

Mathematical Methods and the Born-Oppenheimer Approximation

The “degree of difficulty” is given in parentheses after each problem. Your total score on each problem will be the total number of points you earn on the problem (out of twenty) times the degree of difficulty. The average degree of difficulty will usually be about 1.0 (and is occasionally a bit more). You must choose **two** problems off this set.

I. The Dirac Delta Function (.8)

On the first day of class, I talked about the Dirac Delta function. The defining relation of the Dirac delta function is that

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx \equiv f(x_0) \quad (1.1)$$

A. Show that the Dirac delta function is normalized. That is, show that

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx \equiv 1 \quad (1.2)$$

Next, recall that any function can be expanded in terms of a complete set of orthonormal basis functions—for instance, the eigenfunctions of the harmonic oscillator Hamiltonian (or, for that matter, any Hamiltonian).

B. Let $\{\psi_n(x)\}_{n=0}^{\infty}$ be a complete set of orthonormal functions. Show that the Dirac delta function is given by

$$\sum_{n=0}^{\infty} \psi_n(x_0) \psi_n(x) \equiv \delta(x - x_0). \quad (1.3)$$

The left-hand-side of Eq. (1.3) is called the Poisson kernel. If the sum doesn’t extend all the way to infinity, it is called the Christoffel-Darboux kernel.

C. Use integration by parts to show that

$$\int_{-\infty}^{\infty} f(x) \frac{d(\delta(x - x_0))}{dx} dx \equiv - \left. \frac{df(x)}{dx} \right|_{x=x_0} = -f'(x_0) \quad (1.4)$$

In three dimensions, we have

$$\delta(\mathbf{r} - \mathbf{r}_0) \equiv \delta(x - x_0) \delta(y - y_0) \delta(z - z_0) \quad (1.5)$$

D. Confirm that

$$\int f(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_0) d\mathbf{r} = f(\mathbf{r}_0). \quad (1.6)$$

(Here, and in the following, integrations without limits are assumed to be over all space. So

$$\int f(\mathbf{r}) d\mathbf{r} \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) dx dy dz \quad (1.7)$$

E. In Spherical coordinates,

$$\delta(\mathbf{r} - \mathbf{r}_0) = \frac{1}{r^2 \sin \theta} \delta(r - r_0) \delta(\theta - \theta_0) \delta(\phi - \phi_0) \quad (1.8)$$

Explain why the $\frac{1}{r^2 \sin \theta}$ term is needed.

When I introduced the Dirac delta function, I said that

$$\int \rho(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_0) d\mathbf{r} = \rho(\mathbf{r}_0) \quad (1.9)$$

represented the probability of observing an electron at \mathbf{r}_0 , if $\rho(\mathbf{r})$ is the electron density—the distribution function for electrons.

F. Let $P(\mathbf{r})$ be the charge density for a single classical particle. What sort of physical system is characterized by

$$P(\mathbf{r}) \equiv Ze\delta(\mathbf{r} - \mathbf{r}_0)? \quad (1.10)$$

The Dirac delta function is usually constructed using “delta function sequences”. The simplest such sequence is the function

$$f_a(x) = \begin{cases} 0 & x < -a \\ \frac{1}{2a} & -a \leq x \leq a \\ 0 & x > a \end{cases} \quad (1.11)$$

G. Show that

$$\lim_{a \rightarrow 0^+} f_a(x) \rightarrow \delta(x) \quad (1.12)$$

by showing that this function satisfies the key relation, Eq. (1.1).

From this result, do you see why I said that Eq. (1.9) represents the “expectation value” for observing an electron at the point \mathbf{r}_0 ?

Hint for part **G**: Remember, from calculus, that the average value of a function on an interval is

$$\text{average value of } f \text{ in the interval } [a, b] \equiv \bar{f} \equiv \frac{1}{b-a} \int_a^b f(x) dx \quad (1.13)$$

Equation (1.13) is easy to remember: it is just the continuous version of the formula for the mean of a data set,

$$\bar{f} = \frac{1}{N} \sum_{n=1}^N f_n \quad (1.14)$$

II. Eigenfunctions, Eigenvalues, and the Variational Principle (.7)

Let \hat{A} be a Hermitian operator. For the purposes of this exercise, you may assume that \hat{A} is a real-valued operator, whence all its eigenfunctions can be chosen to be real.

A. Show that if $\psi_i(x)$ and $\psi_j(x)$ are eigenfunctions of this operator with different eigenvalues, then they are orthogonal. That is,

$$\int \psi_i(x) \psi_j(x) dx = 0 \quad (11.1)$$

- B. Suppose $\psi_i(x)$ and $\psi_j(x)$ are normalized eigenfunctions with the same eigenvalue, and that $\psi_i(x) \neq \psi_j(x)$. Show that we can find a two linear combinations of these two functions, $\chi_a(x) = c_{ai}\psi_i(x) + c_{aj}\psi_j(x)$ and $\chi_b(x) = c_{bi}\psi_i(x) + c_{bj}\psi_j(x)$ so that

$$\int \chi_a(x) \chi_b(x) dx = 0 \quad (\text{II.2})$$

- C. From A and B, we can prove that the eigenvectors of a Hermitian operator can always be chosen so that they are orthogonal. Why?

A very important result in chemistry is the variational principle. Namely, the ground state wave function is found by minimizing $\frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}$ with respect to all wave functions.

- D. Prove that

$$E_{\text{ground state}} = \min_{\Psi} \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (\text{II.3})$$

Hint: any wave function can be written as

$$\Psi(\mathbf{r}) = \sum_{i=0}^{\infty} c_i \psi_i(\mathbf{r}) \quad (\text{II.4})$$

where $\psi_i(\mathbf{r})$ are the eigenfunctions of \hat{H} .

- E. Now, show that we can find the energy of the first excited state by minimizing the energy with respect to all wave functions that are orthogonal to the ground state. That is,

$$E_{\text{first excited state}} = \min_{\{\Psi | \langle \Psi | \Psi_{g.s.} \rangle = 0\}} \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (\text{II.5})$$

III. Some Identities in Vector Calculus (.8)

In class, I said that

$$\nabla^2 f(\mathbf{r}) g(\mathbf{r}) \equiv f(\mathbf{r}) \nabla^2 g(\mathbf{r}) + 2 \nabla g(\mathbf{r}) \cdot \nabla f(\mathbf{r}) + g(\mathbf{r}) \nabla^2 f(\mathbf{r}) \quad (\text{III.1})$$

- A. Derive Eq. (III.1)

The generalization to three function is sometimes useful

$$\begin{aligned} \nabla^2 f(\mathbf{r}) g(\mathbf{r}) h(\mathbf{r}) &\equiv g(\mathbf{r}) h(\mathbf{r}) \nabla^2 f(\mathbf{r}) + f(\mathbf{r}) h(\mathbf{r}) \nabla^2 g(\mathbf{r}) + f(\mathbf{r}) g(\mathbf{r}) \nabla^2 h(\mathbf{r}) \\ &\quad + 2h(\mathbf{r}) \nabla f(\mathbf{r}) \cdot \nabla g(\mathbf{r}) + 2g(\mathbf{r}) \nabla f(\mathbf{r}) \cdot \nabla h(\mathbf{r}) + 2f(\mathbf{r}) \nabla g(\mathbf{r}) \cdot \nabla h(\mathbf{r}) \end{aligned} \quad (\text{III.2})$$

There are some very useful identities from multivariate calculus that you should remember. Chief among these is Gauss's divergence theorem,

$$\int \nabla \cdot \mathbf{F}(\mathbf{r}) d\mathbf{r} = \oint_{R \rightarrow \infty} \mathbf{A} \cdot \mathbf{n} da \quad (\text{III.3})$$

If you don't remember what a surface integral (that thing on the right side of Eq. (III.3)) is, look it up. The key idea is to parametrize the surface as

$$\mathbf{S}(u, v) \equiv x(u, v) \mathbf{i} + y(u, v) \mathbf{j} + z(u, v) \mathbf{k} \quad (\text{III.4})$$

(\mathbf{i} , \mathbf{j} , and \mathbf{k} are the unit vectors in the x , y , and z directions, respectively). Then $\mathbf{N}[\mathbf{u}, \mathbf{v}] \equiv \frac{\partial \mathcal{S}}{\partial \mathbf{u}} \times \frac{\partial \mathcal{S}}{\partial \mathbf{v}}$. Then

$$\oint\!\!\!\oint \mathbf{A} \cdot \mathbf{n} da \equiv \iint \mathbf{F}(u, v) \cdot \mathbf{N}(u, v) du dv, \quad (\text{III.5})$$

where the double integral is performed over the entire region of space.

The main case we'll use in this course is when integration is over all of space, and we denote this with the notation $R \rightarrow \infty$ on the right-hand-side of Eq. (III.3). In this case, the result simplifies to

$$\oint\!\!\!\oint_{R \rightarrow \infty} \mathbf{A} \cdot \mathbf{n} da = \int_0^\pi \int_0^{2\pi} \lim_{r \rightarrow \infty} [(\mathbf{A} \cdot \hat{\mathbf{r}}) r^2] \sin \theta \, d\theta d\phi \quad (\text{III.6})$$

where $\hat{\mathbf{r}}$ is the unit vector in the outward-looking radial direction.

Similarly important are Green's first identity

$$\iiint \phi(\mathbf{r}) \nabla^2 \psi(\mathbf{r}) + \nabla \phi(\mathbf{r}) \cdot \nabla \psi(\mathbf{r}) d\mathbf{r} \equiv \oint\!\!\!\oint \phi(\mathbf{r}) \nabla \psi(\mathbf{r}) \cdot \mathbf{n} da \quad (\text{III.7})$$

and Green's theorem

$$\iiint \phi(\mathbf{r}) \nabla^2 \psi(\mathbf{r}) - \psi(\mathbf{r}) \nabla^2 \phi(\mathbf{r}) d\mathbf{r} \equiv \oint\!\!\!\oint (\phi(\mathbf{r}) \nabla \psi(\mathbf{r}) - \psi(\mathbf{r}) \nabla \phi(\mathbf{r})) \cdot \mathbf{n} da \quad (\text{III.8})$$

B. Derive Green's theorem (III.8) from Green's first identity (III.7)

Equations (III.7) and (III.8) are the three-dimensional analogues of integration by parts. Eq. (III.3) is the analogue of the fundamental theorem of calculus.

C. Show that each of the following formulae gives the correct kinetic energy for a one-particle system

$$\int \frac{\hbar^2 \nabla \Psi^*(\mathbf{r}) \cdot \nabla \Psi(\mathbf{r})}{2m} d\mathbf{r} = \int -\frac{\hbar^2 \Psi^*(\mathbf{r}) \nabla^2 \Psi(\mathbf{r})}{2m} d\mathbf{r} = T[\Psi] \quad (\text{III.9})$$

D. Similarly, if the wave function is real, so that $\Psi^*(\mathbf{r}) = \Psi(\mathbf{r})$, show that

$$\begin{aligned} T[\Psi] &= \frac{\hbar^2}{8m} \int \frac{\nabla |\Psi(\mathbf{r})|^2 \cdot \nabla |\Psi(\mathbf{r})|^2}{|\Psi(\mathbf{r})|^2} d\mathbf{r} \\ &= -\frac{\hbar^2}{8m} \int |\Psi(\mathbf{r})|^2 \nabla^2 (\ln |\Psi(\mathbf{r})|^2) d\mathbf{r} \end{aligned} \quad (\text{III.10})$$

In parts C and D, you'll find it helpful to assume that the integral extends over all space and that the wave function is normalized. Because

$$\int |\Psi(\mathbf{r})|^2 d\mathbf{r} = 1, \quad (\text{III.11})$$

the wave function must decay faster than $\Psi(\mathbf{r}) \sim \frac{1}{r^{3/2}}$ for large r . To see this, define the spherical average of $\Psi(\mathbf{r})$ with

$$\overline{\Psi(r)} \equiv \iint \Psi(\mathbf{r}) \sin \theta \, d\theta d\phi \quad (\text{III.12})$$

E. Show that if $\overline{\Psi(r)}$ decays as $\frac{1}{r^{3/2}}$ for large r , then

$$\int |\Psi(\mathbf{r})|^2 d\mathbf{r} \rightarrow \infty \quad (\text{III.13})$$

Then, using

$$\frac{\partial |\Psi(r, \theta, \phi)|^2}{\partial r} \equiv \nabla |\Psi(\mathbf{r})|^2 \cdot \hat{\mathbf{r}}, \quad (\text{III.14})$$

we have that

$$\left| \frac{\partial |\Psi(\mathbf{r})|^2}{\partial r} \right| < \left| \frac{\partial \left(\frac{1}{r^3} \right)}{\partial r} \right| < \left| \frac{1}{r^4} \right| \quad (\text{III.15})$$

at long range. This can be used to

F. Show that

$$\int \nabla^2 |\Psi(\mathbf{r})|^2 d\mathbf{r} = 0. \quad (\text{III.16})$$

Equation (III.16) is needed to prove the desired results, Eqs. (III.9) and (III.10).

In three dimensions,

$$\nabla^2 \left(\frac{1}{r} \right) = -4\pi \delta(\mathbf{r}) \quad (\text{III.17})$$

This equation is often cited as the “reason” electronic and gravitational forces are inverse-square forces (and thus have $\frac{1}{r}$ potentials). If, as many physicists now believe, our universe has more than 3 “spatial” dimensions, Eq. (III.17) would not be valid and we might observe this effect by an as-yet unobserved perturbation away from inverse-square forces. (A second, equivalent, consequence would be that photons would have mass.) Similarly, if even the most elementary particles in the universe are not point-charges and point-masses (many physicists believe they look like high-dimension strings, membranes, etc.), then Eq. (III.17) would not represent the potential due to an elementary charged particle, and we might observe deviations from inverse-square forces. In fact, the analogue of Eq. (III.17) in $d \geq 3$ dimensions is

$$\nabla^2 \left(\frac{1}{r^{d-2}} \right) = \frac{-1}{r^{d-2}} \Omega_d \delta(\mathbf{r}) \quad (\text{III.18})$$

where

$$\Omega_d = \int_0^{2\pi} \int_0^\pi \dots \int_0^\pi \int_0^\pi \sin(\theta_1) \sin^2(\theta_2) \dots \sin^{d-2}(\theta_{d-2}) d\theta_{d-2} d\theta_{d-3} \dots d\theta_2 d\phi \quad (\text{III.19})$$

If the physicists are right, we should observe that gravity and electromagnetic forces are slightly weaker (slightly shorter ranged) than their commonly supposed $\frac{1}{r^2}$ dependence.

G. Show that Eq. (III.18) is the same as (III.17) when $d = 3$.

At some level, then, Equation (III.17) is the reason our universe acts the way it does. We should derive this result.

H. First, show that (in any number of dimensions)

$$\nabla \left(\frac{1}{|\mathbf{r}|} \right) = - \frac{\mathbf{r}}{|\mathbf{r}|^3}. \quad (\text{III.20})$$

In three dimensions, the Laplacian of a function that does not depend on the angular coordinates, θ and ϕ , can be written

$$\nabla^2 f(r) = \left(\frac{1}{r^2} \right) \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} \quad (\text{III.21})$$

I. Show that $\nabla^2 \left(\frac{1}{r} \right) = 0$ unless $r = 0$, where it is indeterminate.

J. Complete the derivation of Eq. (III.17) by showing that

$$\iiint \nabla^2 \frac{1}{r} d\mathbf{r} = -4\pi \quad (\text{III.22})$$

Since $\nabla^2 \left(\frac{1}{r} \right) = 0$ except when $r = 0$, but is normalized to -4π , Equation (III.17) must hold.

IV. The Born-Oppenheimer Approximation (.7)

The goal of this exercise is to show that the ground state total energy computed using the Born-Oppenheimer approximation is a lower bound to the true energy.

Let's consider the exact ground state wave function, $\Psi_{exact}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$, which solves

$$(\hat{T}_n + \hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}))\Psi_{exact} = E_{exact}\Psi_{exact}. \quad (\text{IV.1})$$

From the variational principle, the ground state energy could also be found by

$$E_{exact} = \min_{\Psi} \frac{\langle \Psi | \hat{T}_n + \hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) | \Psi \rangle_{\mathbf{r}, \mathbf{R}}}{\langle \Psi | \Psi \rangle_{\mathbf{r}, \mathbf{R}}}. \quad (\text{IV.2})$$

The notation $\langle | \rangle_{\mathbf{r}, \mathbf{R}}$ indicates that the integration variables include both the electronic and the nuclear coordinates.

By comparison, in the Born-Oppenheimer approximation, we approximate the wave function with

$$\Psi^{BO} \equiv \psi_e^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) \chi_n^{BO}(\{\mathbf{R}_\alpha\}) \quad (\text{IV.3})$$

and determine ψ_e using the “electronic Schrödinger equation”

$$(T_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}))\psi_e^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) = U^{BO}(\{\mathbf{R}_\alpha\})\psi_e^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}). \quad (\text{IV.4})$$

For any given set of nuclear positions, $\{\mathbf{R}_\alpha\}$ is constant, and we could use the variational principle to determine the potential energy surface, $U(\{\mathbf{R}_\alpha\})$,

$$U^{BO}(\{\mathbf{R}_\alpha\}) = \min_{\psi_e} \frac{\langle \psi_e | \hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) | \psi_e \rangle_{\mathbf{r}}}{\langle \psi_e | \psi_e \rangle_{\mathbf{r}}} \quad (\text{IV.5})$$

Unlike the case in Eq. (IV.2), here the notation, $\langle | \rangle_{\mathbf{r}}$, indicates that the expression is only integrated with respect to electronic coordinates.

We can consider $\Psi_{exact}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$ to be an “approximate wave function” for the electronic Hamiltonian.

Then, from Eq. (IV.5), we have that

$$\frac{\langle \Psi_{exact} | \hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) | \Psi_{exact} \rangle_{\mathbf{r}}}{\langle \Psi_{exact} | \Psi_{exact} \rangle_{\mathbf{r}}} \geq \frac{\langle \psi_e | \hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) | \psi_e \rangle_{\mathbf{r}}}{\langle \psi_e | \psi_e \rangle_{\mathbf{r}}} = U(\{\mathbf{R}_\alpha\}) \quad (\text{IV.6})$$

and so

$$\langle \Psi_{exact} | \hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) | \Psi_{exact} \rangle_{\mathbf{r}} \geq U(\{\mathbf{R}_\alpha\}) \langle \Psi_{exact} | \Psi_{exact} \rangle_{\mathbf{r}}. \quad (\text{IV.7})$$

Now, using the fact that

$$E_{exact} = \frac{\langle \Psi_{exact} | \hat{T}_n + \hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) | \Psi_{exact} \rangle_{\mathbf{r}, \mathbf{R}}}{\langle \Psi_{exact} | \Psi_{exact} \rangle_{\mathbf{r}, \mathbf{R}}}, \quad (\text{IV.8})$$

A. Show that

$$E_{exact} \langle \Psi_{exact} | \Psi_{exact} \rangle_{\mathbf{r}, \mathbf{R}} \geq \langle \Psi_{exact} | \hat{T}_n + U^{BO}(\{\mathbf{R}_\alpha\}) | \Psi_{exact} \rangle_{\mathbf{r}, \mathbf{R}} \quad (\text{IV.9})$$

(Hint: use Eqs. (IV.7) and (IV.8).)

Now, recall that the nuclear wave function in the Born-Oppenheimer approximation is determined from the Schrödinger equation,

$$(\hat{T}_n + U^{BO}(\{\mathbf{R}_\alpha\}))\chi_n^{BO}(\{\mathbf{R}_\alpha\}) = E^{BO}\chi_n^{BO}(\{\mathbf{R}_\alpha\}). \quad (\text{IV.10})$$

B. Write the variational principle corresponding to Eq. (IV.10).

C. Use Equation (IV.9) and the result from part B to show that

$$E_{exact} \geq E^{BO}. \quad (\text{IV.11})$$

V. The Born-Huang Approximation: the importance of non-adiabatic effects (1.0)

Sometimes the Born-Oppenheimer approximation just isn't good enough. In that case, one often uses the Born-Huang approximation, which is a "diagonal correction" to the Born-Oppenheimer approximation. In this problem we'll derive the Born-Huang approximation and discuss its implications for chemical reactivity.

As you know, any wave function can be expanded in terms of a complete set of basis functions. Since we are interested in "correcting" the Born-Oppenheimer approximation, we consider the eigenfunctions of the Born-Oppenheimer Hamiltonians,

$$[\hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})] \psi_m^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) = U_m^{BO}(\{\mathbf{R}_\alpha\}) \psi_m^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) \quad (\text{V.1})$$

and

$$(\hat{T}_n + U_m^{BO}(\{\mathbf{R}_\alpha\})) \chi_{m,\mu}^{BO}(\{\mathbf{R}_\alpha\}) = E_{m,\mu}^{BO} \chi_{m,\mu}^{BO}(\{\mathbf{R}_\alpha\}). \quad (\text{V.2})$$

In these equations, $\{\mathbf{r}_i\}_{i=1}^N$ denote the positions of the electrons, $\{\mathbf{R}_\alpha\}_{\alpha=1}^P$ denote the positions of the atomic nuclei, $V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$ denotes the sum of the electron-electron, nuclear-electron, and nuclear-nuclear potential energy. For any given position of the atomic nuclei, $U_m^{BO}(\{\mathbf{R}_\alpha\})$ and $\psi_m^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$ denote the energy and wave function of the m^{th} excited state of the system. That is, for any given set of nuclear positions, $\{\mathbf{R}_\alpha\}$, $\psi_0^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$ is the ground state electronic wave function, $\psi_1^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$ is the electronic wave function of the first excited state, etc..

For any given electronic state, the nuclear wave functions are determined by solving the nuclear Schrödinger equation, Eq. (V.2). Here, $\chi_{m,0}^{BO}(\{\mathbf{R}_\alpha\})$ and $E_{m,0}^{BO}$ denote the ground state nuclear wave function and the ground state total energy (in the Born-Oppenheimer approximation, of course) for the m^{th} electronic excited state of the system, $\chi_{m,1}^{BO}(\{\mathbf{R}_\alpha\})$ and $E_{m,1}^{BO}$ denote the wave function and energy of the first nuclear excited state on the potential energy surface of the m^{th} electronic excited state, and so on. Taken together, $\psi_m^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$ and $\chi_{m,\mu}^{BO}(\{\mathbf{R}_\alpha\})$ form a complete orthonormal set. We choose to expand the exact wave function in this basis, so that

$$\Psi^{exact}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) \equiv \sum_{\mu=0}^{\infty} \sum_{m=0}^{\infty} k_{m\mu}^{exact} \cdot \psi_m^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) \chi_{m,\mu}^{BO}(\{\mathbf{R}_\alpha\}) \quad (\text{V.3})$$

where the expansion coefficients, $k_{m\mu}^{exact}$, can be determined by evaluating

$$k_{m\mu}^{exact} = \langle \psi_m^{BO} \chi_{m,\mu}^{BO} | \Psi^{exact} \rangle_{\mathbf{r}, \mathbf{R}}. \quad (\text{V.4})$$

The notation here is the same as that in problem IV: $\langle | \rangle_r$ denotes that the integration is performed only over the electronic variables (the nuclear coordinates are fixed), $\langle | \rangle_R$ denotes integration over only the nuclear coordinates (the electronic coordinates are fixed), and $\langle | \rangle_{\mathbf{r}, \mathbf{R}}$ denotes integration over both electronic and nuclear coordinates.

A. Derive Eq. (V.4).

There is a subtle point related to the above treatment. When the electronic energy is greater than zero, $\psi_m^{BO}(\{\mathbf{r}_i\}) \sim e^{i\mathbf{k}\cdot\mathbf{r}_i}$, which can not be normalized. In addition, because the “entire molecule” can move freely, the nuclear wave functions have a component (corresponding to the free translation of the center of mass) like $\chi_{m,\mu}^{BO}(\{\mathbf{R}_\alpha\}) = e^{i(\mathbf{k}\cdot\mathbf{R}_{COM})} = \cos(\mathbf{k}\cdot\mathbf{R}_{COM}) + i\sin(\mathbf{k}\cdot\mathbf{R}_{COM})$, where \mathbf{R}_{COM} is the position of the center of mass of the particle. There are several ways to solve these problems. One way is to go through the (somewhat complicated) mathematics necessary to treat “continuum wave functions”—wave functions that cannot be normalized. One can also “separate out” the motion of the center of mass from the nuclear equation (which solves the problem for the nuclei, albeit not the electrons). However, chemists tend to take a “pragmatic” approach. Put the molecule in a large sphere of radius L , where the potential inside the sphere is just that due to the electrons and nuclei ($V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$) and the potential outside the sphere is infinite for both the electrons and the nuclei. What used to be the “continuum” electronic and nuclear wave functions now resemble the states of the particle in a box. In particular, the part of the nuclear wave function associated with the center of mass is exactly that due to a particle in a box. Now, if L is taken to be very large, the likelihood of any electron or nucleus being close to the boundary of the “box” is very small, so the chemical properties of the system are unaffected by this change. Finally, taking the limit $L \rightarrow \infty$, we recover the exact result. This construction is the wave the wave functions used here should be considered.

With this ansatz, all the wave functions in the problem can be normalized. Using the variational principle, then, we can find exact wave function for the system by minimizing the energy, so

$$E_{exact} \equiv \min_{\Psi} \frac{\langle \Psi | \hat{T}_n(\{\mathbf{R}_\alpha\}) + \hat{T}_e(\{\mathbf{r}_i\}) + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) | \Psi \rangle_{\mathbf{r}, \mathbf{R}}}{\langle \Psi | \Psi \rangle_{\mathbf{r}, \mathbf{R}}}. \quad (\text{V.5})$$

If we expand Ψ in terms of the Born-Oppenheimer states,

$$\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) \equiv \sum_{\mu=0}^{\infty} \sum_{m=0}^{\infty} k_{m\mu} \cdot \psi_m^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) \chi_{m,\mu}^{BO}(\{\mathbf{R}_\alpha\}), \quad (\text{V.6})$$

we find that

$$E_{exact} \equiv \min_{\{k_{m,\mu}\}} \frac{\sum_{\mu=0}^{\infty} \sum_{m=0}^{\infty} \sum_{\nu=0}^{\infty} \sum_{n=0}^{\infty} k_{n,\nu}^* k_{m,\mu} \langle \chi_{n,\nu}^{BO} \psi_n^{BO} | \hat{T}_n(\{\mathbf{R}_\alpha\}) + U_m^{BO}(\{\mathbf{R}_\alpha\}) | \chi_{m,\mu}^{BO} \psi_m^{BO} \rangle_{\mathbf{r}, \mathbf{R}}}{\sum_{m,\mu=0}^{\infty} |k_{m,\mu}|^2} \quad (\text{V.7})$$

B. Derive Eq. (V.7).

Recall that the nuclear kinetic energy operator is

$$\hat{T}_n(\{\mathbf{R}_\alpha\}) \equiv \sum_{\alpha=1}^P \frac{-1}{2M_\alpha} \nabla_\alpha^2, \quad (\text{V.8})$$

where P is the number of nuclei and M_α is the mass number for the α^{th} nucleus. Using the result from Eq. (III.1), we can write the summand in the numerator of Eq. (V.7) as

$$\begin{aligned}
& \iiint \cdots \int \left[(\chi_{n,\nu}^{BO} \psi_n^{BO})^* \psi_m^{BO} \hat{T}_n \chi_{m,\mu}^{BO} + (\chi_{n,\nu}^{BO} \psi_n^{BO})^* \chi_{m,\mu}^{BO} \hat{T}_n \psi_m^{BO} \right] d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{R}_1 \dots d\mathbf{R}_P \\
& + \iiint \cdots \int \left[(\chi_{n,\nu}^{BO} \psi_n^{BO})^* \sum_{\alpha=1}^P \frac{-1}{M_\alpha} (\nabla_\alpha \chi_{m,\mu}^{BO} \cdot \nabla \psi_m^{BO}) \right] d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{R}_1 \dots d\mathbf{R}_P \\
& + \left\langle \chi_{n,\nu}^{BO} \psi_n^{BO} \left| U_m^{BO}(\{\mathbf{R}_\alpha\}) \right| \chi_{m,\mu}^{BO} \psi_m^{BO} \right\rangle_{\mathbf{r}, \mathbf{R}} \\
& = \delta_{nm} \delta_{\nu\mu} E_{m,\mu}^{BO} + \iiint \cdots \int (\chi_{n,\nu}^{BO} \psi_n^{BO})^* \chi_{m,\mu}^{BO} \hat{T}_n \psi_m^{BO} d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{R}_1 \dots d\mathbf{R}_P \\
& + \iiint \cdots \int \left[(\chi_{n,\nu}^{BO} \psi_n^{BO})^* \sum_{\alpha=1}^P \frac{-1}{M_\alpha} (\nabla_\alpha \chi_{m,\mu}^{BO} \cdot \nabla \psi_m^{BO}) \right] d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{R}_1 \dots d\mathbf{R}_P
\end{aligned} \tag{V.9}$$

C. Derive Eq. (V.9).

(Hint: remember that the eigenfunctions are all orthonormal.) The Kronecker delta, δ_{ij} , is just the identity matrix. That is,

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases} \tag{V.10}$$

Before going any further, let's consider the interpretation of the last two terms in Eq. (V.9). Unlike the Born-Oppenheimer approximation, in these terms the electronic excited states, $\psi_{m>0}^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$ are coupled to the ground state, $\psi_0^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$. The interpretation is that the motion of the nuclei can “induce” a transition from the ground electronic state to an electronic excited state. Noting that the energy of the nuclear energy states is essentially the energy of molecular vibrations, we expect this “transfer of energy” from the nuclei to the electrons can only occur when two electronic states are very close in energy, so that the “gap” between two electronic states is about the same size as the spacing between the vibrational energy levels in the molecule. (Remember, energy is conserved. So, if the molecule moves from the electronic ground state to an electronic excited state $\psi_0^{BO} \rightarrow \psi_{m>0}^{BO}$, the nuclear wave function must move from an excited state to the ground state $\chi_{0,\mu}^{BO} \rightarrow \chi_{m>0,\nu<\mu}^{BO}$. This is why it is only when the separation between nuclear energy levels (and the largest separations are associated with molecular vibration) and the separation between electronic energy levels (which are usually a hundred times greater than the nuclear energy spacings) are similar in size that non-adiabatic effects have chemically important consequences.)

The last term in Eq. (V.9) is called the non-adiabatic term. When it is included, one cannot separate the motion of the electrons from the motion of the nuclei. (Because both the nuclear and the electronic wave functions are operated upon, we cannot separate variables.) So let us ignore the last term in Eq. (V.9). Equation (V.7) then becomes

$$E_{exact} \equiv \min_{\{k_{m,\mu}\}} \frac{\sum_{\mu=0}^{\infty} \sum_{m=0}^{\infty} \left(|k_{m,\mu}|^2 E_{m,\mu}^{BO} + k_{n,\nu}^* k_{m,\mu} \sum_{\nu=0}^{\infty} \sum_{n=0}^{\infty} \iiint \cdots \int (\chi_{n,\nu}^{BO} \psi_n^{BO})^* \chi_{m,\mu}^{BO} \hat{T}_n \psi_m^{BO} d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{R}_1 \dots d\mathbf{R}_P \right)}{\sum_{m,\mu=0}^{\infty} |k_{m,\mu}|^2} \tag{V.11}$$

Equation (V.11) gives Golden's minimal adiabatic approximation. It is the most accurate “adiabatic” approximation known, but it is difficult to evaluate because the second term in the numerator couples different electronic states of the molecule. (This of course, is quite useful, since it means that the “physics” of the minimal adiabatic approximation is correct. In particular, unlike the simpler approximation we shall consider, Eq. (V.11) predicts the Jahn-Teller distortion of degenerate ground states.)

Born and Huang derived a simpler approximation by assuming that

$$\iiint \cdots \int (\psi_n^{BO})^* \hat{T}_n \psi_m^{BO} d\mathbf{r}_1 \dots d\mathbf{r}_N \equiv \mathcal{T}_m(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_P) \delta_{nm} \tag{V.12}$$

That is, they assumed that the kinetic energy operator does not cause transitions between electronic states. This is clearly incorrect, but it is often a good approximation.

D. Show that, assuming the Born-Huang approximation, Eq. (V.12), is valid, then

$$E^{BH} \equiv \min_{\{k_{m,\mu}\}} \frac{\sum_{\mu=0}^{\infty} \sum_{m=0}^{\infty} \sum_{\nu=0}^{\infty} k_{m,\nu}^* k_{m,\mu} \left\langle \chi_{m,\nu}^{BO} \left| \hat{T}_n(\{\mathbf{R}_\alpha\}) + U_m^{BO}(\{\mathbf{R}_\alpha\}) + \mathcal{T}_m(\{\mathbf{R}_\alpha\}) \right| \chi_{m,\mu}^{BO} \right\rangle_{\mathbf{R}}}{\sum_{m,\mu=0}^{\infty} |k_{m,\mu}|^2} \quad (\text{V.13})$$

Now, because $\{\chi_{m,\mu}^{BO}(\{\mathbf{R}_\alpha\})\}_{\mu=0}^{\infty}$ is a complete orthonormal set of functions, we can write the nuclear wave function in the Born-Huang approximation, $\chi_{m,\nu}^{BH}(\{\mathbf{R}_\alpha\})$, as

$$\chi_{m,\nu}^{BH}(\{\mathbf{R}_\alpha\}) = \sum_{\mu=0}^{\infty} k_{m,\mu}^{BH,\nu} \chi_{m,\mu}^{BO}(\{\mathbf{R}_\alpha\}). \quad (\text{V.14})$$

E. Use the variational principle to show that we can find the Born-Huang approximation by solving the Born-Huang nuclear Schrödinger equation,

$$(\hat{T}_n + U_m^{BO}(\{\mathbf{R}_\alpha\}) + \mathcal{T}_m(\{\mathbf{R}_\alpha\})) \chi_{m,\nu}^{BH}(\{\mathbf{R}_\alpha\}) = E_{m,\nu}^{BH} \chi_{m,\nu}^{BH}(\{\mathbf{R}_\alpha\}). \quad (\text{V.15})$$

The electronic wave function in the Born-Huang approximation is just the Born-Oppenheimer electronic wave function, so the total Born-Huang wave function is

$$\Psi_{m,\nu}^{BH}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) = \psi_m^{BO}(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) \chi_{m,\nu}^{BH}(\{\mathbf{R}_\alpha\}). \quad (\text{V.16})$$

One of the reasons the Born-Huang approximation is useful is because, unlike the Born-Oppenheimer approximation, the total Born-Huang energy of the ground state is an upper bound to the true energy. That is,

$$E_{0,0}^{BH} \geq E_{0,0}^{exact} \geq E_{0,0}^{BO}. \quad (\text{V.17})$$

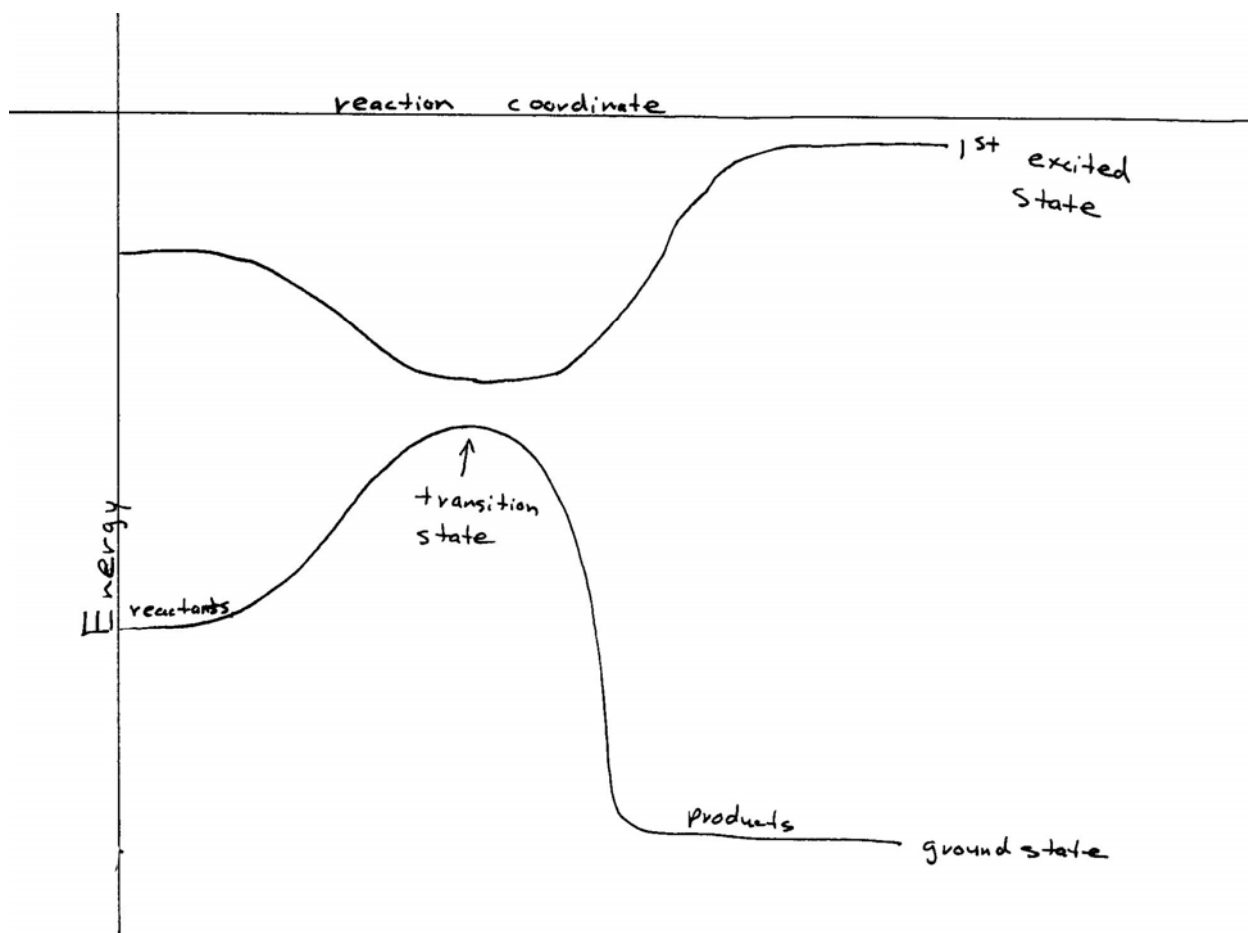
Proving this result is much more difficult than problem IV, so I won't ask you to do so. If you're interested, ask me and I'll show you how it's done.

Recall that the primary error in both the Born-Oppenheimer and the Born-Huang approximation is that we do not account for the ability for the nuclear motion to cause a transition in the electronic state of the molecule. This approximation was judged to be especially severe when the spacing between electronic states was small, so that the spacing between nuclear energy levels and electronic energy levels is similar.

If we plot the energy of the ground state and the first excited state of a substitution reaction,



along the reaction path, the curves typically look like



(The idea that more stable states have larger gaps to the first excited state is called the maximum hardness principle, and was proposed by Pearson in 1988.)

Using Eq. (IV.7), we can define the error in the Born-Oppenheimer approximation,

$$\Delta^{BO}(\{\mathbf{R}_\alpha\}) = U^{BO}(\{\mathbf{R}_\alpha\}) - \frac{\langle \Psi^{exact} | \hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) | \Psi^{exact} \rangle_r}{\langle \Psi^{exact} | \Psi^{exact} \rangle_r} \leq 0. \quad (\text{V.19})$$

It is usually approximately true that $\Delta(\{\mathbf{R}_\alpha\})$ is inversely proportional to the gap between the ground and first excited state. Assuming this to be the case

F. Compare the accuracy of the Born-Oppenheimer approximation for the transition state, reactants, and products. Where might the approximation be particularly problematic? Will the Born-Oppenheimer approximation tend to overestimate or underestimate reaction rates? What about the Born-Huang approximation?

In the Born-Huang approximation, we have

$$\Delta^{BH}(\{\mathbf{R}_\alpha\}) = U^{BO}(\{\mathbf{R}_\alpha\}) + \mathcal{T}(\{\mathbf{R}_\alpha\}) - \frac{\langle \Psi^{exact} | \hat{T}_e + V(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) | \Psi^{exact} \rangle_r}{\langle \Psi^{exact} | \Psi^{exact} \rangle_r} \geq 0. \quad (\text{V.20})$$

Usually, $|\Delta^{BH}(\{\mathbf{R}_\alpha\})| < |\Delta^{BO}(\{\mathbf{R}_\alpha\})|$. That is, the Born-Huang approximation is usually more accurate than the Born-Oppenheimer approximation. For a proton-transfer reaction (referring to Eq. (V.18), $B = H^+$), we can approximate the Born-Oppenheimer error as

$$\Delta^{BO}(\{\mathbf{R}_\alpha\}) \approx -\frac{4}{3}\left(\frac{m_e}{m_p}\right)\eta(\{\mathbf{R}_\alpha\}) = -\frac{4}{3 \cdot 1832}\frac{1}{\eta(\{\mathbf{R}_\alpha\})} \quad (\text{V.21})$$

where η is the energy gap between the ground and first excited states, **in Hartree**. The Born-Huang error can be approximate in similar fashion,

$$\Delta^{BH}(\{\mathbf{R}_\alpha\}) \approx \frac{4}{3}\left(\frac{m_e}{m_p}\right)^2\eta(\{\mathbf{R}_\alpha\}) = \frac{4}{3 \cdot (1832)^2}\frac{1}{\eta(\{\mathbf{R}_\alpha\})}. \quad (\text{V.22})$$

Based on the preceding figure, we may surmise that it is quite unlikely that non-adiabatic effects are important in the reactant or product states. However, they can be very significant near the transition state. In particular, if the error in energy is greater than $1 \frac{\text{kcal}}{\text{mol}}$, predictions of the rates of chemical reactions could be off by an order of magnitude. (At room temperature, the average kinetic energy of molecules in a gas is about $1 \frac{\text{kcal}}{\text{mol}}$, so if we miss the activation energy by this much, we make a significant error in the reaction rate.)

G. Using the approximations in Eqs. (V.21) and (V.22) calculate how small $\eta(\{\mathbf{R}_\alpha\})$ must be before the error in the approximations exceeds $1 \frac{\text{kcal}}{\text{mol}}$.

VI. The Calculus of Variations and Stationary States (.9)

You were undoubtedly told that the Schrödinger equation (or something like unto it) was the fundamental basis for quantum mechanics. Most modern physicists would disagree with this viewpoint. Rather, variational principles are considered to be much more fundamental. Especially interesting are “stationary” or “stability” principles, and the Schrödinger equation can be derived from such principles. The simplest such principle is the statement that the energy of a “stationary state” is stable with respect to small changes in the wave function. That is, if we take a wave function, Ψ and change it by some small amount, $\delta\Psi$, the energy will only change in a way proportional to the norm (= size) of $|\delta\Psi|^2$. Mathematically, we say that

$$E[\Psi + \delta\Psi] - E[\Psi] \propto \langle \delta\Psi | \delta\Psi \rangle \quad (\text{VI.1})$$

Equation (VI.1) is the reason that excited states of molecules are metastable. In fact, were it not for an exotic phenomena called vacuum polarization (which has to do with the ability of empty space to spontaneously generate photons, provided they disappear into the vacuum sufficiently quickly), the excited states of an isolated molecule would have an infinite lifetime.

Equation (VI.1) is analogous to the condition for a “point of inflection”, maximum, or minimum in a function, namely

$$f(x + dx) - f(x) \propto (dx)^2 \quad (\text{VI.2})$$

If you write $f(x)$ as a Taylor series, then Eq. (VI.2) is seen to imply that the derivative of f with respect to x is zero, which is precisely the familiar condition for stability. Similarly, when Eq. (VI.1) holds we can say that the functional derivative of E with respect to Ψ is zero.

Recall the definition of the gradient of a function, $\nabla g(\mathbf{r})$ is defined as the vector that maps changes in \mathbf{r} , $d\mathbf{r}$ to changes in $g(\mathbf{r})$, dg , via the formula

$$dg(\mathbf{r}) \equiv g(\mathbf{r} + d\mathbf{r}) - g(\mathbf{r}) = \nabla g(\mathbf{r}) \cdot d\mathbf{r} + \mathcal{O}(d\mathbf{r} \cdot d\mathbf{r}). \quad (\text{VI.3})$$

Here, $\mathcal{O}(d\mathbf{r} \cdot d\mathbf{r})$ denotes that the error in Eq. (VI.3) is of the order $d\mathbf{r} \cdot d\mathbf{r}$ —in the limit of very small $d\mathbf{r}$ such terms are negligible.

We define the functional derivative in the same way—the functional derivative of $A[\Psi]$ with respect to Ψ is defined as that function, $\frac{\delta A[\Psi]}{\delta \Psi(\boldsymbol{\tau})}$, that maps changes in Ψ , $\delta\Psi$ to changes in $A[\Psi]$, $dA[\Psi]$ via the formula

$$dA[\Psi(\boldsymbol{\tau})] \equiv A[\Psi(\boldsymbol{\tau}) + \delta\Psi(\boldsymbol{\tau})] - A[\Psi(\boldsymbol{\tau})] = \int \frac{\delta A[\Psi]}{\delta \Psi(\boldsymbol{\tau})} \delta\Psi(\boldsymbol{\tau}) d\boldsymbol{\tau} + \mathcal{O}(\langle \delta\Psi(\boldsymbol{\tau}) | \delta\Psi(\boldsymbol{\tau}) \rangle). \quad (\text{VI.4})$$

A. Choose $\delta\Psi(\tau) = \varepsilon \cdot \delta(\tau - \tau_0)$ in Eq. (VI.4). Show that

$$\frac{\delta A[\Psi]}{\delta\Psi(\tau_0)} = \lim_{\varepsilon \rightarrow 0} \frac{A[\Psi + \varepsilon\delta(\tau - \tau_0)] - A[\Psi]}{\varepsilon} = \left(\frac{dA[\Psi + \varepsilon\delta(\tau - \tau_0)]}{d\varepsilon} \right)_{\varepsilon=0} \quad (\text{VI.5})$$

Equation (VI.5) is a useful formula for computing functional derivatives. Another method follows directly from the definition: compute $A[\Psi + \delta\Psi] - A[\Psi]$ and write the result in the form of Eq. (VI.4). The coefficient of $\delta\Psi(\tau)$ is the functional derivative. Yet another method follows when we write expand the wave function in a basis set,

$$\Psi(\tau) \equiv \sum_{n=0}^{\infty} k_n \Phi_n(\tau) \quad (\text{VI.6})$$

$$\Psi(\tau) + \delta\Psi(\tau) \equiv \sum_{n=0}^{\infty} (k_n + dk_n) \Phi_n(\tau) \quad (\text{VI.7})$$

so that

$$\delta\Psi(\tau) = \sum_{n=0}^{\infty} dk_n \Phi_n(\tau). \quad (\text{VI.8})$$

Using the relationship

$$\begin{aligned} A(\mathbf{k} + d\mathbf{k}) - A(\mathbf{k}) &= \nabla_{\mathbf{k}} A(\mathbf{k}) \cdot d\mathbf{k} + \mathcal{O}(d\mathbf{k} \cdot d\mathbf{k}) \\ &= \sum_{n=0}^{\infty} \frac{\partial A}{\partial k_n} dk_n + \mathcal{O}(d\mathbf{k} \cdot d\mathbf{k}) \end{aligned} \quad (\text{VI.9})$$

one finds that

$$\frac{\delta A[\Psi]}{\delta\Psi(\tau)} = \sum_{n=0}^{\infty} \frac{\partial A}{\partial k_n} \left(\frac{\delta k_n}{\delta\Psi(\tau)} \right) = \sum_{n=0}^{\infty} \frac{\partial A}{\partial k_n} \left(\frac{\delta \langle \Psi(\tau) | \Phi_n \rangle}{\delta\Psi(\tau)} \right) = \sum_{n=0}^{\infty} \frac{\partial A}{\partial k_n} (\Phi_n). \quad (\text{VI.10})$$

B. Show that Eq. (VI.10) satisfies the definition of the functional derivative, Eq. (VI.4).

We should say a few words about what a functional is. A functional is, as should be clear from the above, just a “function of a function.” For instance, the energy is a **functional** of the wave function, because it can be computed from the wave function using

$$E[\Psi] \equiv \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (\text{VI.11})$$

To remove possible ambiguity, we denote the argument of a functional with in brackets, but the argument of a simple function in parentheses. Functionals and functions are closely related. For instance, introducing a basis set for the wave function reveals that the functional derivative is really just a generalization of the gradient (compare Eqs. (VI.3) and (VI.10)), and turns a functional into a function. For instance, we can write the energy as a function of the coefficients of a basis function expansion: using (VI.6),

$$E(\mathbf{k}) = E(\{k_n\}_{n=0}^{\infty}) \equiv \frac{\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} k_l^* k_m \langle \Phi_l | \hat{H} | \Phi_m \rangle}{\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} k_l^* k_m \langle \Phi_l | \Phi_m \rangle}. \quad (\text{VI.12})$$

The Schrödinger equation,

$$\hat{H}\Psi_k = E_k \Psi_k \quad (\text{VI.13})$$

can be derived from the requirement that $\frac{\delta E[\Psi]}{\delta\Psi(\tau)} = 0$. This is a more fundamental “derivation” of the Schrödinger equation than most (though one can do much by forcing a functional called the “action” to be stationary instead

focusing on the energy). This is important, since it reveals that the eigenvectors and eigenvalues of the Hamiltonian operator represent stationary, or stable, states of the system.

C. Show that the Schrödinger Equation is equivalent to principle of stationary energy, $\frac{\delta E[\Psi]}{\delta \Psi(\mathbf{r})} = 0$. (Read the following hint.)

(I know of two straightforward ways to perform the derivation. One is to use Eq. (VI.5), but this is complicated by the fact that the kinetic energy operator will operate on the delta function. However, using the three-dimension generalization of problem I.C allows one to perform the analysis. The most straightforward approach is to calculate $E[\Psi + \delta\Psi] - E[\Psi]$ and rearrange the result into a form like

$$E[\Psi + \delta\Psi] - E[\Psi] = \int A[\Psi] \delta\Psi(\mathbf{r}) d\mathbf{r} + \mathcal{O}(|\delta\Psi(\mathbf{r})|^2) \quad (\text{VI.14})$$

Again, the analysis is a bit complicated because of the kinetic energy operator. The key is to use Green's theorem (cf. Eq. III.8 in problem set 1) to write integrals like $\langle \Psi | \nabla^2 \delta\Psi \rangle$ in terms of integrals like $\langle \nabla^2 \Psi | \delta\Psi \rangle$. In expanding the denominator you might want to use the geometric series, namely

$$(1+x)^k \equiv 1 + k \cdot x + \frac{k(k-1)}{2!} x^2 + \dots + \binom{k}{n} x^n + \dots \quad (\text{VI.15})$$

What is a stationary state? The following analysis might be helpful. Consider what happens when we change the wave function of a stationary state, Ψ_k , by a small amount, $\varepsilon\phi(\mathbf{r})$. We can assume that $\phi(\mathbf{r})$ and $\Psi_k(\mathbf{r})$ are orthogonal and normalized. Then (you will need to use the geometric series again),

D. Show that

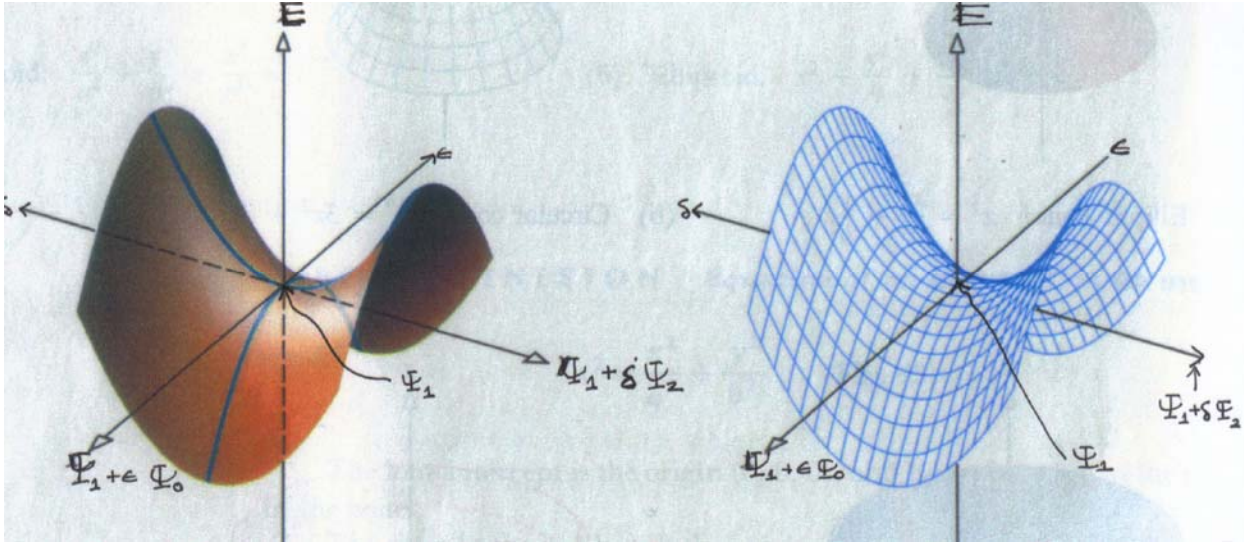
$$E[\Psi_k + \varepsilon\phi] \equiv E_k + \varepsilon^2 (E_\phi - E_k) - \varepsilon^4 (E_\phi - E_k)^2 + \dots \quad (\text{VI.16})$$

where E_k is the energy of the stationary state, Ψ_k , and

$$E_\phi \equiv \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle} = \langle \phi | \hat{H} | \phi \rangle. \quad (\text{VI.17})$$

The interpretation is that stationary states of systems are “saddle points”; the first order change in energy is zero, but the second order change is negative if $\phi(\mathbf{r})$ has a lower energy than $\Psi_k(\mathbf{r})$ but positive if $\phi(\mathbf{r})$ has a larger energy than $\Psi_k(\mathbf{r})$. The ground state is the absolute minimum energy state (and is thus a simple minimum, and not a maximum), but the k^{th} excited state is a true saddle point—one can lower the energy by moving in a k directions (take $\phi(\mathbf{r}) = \sum_{n=0}^{k-1} c_n \Psi_n(\mathbf{r})$) but other directions (take $\phi(\mathbf{r}) = \sum_{n=k+1}^{\infty} c_n \Psi_n(\mathbf{r})$) raise the energy.

The following figure depicts the situation for the first excited state of a system.



The stability of a stationary state is related to “how broad” the saddle point is, since this determines “how big” a perturbation is needed to a system away from its “stationary state” and force a transition to a lower-energy state. Thus, from Eq. (VI.16), one might expect that the rate at which a molecule spontaneously decays from an excited state, E_k , to a lower-energy stationary state, $E_{l < k}$, would be proportional to $(E_k - E_l)$. This gives the right trend (higher excited states are inherently less stable and transitions from excited states to lower-energy states are favored. But the assumption of a linear dependence is wrong (for rather subtle reasons); the lifetime of excited states is proportional to $(E_k - E_l)^3$.

VII. Perturbation Theory for Degenerate States (1.1)

When a set of degenerate orbitals (or wave functions) is placed in an electromagnetic field, the degeneracy is usually partially removed. For this reason, the perturbation theory of degenerate states is a key tool, not only for describing the splitting of spectral lines, but also for describing the removal of the degeneracy of orbitals with different values of the orbital angular momentum quantum number, l , but the same principle quantum number, n . We will derive the first order perturbation theory results for degenerate states in this problem.

We are given an unperturbed Hamiltonian, \hat{H} , and we label the states using the Schrödinger equation, so

$$\hat{H}\Psi_n \equiv E_n\Psi_n. \quad (\text{VII.1})$$

We assume that $\{\Psi_n\}_{n=0}^{\infty}$ are orthonormal and form a complete set of functions. In addition, we assume that the functions,

$$(\text{VII.2})$$

form a g -fold degenerate state, so that

$$\begin{aligned} \langle \Psi_n | \hat{H} | \Psi_n \rangle &= E_{k+1} & k+1 \leq n \leq k+g \\ \langle \Psi_n | \hat{H} | \Psi_n \rangle &\neq E_{k+1} & \text{otherwise.} \end{aligned} \quad (\text{VII.3})$$

Note that for any choice of the coefficients, $\{c_i\}_{i=1}^g$,

$$\hat{H} \sum_{i=1}^g c_i \Psi_i = E_{k+1} \sum_{i=1}^g c_i \Psi_i \quad (\text{VII.4})$$

That is, any linear combination of a set of degenerate eigenvectors is also an eigenvector of the Hamiltonian.

Now, let's consider the “perturbed” Hamiltonian,

$$\hat{H}_{\text{perturbed}}(\lambda) \equiv \hat{H} + \lambda \hat{V}. \quad (\text{VII.5})$$

Unless the perturbation has a special symmetry, for $\lambda \neq 0$ the degeneracy is broken. That is,

$$\langle \Psi_{k+1}^\lambda | \hat{H}_{\text{perturbed}}(\lambda) | \Psi_{k+1}^\lambda \rangle \neq \langle \Psi_{k+2}^\lambda | \hat{H}_{\text{perturbed}}(\lambda) | \Psi_{k+2}^\lambda \rangle \neq \dots \neq \langle \Psi_{k+g}^\lambda | \hat{H}_{\text{perturbed}}(\lambda) | \Psi_{k+g}^\lambda \rangle \quad (\text{VII.6})$$

where

$$(\hat{H} + \lambda \hat{V}) \Psi^\lambda = E^\lambda \Psi^\lambda \quad (\text{VII.7})$$

is the Schrödinger equation describing the perturbed system. This means that, even when $0 < |\lambda| \ll 1$ —that is, the perturbation is very, very, small—the wave functions of the perturbed the system might not resemble those of the unperturbed system. In particular, there is almost always some linear combination of the eigenvectors of the previous system (cf. Eq. (VII.4)) that is not an eigenvector of the perturbed Hamiltonian.

The way around this issue is straightforward, but a bit complicated. You can find standard treatments in most good quantum texts, and what follows is a bit nonstandard, but hopefully easier to follow. The usual procedure is to write the energy and wave function as a Taylor series in λ ,

$$E_n^\lambda = E_n^{\lambda=0} + \lambda \left. \frac{\partial E_n^\lambda}{\partial \lambda} \right|_{\lambda=0} + \frac{1}{2} \lambda^2 \left. \frac{\partial^2 E_n^\lambda}{\partial \lambda^2} \right|_{\lambda=0} + \dots \quad (\text{VII.8})$$

$$\Psi_n^\lambda = \Psi_n^{\lambda=0} + \lambda \left. \frac{\partial \Psi_n^\lambda}{\partial \lambda} \right|_{\lambda=0} + \frac{1}{2} \lambda^2 \left. \frac{\partial^2 \Psi_n^\lambda}{\partial \lambda^2} \right|_{\lambda=0} + \dots \quad (\text{VII.9})$$

The problem is that we don't know which of the many possible unperturbed wave functions (cf. Eq. (VII.4)) to use for the zeroth order approximation to $\{\Psi_n^\lambda\}_{n=k+1}^{k+g}$.

To get around this issue, let's turn the problem on its head. Consider a perturbation of the Hamiltonian which totally removes the degeneracy,

$$\hat{H}^{(0)} = \hat{H} + \mu \hat{V}. \quad (\text{VII.10})$$

(The perturbation is assumed to be very small, but still large enough to remove the degeneracy.) For this reason,

$$\hat{H}^{(0)} \Psi_n^{(0)} = E_n^{(0)} \Psi_n^{(0)} \quad (\text{VII.11})$$

and

$$E_0^{(0)} < E_1^{(0)} < \dots < E_{k+1}^{(0)} < E_{k+2}^{(0)} < \dots. \quad (\text{VII.12})$$

The “perturbed Hamiltonian”,

$$\begin{aligned} \hat{H}^{(\mu)} &\equiv \hat{H}^{(0)} - \mu \hat{V} \\ &= \hat{H}^{(0)} + \mu (-\hat{V}), \end{aligned} \quad (\text{VII.13})$$

is just the original, degenerate, Hamiltonian. The eigenfunctions and eigenvalues of $\hat{H}^{(\mu)}$ are known since they are just the eigenfunctions of the original system. For clarity, we adopt a naming convention similar to that in Eqs. (VII.11) and (VII.12), namely,

$$\hat{H}^{(\mu)} \Psi_n^{(\mu)} = E_n^{(\mu)} \Psi_n^{(\mu)} \quad (\text{VII.14})$$

where, because of the assumed g-fold degeneracy in Eq. (VII.2),

$$E_0^{(\mu)} < E_1^{(\mu)} < \dots < E_{k+1}^{(\mu)} = E_{k+2}^{(\mu)} = \dots = E_{k+g}^{(\mu)} < E_{k+g+1}^{(\mu)} \dots. \quad (\text{VII.15})$$

Equations (VII.11) and (VII.14) are our working equations. The choice of notation is such that a superscript (μ) denotes the Hamiltonian, eigenfunctions, and eigenvectors of the original degenerate system, while a superscript (0) denotes the corresponding quantities for the nondegenerate system described with $\hat{H}^{(0)}$.

Because the eigenfunctions of $\hat{H}^{(\mu)}$ form a complete set, we can write $\Psi_n^{(0)}$ as

$$\Psi_n^{(0)} = \sum_{i=0}^{\infty} c_{in}^{(0)} \Psi_i^{(\mu)}. \quad (\text{VII.16})$$

As is usual in perturbation theory, we expand quantities in Taylor series. That is,

$$\hat{H}^{(\mu)} = \hat{H}^{(0)} + \mu \left. \frac{\partial \hat{H}^{(\lambda)}}{\partial \lambda} \right|_{\lambda=\mu} \quad (\text{VII.17})$$

$$E_n^{(\mu)} = E_n^{(0)} + \mu \left. \frac{\partial E_n^{(\lambda)}}{\partial \lambda} \right|_{\lambda=\mu} + \frac{\mu^2}{2} \left. \frac{\partial^2 E_n^{(\lambda)}}{\partial \lambda^2} \right|_{\lambda=\mu} + \dots \quad (\text{VII.18})$$

$$c_{in}^{(\mu)} = c_{in}^{(0)} + \mu \left. \frac{\partial c_{in}^{(\lambda)}}{\partial \lambda} \right|_{\lambda=\mu} + \frac{\mu^2}{2} \left. \frac{\partial^2 c_{in}^{(\lambda)}}{\partial \lambda^2} \right|_{\lambda=\mu} + \dots \quad (\text{VII.19})$$

We now substitute these expressions into Eq. (VII.11), so

$$\begin{aligned} (\hat{H}^{(\mu)} + \mu \hat{V}) \sum_{i=0}^{\infty} c_{in}^{(0)} \Psi_i^{(\mu)} &= \left(E_n^{(\mu)} - \mu \left. \frac{\partial E_n^{(\lambda)}}{\partial \lambda} \right|_{\lambda=\mu} - \frac{\mu^2}{2} \left. \frac{\partial^2 E_n^{(\lambda)}}{\partial \lambda^2} \right|_{\lambda=\mu} - \dots \right) \sum_{i=0}^{\infty} c_{in}^{(0)} \Psi_i^{(\mu)} \\ (\hat{H}^{(\mu)} + \mu \hat{V}) \sum_{i=0}^{\infty} \left(c_{in}^{(\mu)} - \mu \left. \frac{\partial c_{in}^{(\lambda)}}{\partial \lambda} \right|_{\lambda=\mu} - \frac{\mu^2}{2} \left. \frac{\partial^2 c_{in}^{(\lambda)}}{\partial \lambda^2} \right|_{\lambda=\mu} - \dots \right) \Psi_i^{(\mu)} & \\ = \left(E_n^{(\mu)} - \mu \left. \frac{\partial E_n^{(\lambda)}}{\partial \lambda} \right|_{\lambda=\mu} - \frac{\mu^2}{2} \left. \frac{\partial^2 E_n^{(\lambda)}}{\partial \lambda^2} \right|_{\lambda=\mu} - \dots \right) \sum_{i=0}^{\infty} \left(c_{in}^{(\mu)} - \mu \left. \frac{\partial c_{in}^{(\lambda)}}{\partial \lambda} \right|_{\lambda=\mu} - \frac{\mu^2}{2} \left. \frac{\partial^2 c_{in}^{(\lambda)}}{\partial \lambda^2} \right|_{\lambda=\mu} - \dots \right) \Psi_i^{(\mu)} & \end{aligned} \quad (\text{VII.20})$$

The standard perturbation theory idea is that the coefficients of $\mu^0 = 1$, μ , μ^2 , etc. must all be equal. This is because in order for

$$f(x) = \sum_{n=0}^{\infty} a_n \mu^n \quad (\text{VII.21})$$

to be zero everywhere, all the a_n must equal zero. (Otherwise $|f(x)| > 0$ most places.) Thus, in order for two functions to be equal everywhere,

$$f(\mu) \equiv \sum_{n=0}^{\infty} l_n \mu^n = \sum_{n=0}^{\infty} r_n \mu^n \equiv g(\mu) \quad (\text{VII.22})$$

we must have that $l_\mu = r_\mu$. (Otherwise $f(\mu) - g(\mu) = \sum_{n=0}^{\infty} (l_n - r_n) \mu^n$ would not equal zero.) For this reason,

we must have that the coefficients of $\mu^0 = 1$, μ , μ^2 , etc. on the left-hand-side and right-hand-side of Eq. (VII.20) must be equal.

For $\mu^0 = 1$, this gives the equation

$$\hat{H}^{(\mu)} \sum_{i=0}^{\infty} c_{in}^{(\mu)} \Psi_i^{(\mu)} = E_n^{(\mu)} \sum_{i=0}^{\infty} c_{in}^{(\mu)} \Psi_i^{(\mu)} \quad (\text{VII.23})$$

Let us first consider the case where $E_n^{(\mu)}$ is a non-degenerate state. Then we must have that (to within an arbitrary complex phase factor), $c_{in}^{(\mu)} = \delta_{in}$. That is, $c_{in}^{(\mu)} = 0$ unless $i = n$ when $E_n^{(\mu)}$ is a non-degenerate state.

When $E_n^{(\mu)}$ is a degenerate state, any linear combination of the degenerate eigenfunctions will also be an eigenfunction of $\hat{H}^{(\mu)}$, cf. Eq. (VII.4). So if the states $\Psi_{k+1}^{(\mu)}, \Psi_{k+2}^{(\mu)}, \dots, \Psi_{k+g}^{(\mu)}$ are degenerate, then

$$\hat{H}^{(\mu)} \sum_{i=k+1}^{k+g} c_{in}^{(\mu)} \Psi_i^{(\mu)} = E_{k+1}^{(\mu)} \sum_{i=k+1}^{k+g} c_{in}^{(\mu)} \Psi_i^{(\mu)}. \quad (\text{VII.24})$$

That is, $c_{in}^{(\mu)} = 0$ unless $\Psi_i^{(\mu)}$ and $\Psi_n^{(\mu)}$ have the same energy. Recalling that the Kronecker delta, δ_{ij} , is zero unless its arguments are equal ($i = j$), we can summarize the consequences of Eq. (VII.23) with

$$\hat{H}^{(\mu)} \sum_{i=0}^{\infty} \delta_{E_i E_n} c_{in}^{(\mu)} \Psi_i^{(\mu)} = E_n^{(\mu)} \sum_{i=0}^{\infty} \delta_{E_i E_n} c_{in}^{(\mu)} \Psi_i^{(\mu)}. \quad (\text{VII.25})$$

This indicates that we can replace $c_{in}^{(\mu)}$ with $c_{in}^{(\mu)} \delta_{E_i E_n}$.

- A.** Show that Eq. (VII.25) actually agrees with the conclusions in the preceding two paragraphs for both non-degenerate and degenerate states.

The “meat” of our treatment is based on the equating the coefficients of $\mu^1 = \mu$ in Eq. (VII.20).

- B.** When you equate the coefficients of μ in Eq. (VII.20), show that you obtain the equation

$$\sum_{i=0}^{\infty} c_{in}^{(\mu)} \hat{V} \Psi_i^{(\mu)} = - \sum_{i=0}^{\infty} \left(\frac{\partial E_n^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} c_{in}^{(\mu)} + (E_n^{(\mu)} - E_i^{(\mu)}) \frac{\partial c_{in}^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} \right) \Psi_i^{(\mu)}. \quad (\text{VII.26})$$

The easiest way to solve Eq. (VII.26) is to multiply both sides by $\Psi_m^{(\mu)}$ and integrate. We obtain

$$\begin{aligned} \sum_{i=0}^{\infty} c_{in}^{(\mu)} \langle \Psi_m^{(\mu)} | \hat{V} | \Psi_i^{(\mu)} \rangle &= - \sum_{i=0}^{\infty} \left(\frac{\partial E_n^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} c_{in}^{(\mu)} + (E_n^{(\mu)} - E_i^{(\mu)}) \frac{\partial c_{in}^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} \right) \langle \Psi_m^{(\mu)} | \Psi_i^{(\mu)} \rangle. \\ &= - \sum_{i=0}^{\infty} \left(\frac{\partial E_n^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} c_{in}^{(\mu)} + (E_n^{(\mu)} - E_i^{(\mu)}) \frac{\partial c_{in}^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} \right) \delta_{mi}. \\ &= - \left(\frac{\partial E_n^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} c_{mn}^{(\mu)} + (E_n^{(\mu)} - E_m^{(\mu)}) \frac{\partial c_{mn}^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} \right) \end{aligned} \quad (\text{VII.27})$$

There are two cases in Eq. (VII.27). In the first case, $E_n^{(\mu)} = E_m^{(\mu)}$. The second term on the right-hand-side then vanishes, and we get

$$\sum_{i=0}^{\infty} \langle \Psi_m^{(\mu)} | \hat{V} | \Psi_i^{(\mu)} \rangle c_{in}^{(\mu)} = - \frac{\partial E_n^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} c_{mn}^{(\mu)} \quad (\text{if } E_m^{(\mu)} = E_n^{(\mu)}) \quad (\text{VII.28})$$

Because $c_{in}^{(\mu)} = 0$ unless $E_i^{(\mu)} = E_n^{(\mu)}$, this becomes

$$\sum_{i=k+1}^{k+g} \langle \Psi_m^{(\mu)} | \hat{V} | \Psi_i^{(\mu)} \rangle c_{in}^{(\mu)} = - \frac{\partial E_n^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} c_{mn}^{(\mu)} \quad (\text{if } E_m^{(\mu)} = E_n^{(\mu)}) \quad (\text{VII.29})$$

This equation is solved by finding the eigenvalues and eigenvectors of the matrix whose elements are $\langle \Psi_m^{(\mu)} | \hat{V} | \Psi_i^{(\mu)} \rangle$, with $k+1 \leq i, m \leq k+g$.

- C.** Show that if $\mathbf{c}_n \equiv [c_{k+1,n}^{(\mu)}, c_{k+2,n}^{(\mu)}, \dots, c_{k+g,n}^{(\mu)}]$ is an eigenvector of

$$\mathbf{V} \equiv V_{mi} = \langle \Psi_m^{(\mu)} | \hat{V} | \Psi_i^{(\mu)} \rangle; \quad k+1 \leq i, m \leq k+g \quad (\text{VII.30})$$

then Eq. (VII.29) is satisfied.

The second case in Eq. (VII.27) arises when $E_n^{(\mu)} \neq E_m^{(\mu)}$. In this case, $c_{mn}^{(\mu)} = 0$, and so Eq. (VII.27) simplifies to

$$\frac{\sum_{i=0}^{\infty} \langle \Psi_m^{(\mu)} | \hat{V} | \Psi_i^{(\mu)} \rangle c_{in}^{(\mu)}}{E_m^{(\mu)} - E_n^{(\mu)}} = \frac{\partial c_{mn}^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} \quad (\text{VII.31})$$

where $c_{in}^{(\mu)} = 1$ if $E_i \neq E_n$, but is otherwise determined by Eq. (VII.29).

We can now substitute our results back into our original expressions. Using Eqs. (VII.18) and (VII.19),

$$E_n^{(0)} \approx E_n^{(\mu)} - \mu \frac{\partial E_n^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} \quad (\text{VII.32})$$

and

$$\Psi_n^{(0)} = \sum_{i=0}^{\infty} \left(c_{in}^{(\mu)} - \mu \frac{\partial c_{in}^{(\lambda)}}{\partial \lambda} \Big|_{\lambda=\mu} \right) \Psi_i^{(\mu)} \quad (\text{VII.33})$$

where the coefficients are determined using Eqs. (VII.25), (VII.29), and (VII.31).

- D. Suppose $\hat{H}^{(\mu)}$ has no degenerate states. Show that Eqs. (VII.32) and (VII.33) reduce to the usual formulae from first-order perturbation theory,**

$$E_n^{(0)} = E_n^{(\mu)} + \langle \Psi_n^{(\mu)} | \mu \hat{V} | \Psi_n^{(\mu)} \rangle \quad (\text{VII.34})$$

$$\Psi_n^{(0)} = \Psi_n^{(\mu)} + \sum_{\substack{n=0 \\ n \neq m}}^{\infty} \frac{\langle \Psi_m^{(\mu)} | \mu \hat{V} | \Psi_n^{(\mu)} \rangle}{E_n^{(\mu)} - E_m^{(\mu)}} \Psi_m^{(\mu)}. \quad (\text{VII.35})$$

It is important to understand these results; they will be very important in this course!!

VIII. Electrons in Ammonia (2.0)

When a small amount of an alkali metal is dissolved in liquid ammonia, a blue solution results. Then, as the solution becomes less dilute, a bronze-colored phase separates out (except for Cesium, where the blue and bronze phase have no perceptible difference in density), floating above the “blue phase” until the entire solution is bronze. Precipitation of the “bronze” solution gives alkalide salts. In the rare cases where crystals have been obtained from the blue solution, there is no anion present in the crystal structure. In the following we will explore this system, with special emphasis on sodium metal.

The blue solution is believed to result from free electrons dissolved in the ammonia solution. That is, adding sodium metal to ammonia causes the sodium to ionize (the sodium cation is stabilized by ammonia molecules with the lone pair on nitrogen “pointing” at the cation), forming cations and electrons:



The electrons occupy cavities in the ammonia (the solution is of much lower density than usual alkali salt solutions), which are believed to be about 3-3.4 Å in diameter. The blue color is because a very broad absorption maximum (centered at 1500 nm) spreads into the red portion of the spectrum, so that blue light is absorbed less weakly (reflected more strongly) than red light.

To model this state, we can use a potential of the form

$$V(r) \equiv A(3 - r^2)e^{-\frac{r^2}{3}} \quad (\text{VIII.2})$$

where $A > 0$ approximates the depth of the energy well. (It should be noted that this is a model potential (designed so you can do the integrals easily!)—don’t take the results too literally! The model does reproduce some qualitative features: the well is somewhat “flat” on the bottom and the potential becomes positive as you near the “boundary molecules” confining the electron.)

A. Write the Hamiltonian for an electron confined by this 3-dimensional potential.

(Hint: Use spherical coordinates and separation of variables.)

We can use the variational principle to estimate the ground state wave function and ground state energy of an electron confined by this potential (cf. problem IID). We choose, as an approximate wave function for the ground state,

$$\psi_0(r, \alpha) = N(\alpha) e^{-\alpha r^2}. \quad (\text{VIII.3})$$

This wave function must be normalized; that is,

$$1 = \int_0^\infty \left(N(\alpha) e^{-\alpha r^2} \right)^2 4\pi r^2 dr. \quad (\text{VIII.4})$$

B. Using Eq. (VIII.4), find $N(\alpha)$.

For relevant integrals, you should “look ahead” to the last few pieces of this problem.

According to the variational principle (problem IID), we should choose the parameter, α , so that the energy is minimized. To do this, calculate the expectation value of the Hamiltonian from part A,

$$E_0(\alpha) \equiv \int_0^\infty \psi_0(r, \alpha) \hat{H} \psi_0(r, \alpha) dr. \quad (\text{VIII.5})$$

C. Find the value of α for which Eq. (VIII.5) takes its minimum value.

D. Using the result from part C, what is the approximate ground state energy?

E. What is the approximate ground state wave function?

The observed absorbance can be attributed to an analogue of the $1s \rightarrow 2p$ transition in Hydrogen, since this is the lowest-lying non-forbidden transition for this system.

F. Show that the wave function

$$\psi_1(r, \beta) = M(\beta)(r \cos \theta) e^{-\beta r^2} \quad (\text{VIII.6})$$

is orthogonal to $\psi_0(r, \beta)$.

G. Find the normalization constant, $M(\beta)$.

Because $\psi_1(r, \alpha)$ and $\psi_0(r, \alpha)$ are orthogonal, we can approximate the first excitation energy (cf. problem IIE) by minimizing the energy with respect to β . That is, calculate the expectation value of the Hamiltonian

$$E_1(\beta) \equiv \int_0^\infty \psi_1(r, \beta) \hat{H} \psi_1(r, \beta) dr. \quad (\text{VIII.7})$$

H. Find the value of β for which Eq. (VIII.7) takes its minimum value.

I. Using the result from part G, what is the approximate excited state energy?

J. What is the approximate excited state wave function? Is β smaller or larger than α ? What does this say about “how delocalized” the electron is in its ground and first excited states?

The energy of the transition can be considered to be $h\nu$, where h is Planck’s constant and $\nu = \frac{c}{\lambda}$, where c is the speed of light and, in this case, $\lambda \approx 1500 \text{ nm}$.

K. What is the experimentally observed energy of the transition, E_{\uparrow}^1 ?

L. What value of A in Eq. (VIII.2) causes the theoretical excitation energy, $E_1 - E_0$, to equal the experimental one, E_{\uparrow}^1 .

The “bronze” phase is generally attributed to the “disproportionation reaction”,



You’ll recall that the first day of class, I said that “in general”, first ionization potential of an atom or molecule was less than its second ionization potential, but here this is not true: the disproportionation reaction in Eq. (VIII.8) can only occur if the first ionization potential of the sodium anion (the electron affinity of the sodium atom) is less than the second ionization potential of the anion (the ionization potential of sodium). This is not true in the gas phase, but

occurs in solution because a polar solvent (ammonia) is much better at stabilizing ions (especially the cation) than neutral atoms/molecules.

The bronze color of this solution may be taken as an indication that yellow light is no longer strongly absorbed, and so we might expect that the absorption maximum in the bronze solution has been shifted to higher wavelength (lower energy). There are two possible models. One is that this is a “charge transfer” band, so that light ionizes the anion and shifts an electron to the cation. Such phenomena are beyond the scope of this course.

We can consider, however, the possibility that we are observing an excited state of the anion. Since the first 10 electrons in Sodium occupy a closed shell, we model the anion as 2 electrons moving in some effective nuclear charge, $Z_{eff} \ll 11$. The $1s \rightarrow 2p$ transition in Helium has an energy of 584.3 \AA . Because atomic energy level spacings change in approximate proportion to Z^2 , we may deduce that the absorbed wavelength for two electrons moving in the effective nuclear charge, Z_{eff} , will be approximately given by

$$\lambda = 58.43 \text{ nm} \cdot \left(\frac{Z_{He}}{Z_{eff}} \right)^2 = \frac{223.7 \text{ nm}}{Z_{eff}^2}. \quad (\text{VIII.9})$$

Suppose the new absorption maximum is at 1700 nm .

M. What is the value of Z_{eff} ? Is this reasonable?

As should be evident from the preceding, we often need to evaluate integrals of Gaussian functions. The following techniques are very useful in this regard.

To evaluate

$$I(\alpha) = \int_{-\infty}^{\infty} e^{-\alpha x^2} dx \quad (\text{VIII.10})$$

we first square the integral. So

$$\begin{aligned} (I(\alpha))^2 &= \int_{-\infty}^{\infty} e^{-\alpha x^2} dx \int_{-\infty}^{\infty} e^{-\alpha y^2} dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\alpha(x^2+y^2)} dx dy \end{aligned} \quad (\text{VIII.11})$$

We then evaluate the integral using spherical coordinates, namely,

$$(I(\alpha))^2 = \int_0^{2\pi} \int_0^{\infty} e^{-\alpha r^2} r dr d\theta \quad (\text{VIII.12})$$

N. Show that $I(\alpha) = \sqrt{\frac{\pi}{\alpha}}$

Now, let $J(\alpha) = \frac{1}{2} I(\alpha) = \int_0^{\infty} e^{-\alpha r^2} dr$. Note that

$$\frac{dJ}{d\alpha} = \int_0^{\infty} -r^2 e^{-\alpha r^2} dr \quad (\text{VIII.13})$$

$$\frac{d^2 J}{d\alpha^2} = \int_0^{\infty} r^4 e^{-\alpha r^2} dr \quad (\text{VIII.14})$$

etc.. The result is the formula,

$$\int_0^{\infty} r^{2n} e^{-\alpha r^2} dr \equiv \frac{(2n-1)(2n-3)\cdots(1)}{2(2\alpha)^n} \sqrt{\frac{\pi}{\alpha}}. \quad (\text{VIII.15})$$

The first term in the numerator is sometimes called a “double” or “two-step” factorial, and is often denoted $(2n - 1)!!$.

O. Using Eq. (VIII.13), confirm the formula (VIII.15) for the $n = 1$ case.

The result for odd powers of r is

$$\int_0^{\infty} r^{2n+1} e^{-\alpha r^2} dr = \frac{n!}{2\alpha^{n+1}}. \quad (\text{VIII.16})$$

The validity of this formula is even easier to establish, since it can be done using the fundamental theorem of calculus. For $n = 0$, one has

$$\begin{aligned} K(\alpha) &= \int_0^{\infty} r e^{-\alpha r^2} dr \\ &= \frac{-1}{2\alpha} \int_0^{\infty} (-2\alpha r) e^{-\alpha r^2} dr \\ &= \frac{-1}{2\alpha} \int_0^{\infty} \frac{d}{dr} (e^{-\alpha r^2}) dr \xrightarrow{\text{fund. th. of calc.}} \frac{-1}{2\alpha} \left[\lim_{r \rightarrow \infty} e^{-\alpha r^2} - e^{-\alpha(0)^2} \right] \\ &= \frac{1}{2\alpha}. \end{aligned} \quad (\text{VIII.17})$$

Now, differentiating $K(\alpha)$ with respect to α gives higher order terms in the series.

P. By evaluating $\frac{dK(\alpha)}{d\alpha}$, confirm Eq. (VIII.16) for $n = 1$.

On the one hand, I have no problem if you jot these integrals down somewhere so you never have to do them again. (I think you'll agree, though, that once you know how to do them, they aren't that hard.) On the other hand, I want you to know how to do them. That way, if you ever get stranded on a deserted island and decide to while away the hours scribbling solutions to theoretical chemistry problems in the sand, you won't get stuck.