

4 / Basis functions

4.1. As a computational tool

4.2. As a conceptual tool

4.3. Equilibrium density matrix

4.4. Supplementary notes:

Density matrix, Perturbation theory

We have seen that it is straightforward to calculate the energy levels for atoms using the SCF method, because the spherical symmetry effectively reduces it to a one-dimensional problem. Molecules on the other hand do not have this spherical symmetry and a more efficient approach is needed to make the problem numerically tractable. The concept of basis functions provides a convenient computational tool for solving the Schrodinger equation (or any differential equation for that matter). At the same time it is also a very important conceptual tool that is fundamental to the quantum mechanical viewpoint. In this chapter we attempt to convey both these aspects.

The basic idea is that the wavefunction can in general be expressed in terms of a set of basis functions, $u_m(\vec{r})$

$$\Phi(\vec{r}) = \sum_{m=1}^M \phi_m u_m(\vec{r})$$

We can then represent the wavefunction by a column vector consisting of the expansion coefficients

$$\Phi(\vec{r}) \rightarrow \{\phi_1 \quad \phi_2 \quad \cdots \quad \phi_M\}$$

In spirit, this is not too different from what we did in Chapter 1 where we represented the wavefunction by its values at different points on a discrete lattice:

$$\Phi(\vec{r}) \rightarrow \{\Phi(\vec{r}_1) \quad \Phi(\vec{r}_2) \quad \cdots \quad \Phi(\vec{r}_M)\}$$

However, the difference is that now we have the freedom to choose the basis function $u_m(\vec{r})$: if we choose them so that they look much like our expected wavefunction, then we can represent the wavefunction accurately with just a few terms, thereby reducing the size of the resulting matrix $[H]$ greatly. This makes the approach useful as a computational tool (similar in spirit to the concept of “shape functions” in the finite element method [2.2]) as we illustrate with a simple example in **Section 4.1**.

But the concept of basis functions is far more general. One can view them as the coordinate axes in an abstract Hilbert space as described in **Section 4.2** and we will illustrate the power and versatility of this viewpoint in Section 4.3 using the concept of density matrix in **Section 4.3**.

4.1. (Basis functions) as a computational tool

The basic formulation can be stated fairly simply. We write the wavefunction in terms of any set of basis functions $u_m(\vec{r})$:

$$\Phi(\vec{r}) = \sum_m \phi_m u_m(\vec{r}) \quad (4.1.1)$$

and substitute it into the Schrodinger equation $E \Phi(\vec{r}) = H_{op} \Phi(\vec{r})$ to obtain

$$E \sum_m \phi_m u_m(\vec{r}) = \sum_m \phi_m H_{op} u_m(\vec{r})$$

Multiply both sides by $u_n^*(\vec{r})$ and integrate over all $\vec{r}$ to yield

$$E \sum_m S_{nm} \phi_m = \sum_m H_{nm} \phi_m$$

which can be written as a matrix equation

$$E [S] \{\phi\} = [H] \{\phi\} \quad (4.1.2)$$

$$\text{where } S_{nm} = \int d\vec{r} u_n^*(\vec{r}) u_m(\vec{r}) \quad (4.1.3a)$$

$$H_{nm} = \int d\vec{r} u_n^*(\vec{r}) H_{op} u_m(\vec{r}) \quad (4.1.3b)$$

To proceed further we have to evaluate the integrals and that is the most time-consuming step in the process. But once the matrix elements have been calculated, it is straightforward to obtain the eigenvalues E_α and eigenvectors $\{\Phi_\alpha\}$ of the matrix. The eigenfunctions can then be written down in “real space” by substituting the coefficients back into the original expansion in Eq.(4.1.1):

$$\Phi_\alpha(\vec{r}) = \frac{1}{\sqrt{Z_\alpha}} \sum_n \phi_{\alpha n} u_n(\vec{r}) \quad (4.1.4)$$

where Z_α is a constant chosen to ensure proper normalization:

$$1 = \int d\vec{r} \Phi_\alpha^*(\vec{r}) \Phi_\alpha(\vec{r}) \rightarrow Z_\alpha = \sum_n \sum_m \phi_{n\alpha}^* \phi_{m\alpha} S_{nm} \quad (4.1.5)$$

Eqs.(4.1.1)-(4.1.5) summarize the basic mathematical relations involved in the use of basis functions.

A specific example: To understand the underlying physics and how this works in practice let us look at a specific example. In Section 3.3.3 we stated that the lowest energy eigenvalue of the Schrodinger equation including the two nuclear potentials (Fig.4.1.1) but excluding the self-consistent potential

$$E_{\alpha 0} \Phi_{\alpha 0}(\vec{r}) = \left(-\frac{\hbar^2}{2m} \nabla^2 + U_N(\vec{r}) + U_{N'}(\vec{r}) \right) \Phi_{\alpha 0}(\vec{r}) \quad (4.1.6)$$

is approximately given by ($E_1 \equiv -E_0 = -13.6$ eV)

$$E_{B0} = E_1 + \frac{a+b}{1+s} \quad (4.1.7)$$

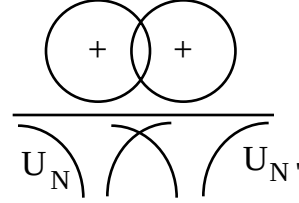
$$\text{where } a = -2E_0 \frac{1-(1+R_0)e^{-2R_0}}{R_0}, \quad b = -2E_0 (1+R_0) e^{-R_0}$$

$$\text{and } s = e^{-R_0} (1+R_0 + (R_0^2/3)), \quad R_0 \equiv R/a_0$$

R being the center-to-center distance between the hydrogen atoms

We will now use the concept of basis functions to show how this result is obtained from Eq.(4.1.6).

Fig.4.1.1. U_N and $U_{N'}$ are the Coulombic potentials due to the left and right nuclei of a H_2 molecule respectively.

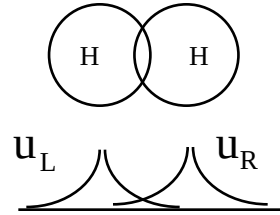


Note that the potential $U(\vec{r}) = U_N(\vec{r}) + U_{N'}(\vec{r})$ in Eq.(4.1.6) is not spherically symmetric, unlike the atomic potentials we discussed in Chapters 2 and 3. This means that we cannot simply solve the radial Schrodinger equation. In general, we have to solve the full three-dimensional Schrodinger equation, which is numerically quite challenging and the problem is made tractable by using basis functions to expand the wavefunction. In the present case we can use just two basis functions

$$\Phi(\vec{r}) = \phi_L u_L(\vec{r}) + \phi_R u_R(\vec{r}) \quad (4.1.8)$$

where $u_L(\vec{r})$ and $u_R(\vec{r})$ represent a hydrogenic 1s orbital centered around the left and right nuclei respectively (see Fig.4.1.2).

Fig.4.1.2. A convenient basis set for the H_2 molecule consists of two '1s' orbitals centered around the left and right nuclei respectively.



This means that

$$E_1 u_L(\vec{r}) = \left(-\frac{\hbar^2}{2m} \nabla^2 + U_N(\vec{r}) \right) u_L(\vec{r}) \quad (4.1.9a)$$

$$\text{and} \quad E_1 u_R(\vec{r}) = \left(-\frac{\hbar^2}{2m} \nabla^2 + U_{N'}(\vec{r}) \right) u_R(\vec{r}) \quad (4.1.9b)$$

The ansatz in Eq.(4.1.8) is motivated by the observation that it clearly describes the eigenstates correctly if we move the two atoms far apart: the eigenstates are then given by

$$(\phi_L \ \phi_R) = (1 \ 0) \quad \text{and} \quad (\phi_L \ \phi_R) = (0 \ 1)$$

It seems reasonable to expect that if the bond length R is not too short (compared to the Bohr radius a_0) Eq.(4.1.8) will still provide a reasonably accurate description of the correct eigenstates with an appropriate choice of the coefficients $(\phi_L \ \phi_R)$.

Since we have used only two functions u_L and u_R to express our wavefunction, the matrices $[S]$ and $[H]$ in Eq.(4.1.2) are simple (2x2) matrices whose elements can be written down from Eq.(4.1.3a,b) making use of Eqs.(4.1.9a,b):

$$S = \begin{bmatrix} 1 & s \\ s & 1 \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} E_1 + a & E_1 s + b \\ E_1 s + b & E_1 + a \end{bmatrix} \quad (4.1.10)$$

$$\text{where} \quad s \equiv \int d\vec{r} u_L^*(\vec{r}) u_R(\vec{r}) = \int d\vec{r} u_R^*(\vec{r}) u_L(\vec{r}) \quad (4.1.11a)$$

$$a \equiv \int d\vec{r} u_L^*(\vec{r}) U_{N'}(\vec{r}) u_L(\vec{r}) = \int d\vec{r} u_R^*(\vec{r}) U_N(\vec{r}) u_R(\vec{r}) \quad (4.1.11b)$$

$$\begin{aligned} b &\equiv \int d\vec{r} u_L^*(\vec{r}) U_N(\vec{r}) u_R(\vec{r}) = \int d\vec{r} u_L^*(\vec{r}) U_{N'}(\vec{r}) u_R(\vec{r}) \\ &= \int d\vec{r} u_R^*(\vec{r}) U_N(\vec{r}) u_L(\vec{r}) = \int d\vec{r} u_R^*(\vec{r}) U_{N'}(\vec{r}) u_L(\vec{r}) \end{aligned} \quad (4.1.11c)$$

Hence Eq.(4.2.2) becomes

$$E \begin{pmatrix} \phi_L \\ \phi_R \end{pmatrix} = \begin{bmatrix} 1 & s \\ s & 1 \end{bmatrix}^{-1} \begin{bmatrix} E_1 + a & E_1 s + b \\ E_1 s + b & E_1 + a \end{bmatrix} \begin{pmatrix} \phi_L \\ \phi_R \end{pmatrix} \quad (4.1.12)$$

from which it is straightforward to write down the two eigenvalues - the lower one is called the bonding level (B) and the higher one is called the anti-bonding level (A) :

$$E_B = E_1 + \frac{a+b}{1+s} \quad \text{and} \quad E_A = E_1 + \frac{a-b}{1-s} \quad (4.1.13)$$

The quantities a , b and s can be evaluated by plugging in the known basis functions $u_L(\vec{r})$, $u_R(\vec{r})$ and the nuclear potentials $U_N(\vec{r})$ and $U_N'(\vec{r})$ into Eqs.(4.1.11a,b,c). The integrals can be performed analytically to yield the results stated earlier in Eq.(4.1.7).

The wavefunctions corresponding to the bonding and antibonding levels are given by

$$(\phi_L \ \phi_R)_B = (1 \ 1) \quad \text{and} \quad (\phi_L \ \phi_R)_A = (1 \ -1)$$



which represent a symmetric (B) and an antisymmetric (A) combination of two 1s orbitals centered around the two nuclei. Both electrons in a H_2 molecule occupy the symmetric or bonding state whose wavefunction can be written as

$$\Phi_{B0}(\vec{r}) = \frac{1}{\sqrt{Z}} [u_L(\vec{r}) + u_R(\vec{r})] \quad (4.1.14)$$

$$\text{where } u_L(\vec{r}) = \frac{1}{\sqrt{\pi a_0^3}} \exp\left[-\frac{|\vec{r} - \vec{r}_L|}{a_0}\right], \quad \vec{r}_L = -(R_0/2) \hat{z}$$

$$u_R(\vec{r}) = \frac{1}{\sqrt{\pi a_0^3}} \exp\left[-\frac{|\vec{r} - \vec{r}_R|}{a_0}\right], \quad \vec{r}_R = +(R_0/2) \hat{z}$$

The constant Z has to be chosen to ensure correct normalization of the wavefunction:

$$1 = \int d\vec{r} \Phi_{B0}^*(\vec{r}) \Phi_{B0}(\vec{r}) = \frac{2(1+s)}{Z} \rightarrow Z = 2(1+s)$$

The electron density $n(\vec{r})$ in a H_2 molecule is given by $|\Phi_{B0}(\vec{r})|^2$, multiplied by two since we have two electrons (one up-spin and one down-spin) with this wavefunction. Fig.4.1.3 shows a plot of the electron density along a line joining the two nuclei.

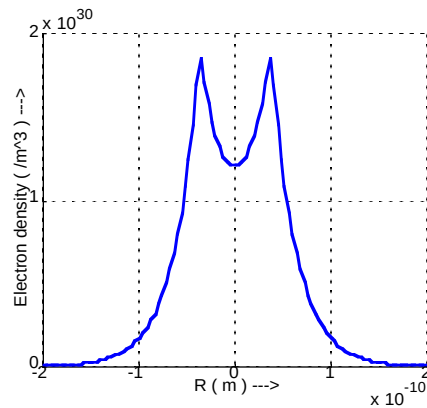


Fig.4.1.3. Plot of electron density along the axis joining the two hydrogen atoms assuming they are separated by the equilibrium bond distance of $R = 0.074$ nm.

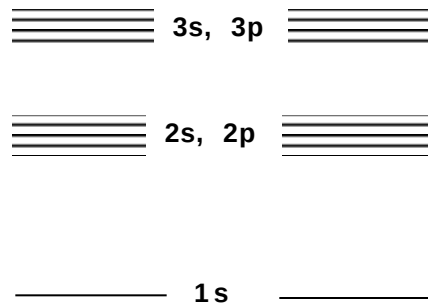
How can we get accurate results using just two basis functions ? If we were to start from the Schrodinger equation and use a discrete lattice representation as we did in Chapters 1 and 2, then we would need a fairly large number of basis functions per atom. For example if the lattice points are spaced by $0.5 \text{ \AA}$ and the size of an atom is $2.5 \text{ \AA}$, then we need $5^3 = 125$ lattice points (each of which represents a basis function), since the problem is a three-dimensional one. What do we lose by using only one basis function instead of 125 ? The answer is that our results are accurate only over a limited range of energies.

To see this, suppose we were to use not just the $1s$ orbital as we did, but also the $2s, 2p_x, 2p_y, 2p_z, 3s, 3p_x, 3p_y$ and $3p_z$ orbitals (see Fig.4.1.5). We argue that the lowest eigenstates will still be essentially made up of ' $1s$ ' wavefunctions and will involve negligible amounts of the other wavefunctions, so that fairly accurate results can be obtained with just one basis function per atom. The reason is that an off-diagonal matrix element M modifies the eigenstates of a matrix

$$\begin{bmatrix} E_1 & M \\ M & E_2 \end{bmatrix}$$

significantly only if it is comparable to the difference between the diagonal elements, that is, if $M \geq |E_1 - E_2|$. The diagonal elements are roughly equal to the energy levels of the isolated atoms, so that $|E_1 - E_2|$ is ~ 10 eV if we consider say the '1s' and the '2s' levels of a hydrogen atom. The off-diagonal element M depends on the proximity of the two atoms and for typical covalently bonded molecules and solids is ~ 2 eV, which is smaller than $|E_1 - E_2|$. As a result the bonding level is primarily composed of '1s' wavefunctions and our treatment based on a 2x2 matrix is fairly accurate. But a proper treatment of the higher energy levels would require more basis functions to be included.

Fig.4.1.5. Eigenstates of a H₂ molecule when the atoms are not too close. All these states could be used as basis functions for a more accurate treatment of the hydrogen molecule.



Ab-initio and semi-empirical models : The concept of basis functions is widely used for ab initio calculations where the Schrodinger equation is solved directly including a self-consistent field. For large molecules or solids such calculations can be computationally quite intensive due to the large number of basis functions involved and the integrals that have to be evaluated to obtain the matrix elements. The integrals arising from the self-consistent field are particularly time consuming. For this reason, semiempirical approaches are widely used where the matrix elements are adjusted through a combination of theory and experiment. Such semi-empirical approaches can be very useful if the parameters turn out to be "transferable", that is, if we can obtain them by fitting one set of observations and then use them to make predictions in other situations. For example, we could calculate suitable parameters to fit the known energy levels of an infinite solid and then use these parameters to calculate the energy levels in a finite nanostructure carved out of that solid.

4.2. (Basis functions) as a conceptual tool

We have mentioned that all practical methods for solving the energy levels of molecules and solids usually involve some sort of expansion in basis functions. However, the concept of basis functions is more than a computational tool. It represents an important conceptual tool for visualizing the physics and developing an intuition for what to expect. Indeed the concept of a wavefunction as a superposition of basis functions is central to the entire structure of quantum mechanics as we will try to explain next.

Vector space vs. Hilbert space : It is useful to compare Eq.(4.1.1) with the expression for an ordinary three-dimensional vector $\vec{V}$ in terms of the three unit vectors $\hat{x}$, $\hat{y}$ and $\hat{z}$:

$$\vec{V} = V_x \hat{x} + V_y \hat{y} + V_z \hat{z} \quad \text{<----->} \quad \Phi = \phi_1 u_1 + \phi_2 u_2 + \phi_3 u_3 + \dots$$

We can view the wavefunction as a state vector in an N-dimensional space called the Hilbert space, N being the total number of basis functions $u_m(\vec{r})$. The ϕ_m 's in Eq.(4.1.1) are like the components of the state vector Φ while the $u_m(\vec{r})$'s are the associated unit vectors along the N coordinate axes. Choosing a different set of basis functions $u_m(\vec{r})$ is like choosing a different coordinate system: the components ϕ_m along the different axes all change, though the state vector remains the same. In principle, N is infinite, but in practice we can often get accurate results with a manageably finite value of N. We have tried to depict this analogy in Figs.4.2.1a and b but it is difficult to do justice to an N-dimensional vector ($N > 3$) on two-dimensional paper. In the Dirac notation, which is very convenient and widely used, the state-vector associated with wavefunction $\Phi(\vec{r})$ is denoted by a "ket" $|\Phi\rangle$ and the unit vectors associated with the basis functions $u_m(\vec{r})$ are also written as kets $|m\rangle$. In this notation the expansion in terms of basis functions (see Eq.(4.1.11)) is written as

$$\begin{array}{ccc} & \text{(Dirac notation)} & \\ \Phi(\vec{r}) = \sum_m \phi_m u_m(\vec{r}) & \text{----->} & |\Phi\rangle = \sum_m \phi_m |m\rangle \quad (4.2.1) \end{array}$$

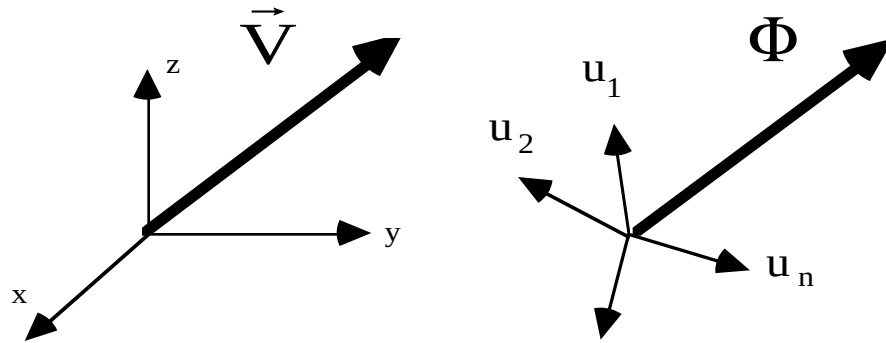


Fig.4.2.1. (a) An ordinary vector $\vec{V}$ in three-dimensional space can be expressed in terms of its components along x , y and z . **(b)** The state-vector $\Phi(\vec{r})$ can be expressed in terms of its components along the basis functions $u_n(\vec{r})$.

Scalar product: A central concept in vector algebra is that of the scalar product :

$$\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z = \sum_{m=x,y,z} A_m B_m$$

The corresponding concept in Hilbert space is that of the overlap of any two functions $f(\vec{r})$ and $g(\vec{r})$:

$$\int d\vec{r} f^*(\vec{r}) g(\vec{r}) \quad \xrightarrow{\text{(Dirac notation)}} \quad \langle f | g \rangle$$

The similarity of the overlap integral to a scalar product can be seen by discretizing the integral :

$$\int d\vec{r} f^*(\vec{r}) g(\vec{r}) \cong a^3 \sum_m f^*(\vec{r}_m) g(\vec{r}_m)$$

In the discrete lattice representation (see Fig.4.1.4) the "component", f_m of $f(\vec{r})$ along $u_m(\vec{r})$ is given by $f(\vec{r}_m)$ just as A_m represents the component of the vector $\vec{A}$ along $\hat{m}$:

$$\int d\vec{r} f^*(\vec{r}) g(\vec{r}) \equiv a^3 \sum_m f_m^* g_m \quad \text{cf.} \quad \vec{A} \cdot \vec{B} = \sum_{m=x,y,z} A_m B_m$$

One difference here is that we take the complex conjugate of one of the functions (this is not important if we are dealing with real functions) which is represented by the "bra" $\langle f|$ as opposed to the "ket" $|g\rangle$. The scalar product is represented by juxtaposing a "bra" and a "ket" as in $\langle f|g\rangle$.

Orthogonality : Coordinate systems are said to be orthogonal if $\hat{n} \cdot \hat{m} = \delta_{nm}$, where the indices m, n stand for x, y or z and δ_{nm} is the Kronecker delta which is defined as

$$\begin{aligned} \delta_{nm} &= 1 & \text{if } n = m \\ &= 0 & \text{if } n \neq m \end{aligned} \quad (4.2.2)$$

This is usually true (for example, $\hat{x} \cdot \hat{y} = \hat{y} \cdot \hat{z} = \hat{z} \cdot \hat{x} = 0$) but it is possible to work with non-orthogonal coordinate systems too. Similarly the basis functions $u_m(\vec{r})$ are said to be orthogonal if the following relation is satisfied :

$$\begin{aligned} & \text{(Dirac notation)} \\ \int d\vec{r} u_n^*(\vec{r}) u_m(\vec{r}) &= \delta_{nm} \quad \text{-----} > \quad \langle n|m \rangle = \delta_{nm} \end{aligned} \quad (4.2.3)$$

Note that the basis functions corresponding to the discrete lattice representation (Fig.4.1.4) are orthogonal, but the one we used for the hydrogen molecule (see Fig.4.1.2) are NON-orthogonal since

$$\int d\vec{r} u_L^*(\vec{r}) u_R(\vec{r}) \equiv s = e^{-R_0} (1 + R_0 + (R_0^2 / 3)) \neq 0.$$

In general

$$\begin{aligned} & \text{(Dirac notation)} \\ \int d\vec{r} u_n^*(\vec{r}) u_m(\vec{r}) &= S_{nm} \quad \text{-----} > \quad \langle n|m \rangle = S_{nm} \end{aligned} \quad (4.2.4)$$

Orthogonalization: Given a non-orthogonal set of basis functions $\{u_n(\vec{r})\}$, we can define another set

$$\tilde{u}_i(\vec{r}) = \sum_n \left[S^{-1/2} \right]_{ni} u_n(\vec{r}) \quad (4.2.5)$$

which will be orthogonal. This is shown as follows

$$\begin{aligned} \int d\vec{r} \tilde{u}_i^*(\vec{r}) \tilde{u}_j(\vec{r}) &= \sum_n \sum_m \left[S^{-1/2} \right]_{in} S_{nm} \left[S^{-1/2} \right]_{mj} \\ &= \left[S^{-1/2} S S^{-1/2} \right]_{ij} \\ &= \delta_{ij} \end{aligned}$$

where we have made use of Eq.(4.2.4). This means that if we use the new set $\{\tilde{u}_i(\vec{r})\}$ as our basis, then the overlap matrix $[S] = [I]$, where $[I]$ is the identity matrix which is a diagonal matrix with ones on the diagonal. This is a property of orthogonal bases functions which makes it conceptually easier to deal with.

Even if we start with a non-orthogonal basis, it is often convenient to orthogonalize it. What we might lose in the process is the local nature of the original basis which makes it convenient to visualize the physics. For example, the $\{u_n(\vec{r})\}$ we used for the hydrogen molecule were localized on the left and right hydrogen atoms respectively. But the orthogonalized basis $\{\tilde{u}_i(\vec{r})\}$ will be linear combinations of the two and thus less local than $\{u_n(\vec{r})\}$. As a rule, it is difficult to find basis functions that are both local and orthogonal. From hereon we will generally assume that the basis functions we use are orthogonal.

Operators : An operator like H_{op} acting on a state vector changes it into a different state vector - we could say that it "rotates" the vector. With ordinary vectors we can represent a rotation by a matrix :

$$\begin{Bmatrix} A'_x \\ A'_y \end{Bmatrix} = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} \begin{Bmatrix} A_x \\ A_y \end{Bmatrix}$$

where for simplicity we have assumed a two dimensional vector. How do we write down the matrix $[R]$ corresponding to an operator R_{op} ? The general principle is the following $R_{nm} = \hat{n} \cdot (R_{op} \hat{m})$. For example, suppose we consider an operator that rotates a vector by an angle θ . We then obtain

$$\begin{aligned} R_{xx} &= \hat{x} \cdot (R_{op} \hat{x}) = \hat{x} \cdot (\hat{x} \cos \theta + \hat{y} \sin \theta) = \cos \theta \\ R_{yx} &= \hat{y} \cdot (R_{op} \hat{x}) = \hat{y} \cdot (\hat{x} \cos \theta + \hat{y} \sin \theta) = \sin \theta \\ R_{xy} &= \hat{x} \cdot (R_{op} \hat{y}) = \hat{x} \cdot (-\hat{x} \sin \theta + \hat{y} \cos \theta) = -\sin \theta \\ R_{yy} &= \hat{y} \cdot (R_{op} \hat{y}) = \hat{y} \cdot (-\hat{x} \sin \theta + \hat{y} \cos \theta) = \cos \theta \end{aligned}$$

The matrix representation for any operator A_{op} in Hilbert space is written using a similar prescription :

(Dirac notation)

$$[A]_{nm} = \int d\vec{r} u_n^*(\vec{r}) (A_{op} u_m(\vec{r})) \quad \text{----->} \quad [A]_{nm} = \langle n | A_{op} | m \rangle \quad (4.2.5)$$

Constant operator : What is the matrix representing a constant operator, one that simply multiplies a state vector by a constant C ? In general, the answer is

$$[C]_{nm} = C \int d\vec{r} u_n^*(\vec{r}) u_m(\vec{r}) = C [S]_{nm} \quad (4.2.6)$$

which reduces to $C [I]$ for orthogonal bases.

Matrix representation of the Schrodinger equation: The matrix representation of the Schrodinger equation obtained at the end of the last section (see Eqs.(4.1.2), (4.1.3a,b))

$$E \Phi(\vec{r}) = H_{op} \Phi(\vec{r}) \quad \text{----->} \quad E [S] \{\phi\} = [H] \{\phi\} \quad (4.2.7)$$

can now be understood in terms of the concepts described in this section. Like the rotation operator in vector space, any differential operator in Hilbert space has a matrix representation. Once we have chosen a set of coordinates or basis functions, H_{op} becomes the matrix $[H]$ while the constant E becomes the matrix $E[S]$:

$$[S]_{nm} = \langle n|m \rangle \equiv \int d\vec{r} u_n^*(\vec{r}) u_m(\vec{r}) \quad (4.2.8a)$$

$$[H]_{nm} = \langle n|H_{op}|m \rangle \equiv \int d\vec{r} u_n^*(\vec{r}) (H_{op} u_m(\vec{r})) \quad (4.2.8b)$$

We could orthogonalize the basis set following Eq.(4.2.5), so that in terms of the orthogonal basis $\{\tilde{u}_i(\vec{r})\}$, the Schrodinger equation has the form of a standard matrix eigenvalue equation:

$$E \{\tilde{\phi}\} = [\tilde{H}] \{\tilde{\phi}\}$$

where the matrix elements of $[\tilde{H}]$ are given by

$$[\tilde{H}]_{ij} = \int d\vec{r} \tilde{u}_i^*(\vec{r}) (H_{op} \tilde{u}_j(\vec{r}))$$

Transformation of bases : Suppose we have expanded our wavefunction in one basis and would like to change to a different basis :

$$\Phi(\vec{r}) = \sum_m \phi_m u_m(\vec{r}) \quad \text{----->} \quad \Phi(\vec{r}) = \sum_i \phi'_i u'_i(\vec{r}) \quad (4.2.9)$$

Such a transformation can be described by a transformation matrix $[C]$ obtained by writing the new basis in terms of the old basis :

$$u'_i(\vec{r}) = \sum_m C_{mi} u_m(\vec{r}) \quad (4.2.10)$$

From Eqs.(4.2.9) and (4.2.10) we can show that

$$\phi_m = \sum_i C_{mi} \phi'_i \quad \text{----->} \quad \{\phi\} = [C] \{\phi'\} \quad (4.2.11a)$$

Similarly we can show that any matrix $[A']$ in the new representation is related to the matrix $[A]$ in the old representation by

$$A'_{ji} = \sum_j \sum_i C_{nj}^* A_{nm} C_{mi} \quad \longrightarrow \quad [A'] = [C]^+ [A] [C] \quad (4.2.11b)$$

Unitary transformation : There is a special class of transformations which conserves the "norm" of a state vector that is,

$$\sum_m \phi_m^* \phi_m = \sum_i \phi_i'^* \phi_i' \quad \longrightarrow \quad \{\phi\}^+ \{\phi\} = \{\phi'\}^+ \{\phi'\} \quad (4.2.12)$$

Substituting for $\{\phi\}$ from Eq.(4.2.11a) into Eq.(4.2.12)

$$\{\phi'\}^+ [C]^+ [C] \{\phi'\} = \{\phi'\}^+ \{\phi'\} \quad \longrightarrow \quad [C]^+ [C] = I \quad (4.2.13)$$

A matrix $[C]$ that satisfies this condition (Eq.(4.2.13)) is said to be unitary and the corresponding transformation is called a unitary transformation.

Hermitian operators : The matrix $[A]$ representing a Hermitian operator A_{Op} is Hermitian (in any representation) which means that it is equal to its conjugate transpose $[A^+]$:

$$[A] = [A]^+ \quad \text{that is, } A_{mn} = A_{nm}^* \quad (4.2.14)$$

If A_{Op} is a function like $U(\vec{r})$ then it is easy to show that it will be Hermitian as long as it is real.

$$[U]_{mn}^* = \left[\int d\vec{r} u_m^*(\vec{r}) U(\vec{r}) u_n(\vec{r}) \right]^* = [U]_{nm}$$

If A_{Op} is a differential operator like d/dx or d^2/dx^2 then it takes a little more work to check if it is Hermitian or not. An easier approach is to use the discrete lattice representation that we discussed in Chapter 2. Eq.(2.3.1) shows the matrix representing d^2/dx^2 and it is clearly Hermitian in this representation. And it can be shown that a matrix that is Hermitian in one representation will remain Hermitian in any other representation. The Hamiltonian operator is Hermitian since it is a sum of

Hermitian operators like d^2/dx^2 , d^2/dy^2 and $U(\vec{r})$. An important requirement of quantum mechanics is that the eigenvalues corresponding to any operator A_{op} representing any observable must be real. This is ensured by requiring all such operators A_{op} to be Hermitian (not just the Hamiltonian operator H_{op} which represents the energy). The eigenvalues of a Hermitian matrix are real.

Another useful property of a Hermitian matrix is that if we form a matrix $[V]$ out of all the normalized eigenvectors

$$[V] = [\{V_1\} \quad \{V_2\} \quad \dots \quad \dots]$$

then this matrix will be unitary, that is, $[V]^\dagger [V] = [I]$. Such a unitary matrix can be used to transform all column vectors $\{\phi\}$ and matrices $[M]$ to a new basis that uses the eigenvectors as the basis :

$$\{\phi\}_{new} = [V]^\dagger \{\phi\}_{old} \quad \longleftrightarrow \quad \{\phi\}_{old} = [V] \{\phi\}_{new} \quad (4.2.15)$$

$$[M]_{new} = [V]^\dagger [M]_{old} [V] \quad \longleftrightarrow \quad [M]_{old} = [V] [M]_{new} [V]^\dagger$$

If $[V]$ is the eigenvector matrix corresponding to a Hermitian matrix like $[H]$, then the new representation of $[H]$ will be diagonal with the eigenvalues E_m along the diagonal

$$[H'] = [V]^\dagger [H] [V] = \begin{bmatrix} E_1 & 0 & 0 & 0 & \dots \\ 0 & E_2 & 0 & 0 & \dots \\ 0 & 0 & E_3 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (4.2.16)$$

For this reason the process of finding eigenfunctions and eigenvalues is often referred to as diagonalization.

4.3. Equilibrium density matrix

The density matrix is one of the central concepts in statistical mechanics which properly belongs to Chapter 7. But the reason I am bringing it up in this chapter is that it provides an instructive example of the concept of basis functions. Let me start by briefly explaining what it means. In Chapter 3 we calculated the electron density, $n(\vec{r})$ in multielectron atoms by summing up the probability densities of each occupied eigenstate ‘ α ’ :

$$n(\vec{r}) = \sum_{\text{occ } \alpha} |\Phi_{\alpha}(\vec{r})|^2$$

This is true at low temperatures for closed systems having a fixed number of electrons which occupy the lowest available energy levels. In general, however, states can be partially occupied and in general the equilibrium electron density can be written as

$$n(\vec{r}) = \sum_{\alpha} f_0(\epsilon_{\alpha} - \mu) |\Phi_{\alpha}(\vec{r})|^2 \quad (4.3.1)$$

where $f_0(E) \equiv [1 + \exp(E/k_B T)]^{-1}$ is the Fermi function (Fig.1.1.3) indicating the extent to which a particular state is occupied: “0” indicates unoccupied states, “1” indicates occupied states, while a value between 0 and 1 indicates that the state is sometimes occupied and sometimes unoccupied.

Could we write a “wavefunction” $\Psi(\vec{r})$ for this multielectron system such that its squared magnitude will give us the electron density $n(\vec{r})$? One possibility is to write it as

$$\Psi(\vec{r}) = \sum_{\alpha} C_{\alpha} \Phi_{\alpha}(\vec{r}) \quad (4.3.2)$$

where $|C_{\alpha}|^2 = f_0(\epsilon_{\alpha} - \mu)$. But this is not quite right. If we square the magnitude of this multielectron “wavefunction” we obtain

$$n(\vec{r}) = |\Psi(\vec{r})|^2 = \sum_{\alpha} \sum_{\beta} C_{\alpha} C_{\beta}^* \Phi_{\alpha}(\vec{r}) \Phi_{\beta}(\vec{r})$$

which is equivalent to Eq.(4.3.1) if and only if,

$$|C_{\alpha}|^2 = f_0(\epsilon_{\alpha} - \mu) \equiv f_{\alpha}, \quad C_{\alpha} C_{\beta}^* = 0, \quad \alpha \neq \beta \quad (4.3.3)$$

This is impossible if we view the coefficients C_{α} as ordinary numbers – in that case $C_{\alpha} C_{\beta}^*$ must equal $\sqrt{f_{\alpha} f_{\beta}}$ and cannot be zero unless both C_{α} and C_{β} are zero. If we wish to write the multielectron wavefunction in the form shown in Eq.(4.3.2) we should view the coefficients C_{α} as stochastic numbers whose correlation coefficients are given by Eq.(4.3.3).

So instead of writing a wavefunction for multielectron systems, it is common to write down a complete matrix $\rho(\alpha, \beta)$ indicating the correlation $C_{\alpha} C_{\beta}^*$ between every pair of coefficients. This matrix ρ is called the density matrix and in the eigenstate representation, we can write its elements as (see Eq.(4.3.3))

$$\rho(\alpha, \beta) = f_{\alpha} \delta_{\alpha\beta} \quad (4.3.4)$$

where $\delta_{\alpha\beta}$ is the Kronecker delta defined as

$$\begin{aligned} \delta_{\alpha\beta} &= 1, & \alpha &= \beta \\ \delta_{\alpha\beta} &= 0, & \alpha &\neq \beta \end{aligned}$$

We can rewrite Eq.(4.3.1) for the electron density, $n(\vec{r})$ in the form

$$n(\vec{r}) = \sum_{\alpha} \sum_{\beta} \rho(\alpha, \beta) \Phi_{\alpha}(\vec{r}) \Phi_{\beta}^*(\vec{r}) \quad (4.3.5)$$

which can be generalized to define

$$\tilde{\rho}(\vec{r}, \vec{r}') = \sum_{\alpha} \sum_{\beta} \rho(\alpha, \beta) \Phi_{\alpha}(\vec{r}) \Phi_{\beta}^*(\vec{r}') \quad (4.3.6)$$

such that the electron density $n(\vec{r})$ is given by its diagonal elements:

$$n(\vec{r}) = \left[\rho(\vec{r}, \vec{r}') \right]_{\vec{r}' = \vec{r}} \quad (4.3.7)$$

Now the point I want to make is that Eq.(4.3.6) **represents a unitary transformation from an eigenstate basis to a real space basis**. This is seen by noting that the transformation matrix $[V]$ is obtained by writing each of the eigenstates (the old basis) as a column vector using the position (the new basis) representation :

$$[V]_{\vec{r}, \alpha} = \Phi_{\alpha}(\vec{r})$$

and that this matrix is unitary : $V^{-1} = V^{+}$

$$\Rightarrow [V^{-1}]_{\alpha, \vec{r}} = [V^{+}]_{\alpha, \vec{r}} = \Phi_{\alpha}^{*}(\vec{r})$$

so that Eq.(4.3.6) can be written in the form of a unitary transformation:

$$\tilde{\rho}(\vec{r}, \vec{r}') = \sum_{\alpha} \sum_{\beta} V(\vec{r}, \alpha) \rho(\alpha, \beta) V^{+}(\beta, \vec{r}') \Rightarrow \tilde{\rho} = V \rho V^{+}$$

This leads to a very powerful concept: The **density matrix** ρ **at equilibrium** can be written as **the Fermi function of the Hamiltonian matrix** (I: Identity matrix of the same size as H):

$$\rho = f_0(H - \mu I) \quad (4.3.8)$$

This is a general matrix relation that is valid **in any representation**. For example if we use the eigenstates α of H as a basis then $[H]$ is a diagonal matrix:

$$[H] = \begin{bmatrix} \epsilon_1 & 0 & 0 & \cdots \\ 0 & \epsilon_2 & 0 & \cdots \\ 0 & 0 & \epsilon_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

and so is $[\rho(E)]$:

$$[\rho(E)] = \begin{bmatrix} f_0(\epsilon_1 - \mu) & 0 & 0 & \dots \\ 0 & f_0(\epsilon_2 - \mu) & 0 & \dots \\ 0 & 0 & f_0(\epsilon_3 - \mu) & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

This is exactly what Eq.(4.3.4) tells us. But the point is that the relation given in Eq.(4.3.8) is valid, not just in the eigenstate representation, but in any representation. Given the matrix representation $[H]$, it takes just three commands in MATLAB to obtain the density matrix:

```
[V, D] = eig (H);
rho = 1./(1 + exp((diag(D) - mu)./kT));
rho = V * diag(rho) * V'
```

The first command calculates a diagonal matrix $[D]$ whose diagonal elements are the eigenvalues of $[H]$ and a matrix $[V]$ whose columns are the corresponding eigenvectors. In other words, $[D]$ is the Hamiltonian $[H]$ transformed to the eigenstate basis $\Rightarrow D = V^+ H V$

The second command gives us the density matrix in the eigenstate representation, which is easy since in this representation both $[H]$ and $[\rho]$ are diagonal. The third command then transforms $[\rho]$ back to the original representation.

Fig.4.3.1 shows the equilibrium electron density for a 1-D box modeled with a discrete lattice of 100 points spaced by $2A$, with $\mu = 0.25$ eV. The Hamiltonian $[H]$ is a (100x100) matrix which can be set up following the prescription in Sections 2.2, 2.3. the density matrix is then evaluated as described above and its diagonal elements give us the electron density $n(x)$ (times the lattice constant, 'a'). Note the standing wave patterns in (c) and (d) which are absent when we use periodic boundary conditions as seen from (e) and (f). Figs.(e) and (f) also show the standing wave patterns in the electron density when a large repulsive potential

$$U_0 \delta(x - (L/2)) \quad \text{where } U_0 = 20 \text{ eV-A}$$

is included at the center of the box.

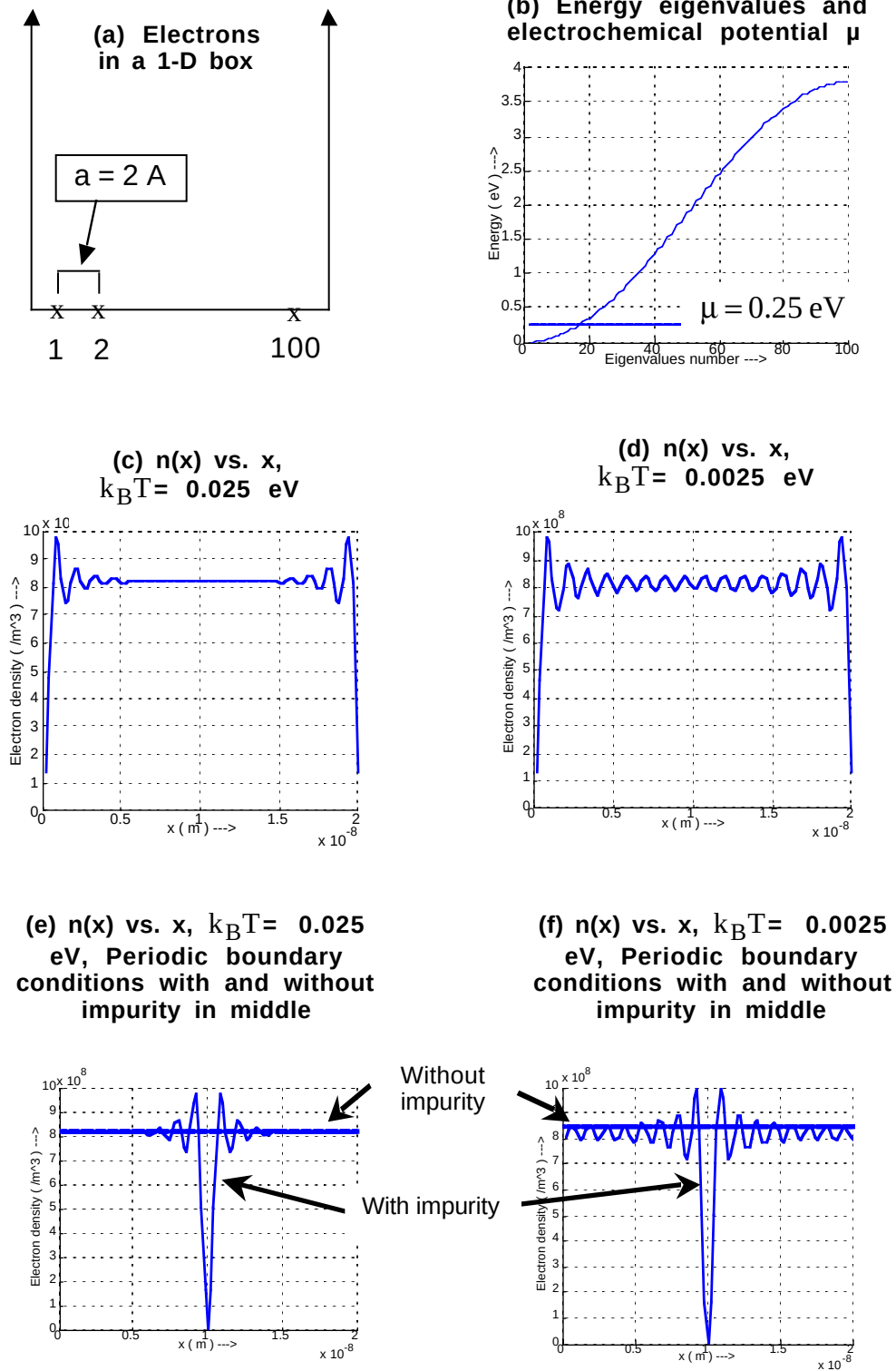


Fig.4.3.1. Equilibrium electron density for a 1-D box modeled with a discrete lattice of 100 points spaced by $2\text{\AA}$.

Note that the density matrix can look very different depending on what basis functions we use. In the eigenstate representation it is diagonal since the Hamiltonian is diagonal, but in the real space lattice representation it has off-diagonal elements. In any basis 'm', the diagonal elements $\rho(m,m)$ tell us the number of electrons occupying the state 'm'. In a real space representation, the diagonal elements $\rho(\vec{r},\vec{r})$ give us the electron density $n(\vec{r})$. The trace (sum of diagonal elements) of ρ , which is invariant in all representations, gives us the total number of electrons, N:

$$N = \text{Trace}(\rho) \quad (4.3.9)$$

If we are only interested in the electron density, then the diagonal elements of the density matrix are all we need. But we cannot “throw away” the off-diagonal elements; they are needed to ensure that the matrix will transform correctly to another representation. Besides, depending on what we wish to calculate, we may need the off-diagonal elements too (see Section 4.4.1).

4.4. Supplementary notes

4.4.1. Density matrix

It is common in quantum mechanics to associate every observable ‘A’ with an operator A_{op} for which we can find a matrix representation $[A]$ in any basis. The expectation value $\langle A \rangle$ for this observable (that is, the average value we expect to get in a series of measurements) is given by

$$\langle A \rangle = \int d\vec{r} \Psi^* A_{op} \Psi$$

Substituting for the wavefunction in terms of the basis functions from Eq.(4.3.2), we can show that

$$\langle A \rangle = \sum_{\alpha} \sum_{\beta} C_{\alpha} C_{\beta}^* \int d\vec{r} \Phi_{\beta}^* A_{op} \Phi_{\alpha}$$

$$\text{so that} \quad \langle A \rangle = \sum_{\alpha} \sum_{\beta} \rho_{\alpha\beta} A_{\beta\alpha} = \text{Trace} [\rho A]$$

We could use this result to evaluate the expectation value of any quantity, even if the system is out of equilibrium, provided we know the density matrix. But what we have discussed here is the equilibrium density matrix. It is much harder to calculate the non-equilibrium density matrix as we will discuss later in the book.

Plane-wave versus sine-cosine representations: Consider now a conductor of length ‘L’ having just two plane wave (‘pw’) states

$$\Psi_{+}(x) = \frac{1}{\sqrt{L}} e^{+ikx} \quad \text{and} \quad \Psi_{-}(x) = \frac{1}{\sqrt{L}} e^{-ikx}$$

The current operator in this basis is given by

$$[J_{op}]_{\pm} = \frac{-q}{L} \begin{bmatrix} '+' & '-' \\ \hbar k/m & 0 \\ 0 & -\hbar k/m \end{bmatrix}$$

and we could write the density matrix as

$$[\rho]_{\pm} = \begin{bmatrix} ' +' & ' - ' \\ f_{+} & 0 \\ 0 & f_{-} \end{bmatrix}$$

where f_{+} and f_{-} are the occupation probabilities for the two states. We wish to transform both these matrices from the ' $\pm$ ' basis to a 'cs' basis using cosine and sine states:

$$\Psi_c(x) = \sqrt{\frac{2}{L}} \cos kx \quad \text{and} \quad \Psi_s(x) = \sqrt{\frac{2}{L}} \sin kx$$

It is straightforward to write down the transformation matrix $[V]$ whose columns represent the old basis (+,-) in terms of the new basis (c,s):

$$[V] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ +i & -i \end{bmatrix}$$

so that in the 'cs' representation

$$[\rho]_{cs} = [V][\rho]_{\pm}[V]^{\dagger} = \frac{1}{2} \begin{bmatrix} 'c' & 's' \\ f_{+} + f_{-} & -i(f_{+} - f_{-}) \\ -i(f_{+} - f_{-}) & f_{+} + f_{-} \end{bmatrix}$$

$$\text{and} \quad [J_{op}]_{cs} = [V][J_{op}]_{\pm}[V]^{\dagger} = \begin{bmatrix} 'c' & 's' \\ 0 & -i\hbar k/mL \\ +i\hbar k/mL & 0 \end{bmatrix}$$

It is easy to check that the current $\langle J \rangle = \text{Trace} [\rho J_{op}]$ is the same in either representation:

$$\langle J \rangle = (-q/L) (\hbar k/m) [f_{+} - f_{-}]$$

This is expected since the trace is invariant under a unitary transformation and thus remains the same, no matter which representation we use. But the point to note is that the current in the cosine-sine representation arises from the **off-diagonal** elements of the current operator and the density matrix, rather than the **diagonal** elements. The off-diagonal elements do not have an intuitive physical meaning like the diagonal elements. As long as the current is carried by the diagonal elements, we can use a semiclassical picture in terms of occupation probabilities. But if the ‘action’ is in the off-diagonal elements then we need a more general quantum framework. (I am indebted to A.W. Overhauser for suggesting this example)

4.4.2. Perturbation theory

Suppose we wish to find the energy levels of a hydrogen atom in the presence of an electric field F applied along the z -direction. Let us use the eigenstates $1s$, $2s$, $2p_x$, $2p_y$ and $2p_z$ as our basis set and write down the Hamiltonian matrix. If the field were absent the matrix would be diagonal :

$$[H_0] = \begin{matrix} & \begin{matrix} 1s & 2s & 2p_x & 2p_y & 2p_z \end{matrix} \\ \begin{matrix} 1s \\ 2s \\ 2p_x \\ 2p_y \\ 2p_z \end{matrix} & \begin{bmatrix} E_1 & 0 & 0 & 0 & 0 \\ 0 & E_2 & 0 & 0 & 0 \\ 0 & 0 & E_2 & 0 & 0 \\ 0 & 0 & 0 & E_2 & 0 \\ 0 & 0 & 0 & 0 & E_2 \end{bmatrix} \end{matrix}$$

where $E_0 = 13.6$ eV, $E_1 = -E_0$ and $E_2 = -E_0/4$. The electric field leads to a matrix $[H_F]$ which has to be added to $[H]$. Its elements are given by

$$[H_F]_{nm} = qF \int_0^\infty dr r^2 \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi u_n^*(\vec{r}) r \cos\theta u_m(\vec{r})$$

Using the wavefunctions

$$u_{1s} = \sqrt{1/\pi a_0^3} e^{-r/a_0}$$

$$u_{2s} = \sqrt{1/32\pi a_0^3} \left(2 - \frac{r}{a_0}\right) e^{-r/2a_0}$$

$$\begin{aligned}
u_{2p_x} &= \sqrt{1/16\pi a_0^3} \left(\frac{r}{a_0} \right) e^{-r/2a_0} \sin\theta \cos\phi \\
u_{2p_y} &= \sqrt{1/16\pi a_0^3} \left(\frac{r}{a_0} \right) e^{-r/2a_0} \sin\theta \sin\phi \\
u_{2p_z} &= \sqrt{1/32\pi a_0^3} \left(\frac{r}{a_0} \right) e^{-r/2a_0} \cos\theta
\end{aligned}$$

we can evaluate the integrals straightforwardly to show that

$$\begin{array}{ccccc}
1s & 2s & 2p_x & 2p_y & 2p_z \\
[H_F] = & \begin{bmatrix} 0 & 0 & 0 & 0 & A \\ 0 & 0 & 0 & 0 & B \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ A & B & 0 & 0 & 0 \end{bmatrix}
\end{array}$$

where A and B are linear functions of the field:

$$A = (128\sqrt{2}/243) a_0 F, \quad B = -3a_0 F$$

Hence

$$\begin{array}{ccccc}
1s & 2s & 2p_z & 2p_x & 2p_y \\
[H_0] + [H_F] = & \begin{bmatrix} E_1 & 0 & A & 0 & 0 \\ 0 & E_2 & B & 0 & 0 \\ A & B & E_2 & 0 & 0 \\ 0 & 0 & 0 & E_2 & 0 \\ 0 & 0 & 0 & 0 & E_2 \end{bmatrix}
\end{array}$$

Note that we have relabeled the rows and columns to accentuate the fact that $2p_x$ and $2p_y$ levels are decoupled from the rest of the matrix and are **unaffected** by the field, while the $1s$ and $2s$ and $2p_z$ levels will be affected by the field.

Degenerate perturbation theory: If the field were absent we would have one eigenvalue E_1 and four degenerate eigenvalues E_2 . How do these eigenvalues change

as we turn up the field ? As we have mentioned before the eigenvalues are more or less equal to the diagonal values unless the off-diagonal term H_{mn} becomes comparable to the difference between the corresponding diagonal terms ($H_{mm} - H_{nn}$). This means that whenever we have two degenerate eigenvalues (that is, $H_{mm} - H_{nn} = 0$) even a small off-diagonal element H_{mn} has a significant effect. We thus expect the 2s and 2p_z levels to be significantly affected by the field since they are degenerate to start with. We can get a very good approximation for these eigenvalues simply by looking at a subset of the [H] matrix containing just these levels :

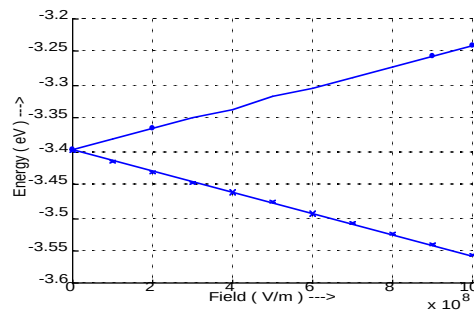
$$[H_0] + [H_F] = \begin{matrix} & \begin{matrix} 2s & 2p_z \end{matrix} \\ \begin{matrix} 2s \\ 2p_z \end{matrix} & \begin{bmatrix} E_2 & B \\ B & E_2 \end{bmatrix} \end{matrix}$$

It is easy to show that the eigenvalues are $E = E_2 \pm B$ and the corresponding eigenvectors are

$$|2s\rangle - |2p_z\rangle \quad \text{and} \quad |2s\rangle + |2p_z\rangle$$

This approximate approach (known as degenerate perturbation theory) describes the exact eigenvalues quite well (see Fig.4.4.1) as long as the off-diagonal elements (like A) coupling these levels to the other levels is much smaller than the energy difference with these levels (like $E_2 - E_1$).

Fig.4.4.1. Energy of 2s - 2p_z level due to an applied electric field, F.
Solid line shows the results obtained by direct diagonalization while 'o' and 'x' show perturbation theory results $E = E_2 \pm B$.



Non-degenerate perturbation theory: How is the 1s eigenvalue affected ? Since there are no other degenerate levels the effect is much less and to first order one could simply ignore the rest of the matrix:

$$[H_0] + [H_F] = \begin{matrix} 1s \\ [E_1] \end{matrix}$$

and argue that the eigenvalue remains E_1 . We could do better by "renormalizing" the matrix as follows. Suppose we partition the $[H]$ matrix and write

$$[H] \begin{Bmatrix} \phi \\ \phi \end{Bmatrix} = E \begin{Bmatrix} \phi \\ \phi \end{Bmatrix} \longrightarrow \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{Bmatrix} \phi_1 \\ \phi_2 \end{Bmatrix} = E \begin{Bmatrix} \phi_1 \\ \phi_2 \end{Bmatrix}$$

where $[H_{11}]$ denotes the part of the matrix we wish to keep (the '1s' block in this case). It is easy to eliminate $\{\phi_2\}$ to obtain

$$[H'] \begin{Bmatrix} \phi_1 \end{Bmatrix} = E \begin{Bmatrix} \phi_1 \end{Bmatrix}$$

where
$$[H'] = [H_{11}] + [H_{12}][E I - H_{22}]^{-1} H_{21}$$

I being an identity matrix of the same size as H_{22} . We haven't gained much if we still have to invert the matrix $E I - H_{22}$ including its off-diagonal elements. But to lowest order we can simply ignore the off-diagonal elements of H_{22} and write down the inverse by inspection. In the present case, this gives us

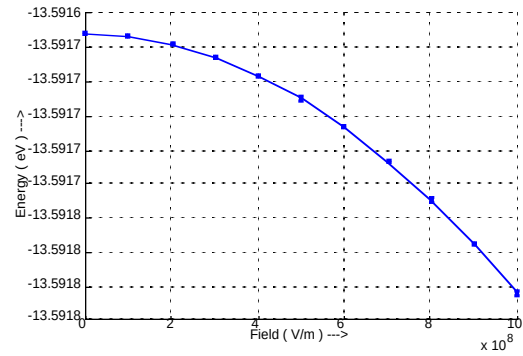
$$[H'] \approx \begin{pmatrix} E_1 & 0 \\ 0 & A \end{pmatrix} \begin{bmatrix} 1/(E - E_2) & 0 \\ 0 & 1/(E - E_2) \end{bmatrix} \begin{pmatrix} 0 \\ A \end{pmatrix} = E_1 + \frac{A^2}{E - E_2}$$

To lowest order, the eigenvalue E is approximately equal to E_1 , so that

$$[H'] \approx E_1 + (A^2 / (E_1 - E_2))$$

showing that the correction to the eigenvalue is quadratic for non-degenerate states, rather than linear as it is for degenerate states. This approximate approach (known as non-degenerate perturbation theory) describes the exact eigenvalues quite well (see Fig.4.4.2).

Fig.4.4.2. Energy of 1s level due to an applied electric field, F . Solid line shows the results obtained by direct diagonalization while the 'x' denotes the perturbation theory results $E = E_1 + A^2 / (E_1 - E_2)$.



Exercises

E.4.1. Plot the electron density $n(x)$ in a hydrogen molecule along the axis joining the two Hydrogen atoms assuming they are separated by the equilibrium bond distance of $R = 0.074$ nm and compare with **Fig.4.1.3**.

E.4.2. Calculate the equilibrium electron density $n(x)$ in a one-dimensional box modeled with discrete lattice of 100 points spaced by $2\text{\AA}$ and compare with each of the results shown in **Fig.4.3.1 (c)-(f)**.

