Lecture #9

Plan:
1. Laguerre polynomials and the radial function
2. Concepts in probability
3. The Heisenberg uncertainty principle and the commutator
4. Variational method
5. Rayleigh-Ritz trial functions
6. Derivation of the secular equation
7. H2+ ion – full problem and associated challenges
8. Born-Oppenheimer approximation the electronic Hamiltonian
9.

Questions you should be able to answer by the end of today’s lecture:
1. Understand the idea behind the variational method for approximating eigenfunctions.
2. Why does bonding occur?
3. How to calculate simple approximations to the molecular orbital in an ionized Hydrogen molecule.
4. How do the molecular orbitals change as the interatomic distance is decreased
5. What physical insight leads to the separation of the electronic and nuclear coordinates?
6. How to choose trial wave functions and use the variation principle.

Math Tools:

1. Basic probability concepts
2. Variation methods
3. Minimization of functionals

References:

1. Merzbacher chapter 8.
2. Lowe, chapter 7.
Hydrogen atom eigenfunctions

\[ u_{nlm} = R_{nl}(r)Y_{lm}(\theta, \varphi) \]

\[ E_n = -\frac{E_s}{n^2} \]

\[ R(r) = e^{-\frac{r}{na_0}} \sqrt{\frac{2}{n}} \frac{2}{2n(n+l)!} \left( \frac{r}{na_0} \right) L_{n-l-1}^{2l+1} \left( \frac{2r}{na_0} \right) \]

Where \( L_{n-l-1}^{2l+1} \) is the generalized Laguerre polynomial (using Mathematica’s definition)

Plotting the \( r^2 R_{10}^2 (r), r^2 R_{20}^2 (r), r^2 R_{30}^2 (r) \) distributions:

\[ R[n, n, l] := e^{-\frac{r}{na_0}} (n+1)! \sqrt{\frac{(n+1)}{n} \frac{1}{2n(n+l)!}} \left( \frac{2r}{na_0} \right)^{n-l-1} \]

\[ \text{LaguerreL}[n-1-1, 2l+1, \frac{2r}{na_0}] \]

\[ \text{Plot} \left[ \{ r^2 R[r, 1, 0]^2, r^2 R[r, 2, 0]^2, r^2 R[r, 3, 0]^2 \}, \{ r, 0, 40 \}, \text{PlotRange} \to \{ \{0, 40\}, \{0, .6\} \} \right] \]
The Heisenberg Uncertainty Principle and it’s connection to the commutator
We mentioned in class that only operators that commute can be measured “simultaneously” and are thus termed compatible this follows from the fact that they have common eigenfunctions. Implying that for a set of non-commuting observables measurement of one physical quantity and the subsequent collapse of the wavefunction to the corresponding specific eigenfunction will invariably associate an uncertainty with respect the measurement of the other.

Review of some basic concepts in probability
The set of possible results of a test, experiment or trial are called outcomes \( \{O_i\} \) we assume that each outcome is assigned a probability of occurrence \( p_i \). The set of probabilities \( \{p_i\} \) is called a probability distribution. A variable \( X \) which takes on the value \( X_i \) if outcome \( O_i \) occurs is called a random variable. A function of a random variable is itself a random variable.

The expectation value (or mean value) of the random variable \( X \) for the given probability distribution is,
\[
\langle X \rangle = \sum_i p_i X_i
\]
The variance of a random variable is given by:
\[
(\Delta X)^2 = \left( \langle X - \langle X \rangle \rangle^2 \right) = \sum_i \left(X - \langle X \rangle \right) p_i = \langle X^2 \rangle - \langle X \rangle^2
\]
The standard deviation is defined by:
\[
\Delta X = \sqrt{\langle X^2 \rangle - \langle X \rangle^2}
\]
Discuss: standard deviation as a measure of uncertainty in a measurement (plot the value of a measurement vs. the experiment number).

The covariance of two random variables is,
\[
\langle \Delta X \Delta Y \rangle = \left( \langle X - \langle X \rangle \rangle \langle Y - \langle Y \rangle \rangle \right) = \langle XY \rangle - \langle X \rangle \langle Y \rangle
\]
if \( X \) and \( Y \) are uncorrelated (independent) the covariance vanishes.

**Mathematica note**
- \( \text{LaguerreL}[n, a, x] \) gives the generalized Laguerre polynomial \( L_n^a(x) \)
- \( \text{SphericalHarmonicY}[l, m, \theta, \phi] \) gives the spherical harmonic \( Y_{lm}(\theta, \phi) \)
Let us now use these basic probability concepts to derive a rigorous form of the Heisenberg uncertainty principle.

The expectation value of an operator gives us the mean value of the physical quantity for a particular state. One would approach this value if a measurement of $A$ would be performed on a large collection of identical states $\Psi$.

$$\langle \hat{A} \rangle = \frac{\langle \Psi \mid \hat{A} \Psi \rangle}{\langle \Psi \mid \Psi \rangle} = \frac{\int_{-\infty}^{\infty} \Psi^* \hat{A} \Psi \ d\psi}{\int_{-\infty}^{\infty} \Psi^* \Psi \ d\psi}$$

The definition of an Hermitian operator is:

$$\langle F\Psi \rangle^* = \int \Psi^* F \Psi \ d\psi$$

This implies that:

$$(\Delta \hat{A})^2 \equiv \left( \hat{A} - \langle \hat{A} \rangle \right)^2 = \langle \left( \hat{A} - \langle \hat{A} \rangle \right) \Psi \mid \left( \hat{A} - \langle \hat{A} \rangle \right) \Psi \rangle$$

and,

$$(\Delta \hat{B})^2 \equiv \left( \hat{B} - \langle \hat{B} \rangle \right)^2 = \langle \left( \hat{B} - \langle \hat{B} \rangle \right) \Psi \mid \left( \hat{B} - \langle \hat{B} \rangle \right) \Psi \rangle$$

define,

$$\Psi_a = \left( \hat{A} - \langle \hat{A} \rangle \right) \Psi$$

$$\Psi_b = \left( \hat{B} - \langle \hat{B} \rangle \right) \Psi$$

Using the Schwarz inequality

$$\langle \Psi_a \Psi_a \rangle \langle \Psi_b \Psi_b \rangle \geq |\langle \Psi_a \Psi_b \rangle|^2$$

which can be interpreted geometrically as the cosine of an angle has to be smaller or equal to 1.

and substituting the expressions defined above,

$$(\Delta \hat{A})^2 (\Delta \hat{B})^2 \geq \left| \langle \left( \hat{A} - \langle \hat{A} \rangle \right) \Psi \mid \left( \hat{B} - \langle \hat{B} \rangle \right) \Psi \rangle \right|^2 = \left| \langle \Psi \mid \left( \hat{A} - \langle \hat{A} \rangle \right) \left( \hat{B} - \langle \hat{B} \rangle \right) \Psi \rangle \right|^2$$

Lets us define,

$$[\hat{A}, \hat{B}] = i\hat{C}$$

where the $i$ has been added to ensure that $C$ is hermitian.

$$\left( \hat{A} - \langle \hat{A} \rangle \right) \left( \hat{B} - \langle \hat{B} \rangle \right) = \frac{\left( \hat{A} - \langle \hat{A} \rangle \right) \left( \hat{B} - \langle \hat{B} \rangle \right) + \left( \hat{B} - \langle \hat{B} \rangle \right) \left( \hat{A} - \langle \hat{A} \rangle \right)}{2} + i \frac{\hat{C}}{2} = \hat{F} + i \frac{\hat{C}}{2}$$
\[(\Delta \hat{A})(\Delta \hat{B}) \geq \left| \langle \psi , \left\{ \hat{F} + \frac{i}{2} \hat{C} \right\} \psi \rangle \right|^2 = \left( \langle \hat{F} \rangle + \frac{i}{2} \langle \hat{C} \rangle \right) \left( \langle \hat{F} \rangle - \frac{i}{2} \langle \hat{C} \rangle \right) = \langle \hat{F} \rangle^2 + \frac{1}{4} \langle \hat{C} \rangle^2 \geq \frac{1}{4} \langle \hat{C} \rangle^2 = \frac{1}{4} \left[ \left[ \hat{A}, \hat{B} \right] \right]^2 \rightarrow (\Delta \hat{A})(\Delta \hat{B}) \geq \frac{1}{2} \left[ \left[ \hat{A}, \hat{B} \right] \right] \]

**The variation method**

The problem: in most cases of physical interest the eigenfunctions of the Hamiltonian cannot be readily found. In such cases one needs to resort to a systematic procedure for searching for functions which approximate the eigenfunctions. The calculus of variations provides us with a powerful tool for solving such problems where \( H \) is known but not \( u_n \) and \( E_n \).

The variation method is based on the idea that if we try to find a trial function \( \phi \) that minimizes average \( E \) it will be greater or equal than the true ground state eigenfunctions \( u_n \).

\[ \overline{E} = \langle \hat{H} \rangle = \int_{-\infty}^{\infty} \phi^* \hat{H} \phi dv \geq E_{\text{ground state}} \]

**Proof:**

Because \( H \) is Hermitian its true eigenfunctions are a complete set therefore any function and in particular the trial function can be expanded in terms of eigenfunctions:

\[ \chi (x) = \sum_{i=1}^{n} c_i u_n (x) \]

\[ \langle \chi | \hat{H} | \chi \rangle = \sum_{i=1}^{n} |c_i|^2 E_n \]

and

\[ \langle \chi | \chi \rangle = \sum_{i=1}^{n} |c_i|^2 \]

For \( \overline{E} = \langle \hat{H} \rangle = E_{\text{ground}} \) it is necessary and sufficient for all \( c_n = 0 \) except for \( n=1 \).

**Note:** it is possible to find eigenvalues other than the ground energy using the variation method:

**The Rayleigh –Ritz trial function**

In this method a trial function is chosen as a linear combination of independent and - in most cases orthogonal and normalized - basis functions \( \chi \). Thus the trial function is of the form:

\[ \chi = \sum_{n=1}^{N} c_n \phi_n \]

The trial function can be viewed as a vector in a basis of \( \phi \) where the \( c \)’s are the components of the vector.
The expectation value is expressed as,

$$
\langle H \rangle = \langle \chi | \hat{H} | \chi \rangle = \sum_{i,j=1}^{N} c_{i}^* \langle \phi_i | \hat{H} | \phi_j \rangle c_j
$$

Note: the expression $$\langle \phi_i | \hat{H} | \phi_j \rangle$$ is a complex number obtained by performing the integral $$\int \phi_i \hat{H} \phi_j \, dv$$ and is called the i,j matrix element of H or $$H_{ij}$$.

The condition for achieving a stationary (or minimum) value of H with respect to variations in the trial wavefunction is:

$$
\frac{\partial}{\partial c_i^*} \langle \hat{H} \rangle = 0 \quad \text{and} \quad \frac{\partial}{\partial c_j} \langle \hat{H} \rangle = 0
$$

$$
\frac{\partial}{\partial c_i^*} \langle \hat{H} \rangle = \frac{\sum_{j=1}^{N} \langle \phi_i | \hat{H} | \phi_j \rangle c_j \left( \sum_{i,j=1}^{N} c_i^* c_j \right) - c_i \sum_{i,j=1}^{N} c_j^* \langle \phi_i | \hat{H} | \phi_j \rangle c_j}{\left( \sum_{i,j=1}^{N} c_i^* c_j \right)^2} = 0
$$

$$
\rightarrow \sum_{j=1}^{N} \langle \phi_i | \hat{H} | \phi_j \rangle c_j \left( \sum_{i,j=1}^{N} c_i^* c_j \right) = c_i \sum_{i,j=1}^{N} c_j^* \langle \phi_i | \hat{H} | \phi_j \rangle c_j
$$

$$
\sum_{j=1}^{N} \langle \phi_i | \hat{H} | \phi_j \rangle c_j = c_i \sum_{i,j=1}^{N} c_j^* \langle \phi_i | \hat{H} | \phi_j \rangle = \langle \hat{H} \rangle c_i = \bar{E} c_i
$$

Yielding n equations with n unknowns since for each $$c_i$$ there is a separate equation, $$\bar{E}$$ being the expectation value corresponding to the optimal choice of the trial function. This equation is known as the secular equation and we are looking for its roots (i.e. the $$\bar{E}$$ values).

$$
\sum_{j=1}^{n} \langle \phi_i | \left( \hat{H} | \phi_j \rangle - \delta_{ij} \bar{E} \right) c_j = 0
$$

In matrix notation,

$$
\begin{bmatrix}
H_{11} - \bar{E} & H_{12} & \cdots & H_{1n} \\
H_{21} & H_{22} - \bar{E} & \cdots & H_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
H_{n1} & H_{n2} & \cdots & H_{nn} - \bar{E}
\end{bmatrix}
\begin{bmatrix}
c_1 \\
c_2 \\
\vdots \\
c_n
\end{bmatrix} = 0
$$

which can only have a non-trivial solution if
\[ \text{Det} \left( \sum_{j=1}^{n} \langle \phi_i | (\hat{H} | \phi_j) \rangle - \delta_{ij} E \right) = 0 \]

\[
\begin{vmatrix}
H_{11} - E & H_{12} & \ldots & H_{1n} \\
H_{21} & H_{22} - E & & \\
& \ddots & \ddots & \\
H_{n1} & & H_{nn} - E
\end{vmatrix} = 0
\]

**Note:** if the basis set is not orthogonal this matrix has some additional terms see Lowe page 195 for more details

*Thus the problem of finding the approximate eigenfunctions and corresponding energy spectrum has been transformed from one that involves functions and minimization to an algebraic problem.*

**The Ritz Theorem:** The expectation value of the Hamiltonian is stationary in the vicinity of its discrete eigenvalues.

All of the intelligence goes into choosing a “good” trial function, which already contains in it a large amount of the true eigenfunction. The trial function is chosen with the aid of physical insight. Hartree Fock method for calculating eigenfunctions of many electron Hamiltonians is also an example of the variation method.

**Understanding the chemical bond - Approximating the wave functions of molecules**

**LCAO linear combination of atomic orbitals**

The problem:

How do we calculate the eigenfunctions of molecules?

The system – simple diatomic molecule

I] Draw the potential

II] Write the full Hamiltonian
\[
\hat{H}(R_A, R_B, r, P_A, P_B, p) = -\frac{\hbar^2}{2m_e} \nabla_r^2 - \frac{\hbar^2}{2m_p} \nabla_A^2 - \frac{\hbar^2}{2m_p} \nabla_B^2 + \left( -\frac{e^2}{r_1} - \frac{e^2}{r_2} + \frac{e^2}{R} \right)
\]

where,
\[
\vec{r}_1 = \vec{r} - \vec{R}_A \\
\vec{r}_2 = \vec{r} - \vec{R}_B
\]

In general the wave functions should contain all of the coordinates of the system.
Because the protons are 1836 times heavier than the electrons they have much smaller velocity this means that at the time scales associated with the electronic movement the protons are stationary.

**The Born-Oppenheimer Approximation**

1. The electronic wavefunctions depend only on the positions of the nuclei and not on their momentum
2. The positions of the protons are a parameter not a QM variable
3. Protons have a fixed separation \( R \) just add a constant to the energy.
4. Reduce the three body problem to an electron in a potential
5. The Hamiltonian is separable

Using the procedure we developed:
Write the Hamiltonian for the electron energy and nuclear repulsion energy.
\[
\hat{H}_{el-nr} = -\frac{\hbar^2}{2m_e} \nabla_r^2 - \frac{e^2}{r_1} + \frac{e^2}{r_2} + \frac{e^2}{R}
\]

\[
\hat{H}_{el} u_{el}(r) = E_{el} u_{el}(r)
\]
\[
E_{el} + E_{nuc} = E_{el} + 1/R
\]

What we have is a functional dependence of the electronic energy on the nuclear separation \( E_{el}(R) \).
\[\hat{H}(R_A, R_B) = -\frac{\hbar^2}{2m_p} \hat{\nabla}_A^2 + -\frac{\hbar^2}{2m_p} \hat{\nabla}_B^2 + \left( E_{\text{elec}}(R) + \frac{e^2}{R} \text{ nuclear repulsion energy} \right)\]

\[\hat{H}_{\text{nuc}}(R_A, R_B) u_{\text{nuc}}(R_A, R_B) = E_{\text{nuc}} u_{\text{nuc}}(R_A, R_B)\]

These eigenfunctions represent the translation, rotation and vibration states of the crystal.

\[u_{\text{tra}}(R_A, R_B, r) = u_{\text{elec}}(r) u_{\text{nuc}}(R_A, R_B)\]

We would like to find the exact energy eigenfunctions because this is difficult, we will find approximate solutions

\[\hat{H}_{\text{el}}u = Eu\]

Assume that our solution can be written as a linear combination of atomic orbitals

\[\chi(\vec{r}) = c_1 \underbrace{u_{100}(\vec{r}_1)}_{\phi_1} + c_2 \underbrace{u_{100}(\vec{r}_2)}_{\phi_2}\]

where

\[\vec{r}_1 = \vec{r} - \vec{R}_A\]
\[\vec{r}_2 = \vec{r} - \vec{R}_B\]

Since this wave function is not an energy eigenfunction, we would like to minimize the average energy or the expectation value with respect to the coefficients,

\[E_{av} = \langle \hat{H} \rangle = \frac{\int \chi \hat{H} \chi d\nu}{\int \chi \chi d\nu}\]

Take a closer look at the integrals,

\[\bar{E}(c_1, c_2) = \frac{\int \bigg[ c_1 \underbrace{u_{100}(\vec{r}_1)}_{\phi_1} + c_2 \underbrace{u_{100}(\vec{r}_2)}_{\phi_2} \bigg] \hat{H} \bigg[ c_1 \underbrace{u_{100}(\vec{r}_1)}_{\phi_1} + c_2 \underbrace{u_{100}(\vec{r}_2)}_{\phi_2} \bigg] d\nu}{\int \bigg[ \underbrace{c_1 \underbrace{u_{100}(\vec{r}_1)}_{\phi_1} + c_2 \underbrace{u_{100}(\vec{r}_2)}_{\phi_2}}_{\phi} \bigg] \bigg[ c_1 \underbrace{u_{100}(\vec{r}_1)}_{\phi_1} + c_2 \underbrace{u_{100}(\vec{r}_2)}_{\phi_2} \bigg] d\nu}\]

We are now going to apply the variation principle to allow us to find the function which minimizes \(E\) (Cauchy Reiman Theorem).

\[\frac{\partial \bar{E}}{\partial c_i} = 0\]
Get a number of integrals which we will label in the following way:

\[
\int u_{100}(\vec{r}_1) \hat{H} u_{100}(\vec{r}_1) \, dr \equiv H_{11}
\]
\[
\int u_{100}(\vec{r}_2) \hat{H} u_{100}(\vec{r}_2) \, dr \equiv H_{22}
\]
\[
\int u_{100}(\vec{r}_1) \hat{H} u_{100}(\vec{r}_2) \, dr \equiv H_{12}
\]
\[
\int u_{100}(\vec{r}_2) \hat{H} u_{100}(\vec{r}_1) \, dr \equiv H_{21}
\]
\[
\int u_{100}(\vec{r}_1) u_{100}(\vec{r}_1) \, dr \equiv S_{11} = 1
\]
\[
\int u_{100}(\vec{r}_2) u_{100}(\vec{r}_2) \equiv S_{22} = 1
\]
since \( u_{100}(\vec{r}_i) \) is orthonormal

\[
\int u_{100}(\vec{r}_1) u_{100}(\vec{r}_2) \, dr \equiv S_{12}
\]
\[
\int u_{100}(\vec{r}_2) u_{100}(\vec{r}_1) \, dr \equiv S_{21}
\]
\[\neq 0 \text{ overlap integral } u_{100}(\vec{r}_i) \, u_{100}(\vec{r}_j) \text{ are non-orthogonal}
\]

**The coulomb energy integral:**

\[
\int u_{100}(\vec{r}_i) \hat{H} u_{100}(\vec{r}_i) \, dr \equiv H_{11} = H_{22}
\]
\[
= \int u_{100}(\vec{r}_i) \left( -\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{r_i} \right) u_{100}(\vec{r}_i) \, dr - \int u_{100}(\vec{r}_i) \left( \frac{e^2}{r_i} \right) u_{100}(\vec{r}_i) \, dr
\]

The second term in the coulomb integral describes the attractive electrostatic interaction between the proton B and the electron around proton A tends to lower the energy.

**The resonance energy contribution:**

\[
\int u_{100}(\vec{r}_i) \hat{H} u_{100}(\vec{r}_i) \, dr \equiv H_{12} = H_{21}
\]
\[
= \int u_{100}(\vec{r}_i) \left( -\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{r_i} \right) u_{100}(\vec{r}_i) \, dr - \int u_{100}(\vec{r}_i) \left( \frac{e^2}{r_i} \right) u_{100}(\vec{r}_i) \, dr
\]

The interaction leads to the splitting and the creation of two energy levels

**The overlap integral** tends to destabilize the molecule.

Minimizing the energy expectation value with respect to the coefficients leads us to the determinant problem,

\[
(H_{11} - ES_{11})c_1 + (H_{12} - ES_{12})c_2 = 0
\]
\[
(H_{21} - ES_{21})c_1 + (H_{22} - ES_{22})c_2 = 0
\]
\[
\begin{bmatrix}
H_{11} & H_{12} \\
H_{21} & H_{22}
\end{bmatrix}
\begin{bmatrix}
c_1 \\
c_2
\end{bmatrix}
= E
\begin{bmatrix}
S_{11} & S_{12} \\
S_{21} & S_{22}
\end{bmatrix}
\begin{bmatrix}
c_1 \\
c_2
\end{bmatrix}
\]

\[
\begin{bmatrix}
H_{11} - ES_{11} & H_{12} - ES_{12} \\
H_{21} - ES_{21} & H_{22} - ES_{22}
\end{bmatrix}
= 0
\]

\[
\begin{pmatrix}
1 \\
1
\end{pmatrix}
\rightarrow
E_+ = \frac{H_{11} + H_{12}}{1 + S_{12}}
\]

\[
\begin{pmatrix}
1 \\
-1
\end{pmatrix}
\rightarrow
E_- = \frac{H_{11} - H_{12}}{1 - S_{12}}
\]

\[
E_+ + \frac{e^2}{R}
\]

\[
E_- + \frac{e^2}{R}
\]

Energy (a.u.= 27.21 eV)
We get two solutions for the approximate eigenfunctions,

\[
\psi_\sigma (\vec{r}) = \frac{1}{\sqrt{2(1+S)}} \left( u_{100} (\vec{r}_1) + u_{100} (\vec{r}_2) \right)
\]

\[
\psi_{\sigma^*} (\vec{r}) = \frac{1}{\sqrt{2(1-S)}} \left( u_{100} (\vec{r}_1) - u_{100} (\vec{r}_2) \right)
\]

\[E_\sigma < E_{\sigma^*}\]

Consequences:

1. Odd and even wave functions the odd wavefunction has a node which is located between the two nuclei.
2. Draw the wavefunction evaluated along the internuclear separation line.
3. Splitting of the energy levels as the internuclear spacing decreases.
4. The total energy of the molecule (including the proton-proton) repulsion.

![Diagram of bonding and antibonding MOs](image)

1. The lower electron eigenfunction or state has a corresponding minimum in the total energy while the higher energy electron state rises monotonically.
2. The internuclear distance at which the system achieves a minimum is the bond length which is $2R_0$.
3. The energy minimum in the lower curve indicates a stable molecule and the corresponding MO is called bonding MO. The other MO is called antibonding.
4. The electron in the bonding state will have a wavefunction that leads to a high probability of finding the electron in between the nuclei – this allows for the attractive e-p forces to offset the repulsive p-p interactions.

5. Near the minimum energy,
   $$E = \frac{K}{2}(R - R_0)^2$$

K – is the bond stiffness.
6. In diatomic MO we distinguished between \( \pi \) and \( \sigma \) type bonds. What is the significance of this labeling scheme? Recall that the AO were labeled with \( s, p, d \) labels which were associated with the orbital angular momentum and the spherically symmetric potential. Since \( L^2 \) commuted with the Hamiltonian we could find a basis set which was common to \( \hat{L}^2, \hat{L}_z \) and the Hamiltonian. In a linear molecule the rotational symmetry around the \( z \) axis is conserved hence there is a conservation of the \( \hat{L}_z \) component of the angular momentum that is why the MO are labeled according to their symmetries.