	GBASIS=PC1
PC2	Triple split double polarized	GBASIS=PC2
	Correlation-consistent	**Good for MP2 and beyond**
cc-pVDZ	Split valence polarized	GBASIS=CC2
cc-pVTZ	Triple split double polarized	GBASIS=CC3
	Effective core potentials	**Good for treating heavy atoms**
SBKJC	Split valence + core potential	GBASIS=SBKJC
Hay-Wadt	Split valence + core potential	GBASIS=HW

7.4 Specifying geometry

Again `GAMESS` has a number of options, several of which are arcane and seldom used. Most common are Cartesian and z-matrix. Here I give examples ignoring any symmetry the molecule might have.

7.4.1 Cartesian

Specify an atom name, atomic number, and cartesian positions in Å. Over-specified, so code will typically center and reorient. Following is for gauche difluoroethane, FCH_2CH_2F.

```
 $CONTRL COORD=CART $END
 $DATA
FCH2CH2F drag calculation
C1
C     6.0    -3.76764     0.33879     0.03727
C     6.0    -2.35246     0.34495     0.03689
F     9.0    -4.72277     0.58147    -1.18012
F     9.0    -1.59909    -0.68487    -0.83662
H     1.0    -4.04387     1.08375     0.75395
H     1.0    -3.92958    -0.71060     0.16941
H     1.0    -2.03786     0.18875     1.04760
H     1.0    -2.09983     1.28759    -0.40187
 $END
```

7.4.2 Z-matrix

Specify atom name, number of atom it is connected to, distance to that atom, number of atom it makes an angle with, value of the angle, number of the atom it makes a dihedral with, and value of the dihedral angle. Values may be given directly or as variables, followed by list of variable specifications. In `Gamess`, would specify z-matrix for FCH_2CH_2F like this:

```
$CONTRL COORD=ZMT $END
$DATA
FCH2CH2F drag calculation
C1
C
C   1   r1
F   2   r2   1   A1
H   2   r3   1   A2   3   D1
H   2   r4   1   A3   3   D2
F   1   r2   2   A1   3   D3
H   1   r3   2   A2   6   D1
H   1   r4   2   A3   6   D2

r1=1.5386
r2=1.39462
r3=1.11456
r4=1.12
A1=109.54214
A2=111.
A3=110.
D1=120.
D2=-120.5
D3=50.
$END
```

Particularly convenient when you'd like to "scan" over the value of some coordinate. Variable can be applied to more than one independent coordinate, if the molecule has symmetry. In general, though, variables should not be reused.

7.5 Specify calculation type

Specified in $CONTRL group by RUNTYP flag:

Calculation	RUNTYP=
Single-point energy	Energy
Single-point energy + force	Gradient
Geometry optimization	Optimize
Frequency calculation	Hessian
Transition state search	Sadpoint
Intrinisc reaction coordinate	IRC

Note too that specifying EXETYP=CHECK will check your input without actually running the job.

7.6 Geometry optimization coordinate system

By default, `GAMESS` performs optimizations in Cartesian coordinates.

To use z-matrix coordinates, specify

```
$CONTRL COORDS=ZMT NZVAR=xx $END
```

where NZVAR = 3N-6.

To have `Gamess` automatically create a set of appropriate "delocalized" internal coordinates, specify

```
$ZMAT DLC=.TRUE.  AUTO=.TRUE. $END
```

7.7 Performing a transition state search

Transition states are always more challenging to find than optimal geometries. Always `at least` a three-step procedure:

1. Guess a transition state structure and compute Hessian matrix (RUNTYP = HESSIAN). Check to be sure there is one imaginary mode approximating desired TS.
2. Perform the transition state search. Must supply computed Hessian matrix, copied from .dat file, as $HESS group. Also should specify $STATPT HESS=READ $END.
3. Check the result! Assuming the transition state search converges, run another frequency calculation to confirm that there is one and only one imaginary mode, and that it corresponds to desired TS.

8 Plane waves and core potentials

8.1 Hydrogen atom in a box

8.2 Periodic boundary conditions

8.3 Supercells - Cartesian and fractional coordinates

8.4 Gaussian vs. Vasp

8.5 Vasp POSCAR

8.6 Vasp INCAR

8.7 Core electron treatment

8.7.1 OPW

8.7.2 PP

8.7.3 PAW

8.8 Comparing energies between calculations

8.9 Wavefunctions and charge densities

8.10 Exploring potential energy surfaces

9 Periodic electronic structure

9.1 Isolated vs. periodic systems

9.2 Bloch's theorem and qualitative band structure

9.3 Band folding

9.4 Multi-dimensional periodicity

9.5 Density of statues

9.6 Bravais lattices

9.7 Quantitative supercell calculations

9.8 Brillouin zone integration

9.8.1 k-point sampling

9.8.2 Fermi smearing

10 Practical supercell calculations

11 Surfaces

11.1 Surface planes

11.2 Slab models

11.3 Surface energy

11.4 Surface potentials and Fermi energies

11.5 Surface adsorption

11.6 Coverage-dependent adsorption

11.7 Reaction barriers

12 First-principles thermodynamics

12.1 Connection Between QM and Thermodynamics

We have focused to this point on the many approaches and details of calculating the *internal electronic energy* of a single molecule, that is, the energy associated with taking infinitely separated constituent nuclei and electrons at rest and forming a molecule:

$$2\ \mathrm{H}^+ + 8\ \mathrm{O}^{8+} + 10\ \mathrm{e}^- \rightarrow \mathrm{H_2O} \qquad E^{\mathrm{elec}} \tag{28}$$

E^{elec} is typically calculated within the Born-Oppenheimer approximation, i.e.\ within the approximation that the nuclei are fixed in space at the minimum energy configuration. Even at 0~K, by quantum mechanics the atoms must vibrate about this minimum, and this intrinsic vibration imparts a *zero-point vibrational energy* (ZPVE) to the molecule, and the 0~K internal energy of a molecule is thus:

$$E^0 = E^{\mathrm{elec}} + ZPVE \tag{29}$$

ZPVE can be calculated reliably within the harmonic approximation, according to

$$\mathrm{ZPVE} = \frac{1}{2} h \sum_{i=1}^{3n-6} \nu_i \tag{30}$$

where ν_i are the harmonic vibrational frequencies, obtained from a vibrational frequency analysis. E^0 is the minimum physically meaninfful energy of the molecule.

Energy can be deposited in a molecule in many other ways as well, e.g.\ as translational and rotational kinetic energy, in excited vibrational modes, in the interaction of a molecule with an external electric or magnetic or gravitational field, or .... If we assume that the energy in these various degrees of freedom are separable, we can write:

$$E_i = E^0 + E^{\mathrm{trans}} + E^{\mathrm{rot}} + E^{\mathrm{vib}} + E^{\mathrm{elec*}} + E^{\mathrm{ext}} \tag{31}$$

To fully describe microscopic energetic state of a molecule, would have to specify all of these.

Typically, though, we are more interested in the collective properties of many molecules at equilibrium, like the internal energy U or enthalpy H or Gibbs energy G, under some external constraints like temperature T or volu me V. These thermodynamic quantities are averages over the energy states of an *ensemble* of molecules. The way this averaging is performed is the realm of *statistical thermodynamics.*

Most important for us will be the *canonical ensemble*, in which the free variables are the number of molecules N, the total volume V, and the temperature T. Offer without proof, in the canonical ensemble the probability for a molecule to be in some energy state E_i above E^0 is given by the Boltzmann factor,

$$P(E_i) \propto e^{-E_i\beta} = e^{-E_i/k_BT}, \qquad \beta = 1/k_BT \tag{32}$$

Defines an exponentially decaying probability function for a state to be occupied at some temperature. In a sense, *temperature* is the property of a system following this distribution.

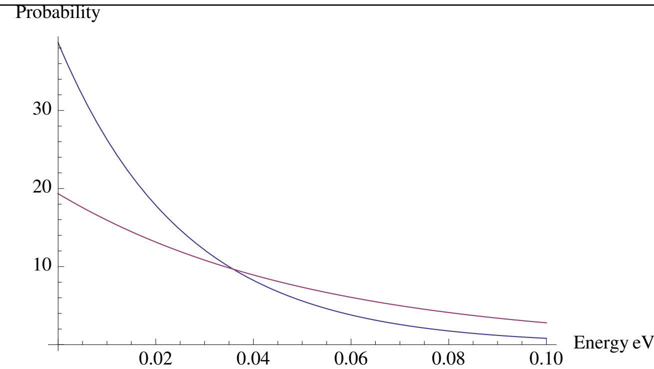


Figure 3: Boltzmann distribution at two different temperatures

12.1.1 Averages and partition functions

Let's use this to calculate the internal energy U of a molecule at some temperature.

$$U(T) = \frac{\sum_i E_i P(E_i)}{\sum_i P(E_i)} \tag{33}$$

where the denominator ensures that the probability is normalized.

$$\begin{aligned} U(T) &= \frac{\sum_i E_i e^{-E_i\beta}}{\sum_i e^{-E_i\beta}} && (34)\\ &= \frac{\frac{\partial}{\partial\beta}\sum_i e^{-E_i\beta}}{\sum_i e^{-E_i\beta}} && (35)\\ &= -\frac{\partial \ln \sum_i e^{-E_i\beta}}{\partial\beta} && (36) \end{aligned}$$

The sum over energy states is evidently a special quantity, called the partition function:

$$q = \sum_i e^{-E_i\beta} \tag{37}$$

All thermodynamic quantities can be written in terms of the partition function!

12.1.2 Harmonic oscillator example

Harmonic oscillator is a reasonable model of a molecular vibration. Energy spectrum given by

$$E_v = (v + 1/2)h\nu, \qquad v = 0, 1, 2, ... \tag{38}$$

Let's define the energy qunatum $h\nu = \epsilon_0$ and reset the energy scale so that zero is at $1/2h\nu$:

$$\begin{aligned} E_v &= v\epsilon_0, \qquad v = 0, 1, 2, ... && (39)\\ q(T) &= \sum_{v=0}^{\infty} e^{-v\epsilon_0\beta} && (40)\\ &= \frac{1}{1 - e^{-\epsilon_0\beta}} && (41) \end{aligned}$$

where we take advantage of the fact that the sum is a geometric series to evaluate it in closed form.

Plot partition function vs T, increasing function.

Internal energy:

$$U(T) = -\frac{\partial \ln q}{\partial \beta} \tag{42}$$
$$= \frac{\epsilon_0}{e^{\epsilon_0 \beta} - 1} \tag{43}$$

Heat capacity:

Entropy

12.2 Molecular Ideal Gas

Nice example above for a simple model. To get thermodynamics of an ideal gas, in principle need to sum over all the types of energy states (translational, rotational, vibrational, ...) of every molecule. Seemingly impossible task. One simplification is if we can write energy as sum of energies of individual elements (molecules) of system:

$$E_j = \epsilon_j(1) + \epsilon_j(2) + ... + \epsilon_j(N) \tag{44}$$
$$Q(N, V, T) = \sum_j e^{-E_j \beta} \tag{45}$$
$$= \sum_j e^{-(\epsilon_j(1)+\epsilon_j(2)+...+\epsilon_j(N))\beta} \tag{46}$$

If molecules/elements of system can be distinguished from each other (like atoms in a fixed lattice), expression can be factored:

$$Q(N, V, T) = \left(\sum_j e^{-\epsilon_j(1)\beta}\right) \cdots \left(\sum_j e^{-\epsilon_j(N)\beta}\right) \tag{47}$$
$$= q(1) \cdots q(N) \tag{48}$$

Assuming all the elements are the same: (49)

$$= q^N \tag{50}$$

If not distinguishable (like molecules in a liquid or gas, or electrons in a solid), problem is difficult, because identical arrangements of energy amongst elements should only be counted once. Approximate solution, good almost all the time:

$$Q(N, V, T) = q^N / N! \tag{51}$$

Sidebar: "Correct" factoring depends on whether individual elements are fermions or bosons, leads to funny things like superconductivity and superfluidity.

This $q(V, T)$ is the *molecular partition function*, and is calculated by summing over the individual energy states of a single molecule (starting at E_0).

Further simplified by factoring into contributions from various ($3N$) molecular degrees of freedom:

$$q(V, T) = \left(\sum_{\text{trans}} e^{-e_{\text{trans}}\beta}\right)\left(\sum_{\text{rot}} e^{-e_{\text{rot}}\beta}\right)\left(\sum_{\text{vib}} e^{-e_{\text{vib}}\beta}\right)\left(\sum_{\text{elec}} e^{-e_{\text{elec}}\beta}\right) \tag{52}$$
$$= q_{\text{trans}} q_{\text{rot}} q_{\text{vib}} q_{\text{elec}} \tag{53}$$
$$U = E_0 + U_{\text{trans}} + U_{\text{rot}} + U_{\text{vib}} + U_{\text{elec}} \tag{54}$$

Similarly for other thermodynamic quantities, for example,

$$C_v = \left(\frac{\partial U}{\partial T}\right)_V = C_{v,\mathrm{trans}} + C_{v,\mathrm{rot}} + C_{v,\mathrm{vib}} + C_{v,\mathrm{elec}} \tag{55}$$

Thermodynamic quantities are sums of contributions from indvidual degrees of freedom.

Have to somehow *model* these motions and have to use our quantum mechanical results to parameterize the models.

12.2.1 Translational partition function

Need a model molecules freely translating about in a box. How about the *particle in a box*?

$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2}, \qquad n = 1, 2, 3, ... \tag{56}$$

Have to construct partition function for one molecule. For gas molecules at normal conditions, energy spacing is tiny. Spare the details, but find that q_{trans} can be written in terms of a *thermal wavelength* Λ:

$$\Lambda = \frac{h}{\sqrt{2\pi m k_B T}} \tag{57}$$

$$q_{\mathrm{trans}} = \frac{V^\circ}{\Lambda^3} \tag{58}$$

Λ depends only a molecule mass (that's easy!) and is of the order the box dimensions at which quantization is evident. Typically a tiny number (e.g. \$1.7× 10^{-11}\$~m for Ar in a 1 liter volume at 298 K. q_{trans} is, on the other hand, enormous: lots of translational freedom. V° defines the standard state volume.

Given this, can find all translational contributions to thermodynamics. S_{trans} gives the Sackur-Tetrode equation, the absolute entropy of a monatomic gas:

12.2.2 Rotational partition function

Model molecule as a rigidly rotating body. Body has three orthogonal moments of inertia I determined by the molecular structure.

12.2.3 Vibrational partition function

See harmonic oscillator above

12.2.4 Electronic partition function

Governed by Fermi-Dirac distribution Electronic degeneracy

12.3 CO T-dependent thermo example

13 Implicit solvation

14 Density functional theory

14.1 Electron density ρ as fundamental quantity

14.2 Thomas-Fermi-Dirac model

14.3 Hartree-Fock-Slater model

14.4 Hohenberg-Kohn theorems

14.5 Kohn-Sham construction

14.6 Exchange-correlation functionals

So Kohn et al. showed that the DFT approach is theoretically well-grounded and provided one way to practically apply it. Promise is that if we can find an approximation to the (unknown) true v_{xc} with the right balance of simplicity and accuracy, we will have one sweet theory. Has to incorporate both exchange, like Slater tried to do, and correlation.

How to proceed? Lots of approaches, and jargon here is at least as bad as in wavefunction-based methods. Perdew 2006 describes the "Jacob's ladder" of approximations:

14.6.1 LDA

One well-defined limit is the homogeneous electron gas, and this is the usual starting point for modern approximate DFT methods. Assume exchange and correlation potentials at any given point depend only on the value of ρ there (or spin-up and spin-down ρ, if spin-polarized). We know from Slater and Dirac's work what the exchange potential is for this system.

It is possible to determine numerically the correlation energy for a given density from quantum Monte Carlo calculations. Ceperley and Alder (PRL 1980, 45, 566) did this to very high accuracy, and others (Vosko, Wilk, and Nusair, "VWN", and Perdew and Wang, "PW") fit these numerical results to analytical models in ρ. This combination of local exchange and correlation defines the LDA model.

LDA offers modest improvement over HFS for molecules. "Homogeneous" approximation pretty severe for an atom or molecule. Nonetheless, works surprisingly well for structures and charge distributions, but has problems in calculating accurate bond energies, typically overbinding. Also tends to underestimate the HOMO-LUMO gap in molecules and analogous band gap in solids.

14.6.2 GGA

14.6.3 Meta-GGA

14.6.4 Hyper GGA and hybrid functionals

- "Screened" exchange

14.6.5 Beyond hyper GGA

14.7 Implementations

14.8 Performance

15 Electron correlation methods