

CHAPTER 1

Mathematical Review

1.1 Linear Algebra

1.1.1 Three-Dimensional Vector Algebra

Exercise 1.1

a) Show that $O_{ij} = \hat{e}_i \cdot \mathcal{O} \hat{e}_j$. b) If $\mathcal{O} \vec{a} = \vec{b}$ show that $b_i = \sum_j O_{ij} a_j$.

Solution 1.1

1. Using (1.7) and (1.13), we get that

$$\hat{e}_i \cdot \mathcal{O} \hat{e}_j = \hat{e}_i \cdot \sum_{k=1}^3 \hat{e}_k O_{kj} = \sum_{k=1}^3 \hat{e}_i \cdot \hat{e}_k O_{kj} = \sum_{k=1}^3 \delta_{ik} O_{kj} = O_{ij}. \quad (1.1-1)$$

2. Similarly,

$$\vec{b} = \sum_{i=1}^3 b_i \hat{e}_i = \mathcal{O} \vec{a} = \mathcal{O} \sum_{j=1}^3 a_j \hat{e}_j = \sum_{j=1}^3 a_j \mathcal{O} \hat{e}_j = \sum_{j=1}^3 a_j \sum_{i=1}^3 \hat{e}_i O_{ij} = \sum_{i=1}^3 \left(\sum_{j=1}^3 O_{ij} a_j \right) \hat{e}_i.$$

From the uniqueness of linear expression by a basis, we arrive at

$$b_i = \sum_{j=1}^3 O_{ij} a_j. \quad (1.1-2)$$

Exercise 1.2

Calculate $[\mathbf{A}, \mathbf{B}]$ and $\{\mathbf{A}, \mathbf{B}\}$ when

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & -1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Solution 1.2

$$[\mathbf{A}, \mathbf{B}] \equiv \mathbf{AB} - \mathbf{BA} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & -1 & 3 \\ -3 & 0 & -1 \end{pmatrix} - \begin{pmatrix} 0 & 1 & -3 \\ -1 & -1 & 0 \\ 1 & 3 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 4 \\ 2 & 0 & 3 \\ -4 & -3 & 0 \end{pmatrix},$$
$$\{\mathbf{A}, \mathbf{B}\} \equiv \mathbf{AB} + \mathbf{BA} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & -1 & 3 \\ -3 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & -3 \\ -1 & -1 & 0 \\ 1 & 3 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -2 \\ 0 & -2 & 3 \\ -2 & 3 & -2 \end{pmatrix}.$$

1.1.2 Matrices

Exercise 1.3

If \mathbf{A} is an $N \times M$ matrix and \mathbf{B} is a $M \times K$ matrix show that $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$.

Solution 1.3

It is obvious that

$$(\mathbf{B}^\dagger \mathbf{A}^\dagger)_{ij} = \sum_{k=1}^M (\mathbf{B}^\dagger)_{ik} (\mathbf{A}^\dagger)_{kj} = \sum_{k=1}^M B_{ki}^* A_{jk}^* = \left(\sum_{k=1}^M A_{jk} B_{ki} \right)^* = [(\mathbf{AB})^*]_{ji} = [(\mathbf{AB})^\dagger]_{ij}, \quad (1.3-1)$$

which means that $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$.

Exercise 1.4

Show that

- $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$.
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$.
- If \mathbf{U} is unitary and $\mathbf{B} = \mathbf{U}^\dagger \mathbf{A} \mathbf{U}$, then $\mathbf{A} = \mathbf{U} \mathbf{B} \mathbf{U}^\dagger$.
- If the product $\mathbf{C} = \mathbf{AB}$ of two Hermitian matrices is also Hermitian, then \mathbf{A} and \mathbf{B} commute.
- If \mathbf{A} is Hermitian then \mathbf{A}^{-1} , if it exists, is also Hermitian.
- If $\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, then $\mathbf{A}^{-1} = \frac{1}{(A_{11}A_{22} - A_{12}A_{21})} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$.

Solution 1.4

- At this time, we assume that \mathbf{A} is an $N \times M$ matrix while \mathbf{B} is a $M \times N$ matrix. Then,

$$\text{tr}(\mathbf{AB}) = \sum_{i=1}^N (\mathbf{AB})_{ii} = \sum_{i=1}^N \sum_{k=1}^M A_{ik} B_{ki} = \sum_{k=1}^M \sum_{i=1}^N B_{ki} A_{ik} = \sum_{k=1}^M (\mathbf{BA})_{kk} = \text{tr}(\mathbf{BA}). \quad (1.4-1)$$

From this issue, we assume that both \mathbf{A} and \mathbf{B} are $N \times N$ matrices.

- We find that

$$\mathbf{AB}(\mathbf{B}^{-1} \mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{1}.$$

Since the inverse of a matrix is unique, we immediately get that

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}. \quad (1.4-2)$$

- Due to $\mathbf{B} = \mathbf{U}^\dagger \mathbf{A} \mathbf{U}$, we can find

$$\mathbf{A} = \mathbf{1} \mathbf{A} \mathbf{1} = (\mathbf{U} \mathbf{U}^\dagger) \mathbf{A} (\mathbf{U} \mathbf{U}^\dagger) = \mathbf{U} (\mathbf{U}^\dagger \mathbf{A} \mathbf{U}) \mathbf{U}^\dagger = \mathbf{U} \mathbf{B} \mathbf{U}^\dagger. \quad (1.4-3)$$

- Because $\mathbf{C} = \mathbf{AB}$ of two Hermitian matrices is also Hermitian, we know that

$$\mathbf{C}^\dagger = (\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger = \mathbf{C} = \mathbf{AB}.$$

With $\mathbf{A}^\dagger = \mathbf{A}$, $\mathbf{B}^\dagger = \mathbf{B}$, we find

$$\mathbf{B}^\dagger \mathbf{A}^\dagger = \mathbf{BA} = \mathbf{AB}. \quad (1.4-4)$$

In other words, \mathbf{A} and \mathbf{B} commute.

- It is obvious that if \mathbf{A}^{-1} exists,

$$(\mathbf{A}^{-1})^\dagger \mathbf{A}^\dagger = (\mathbf{AA}^{-1})^\dagger = \mathbf{1}^\dagger = \mathbf{1}.$$

We know $(\mathbf{A}^\dagger)^{-1} = (\mathbf{A}^{-1})^\dagger$. Then, with $\mathbf{A} = \mathbf{A}^\dagger$, we find that

$$(\mathbf{A}^{-1})^\dagger = (\mathbf{A}^\dagger)^{-1} = \mathbf{A}^{-1}. \quad (1.4-5)$$

Namely, \mathbf{A}^{-1} is also Hermitian if it exists.

f. If $A_{11}A_{22} - A_{12}A_{21} \neq 0$, we can find

$$\mathbf{A} \times \frac{1}{A_{11}A_{22} - A_{12}A_{21}} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which means that if $A_{11}A_{22} - A_{12}A_{21} \neq 0$,

$$\mathbf{A}^{-1} = \frac{1}{A_{11}A_{22} - A_{12}A_{21}} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}. \quad (1.4-6)$$

1.1.3 Determinants

Exercise 1.5

Verify the above properties for 2×2 determinants.

Solution 1.5

From (1.39), we can verify the above properties for 2×2 determinants.

1. If each element in the first row of $|\mathbf{A}|$ is zero, we will find that

$$\begin{vmatrix} 0 & 0 \\ A_{21} & A_{22} \end{vmatrix} = 0 \times A_{22} - 0 \times A_{21} = 0.$$

In the same way, we can verify that if each element in a row or in a column is zero, the value of the 2×2 determinant $|\mathbf{A}|$ is zero.

2. If $(\mathbf{A})_{ij} = A_{ii}\delta_{ij}$, we will find that

$$\begin{vmatrix} A_{11} & 0 \\ 0 & A_{22} \end{vmatrix} = A_{11}A_{22} - 0 \times 0 = A_{11}A_{22} = \prod_{i=1}^2 A_{ii}.$$

3. We can interchange two rows of the 2×2 determinant $|\mathbf{A}|$ and find that

$$\begin{vmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{vmatrix} = A_{21}A_{12} - A_{11}A_{22} = -(A_{11}A_{22} - A_{12}A_{21}) = -\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}.$$

In the same way, we can verify that a single interchange of any two rows (or columns) of a determinant changes its sign.

4. Note that $(\mathbf{A}^\dagger)_{ij} = A_{ji}^*$,

$$(|\mathbf{A}^\dagger|)^* = \begin{vmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{vmatrix}^* = (A_{11}^*A_{22}^* - A_{12}^*A_{21}^*)^* = A_{11}A_{22} - A_{12}A_{21} = |\mathbf{A}|.$$

5. It is evident that for two 2×2 matrices \mathbf{A} and \mathbf{B} , the determinant of their product is

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

Thus, we can find that

$$\begin{aligned} \det(\mathbf{AB}) &= \begin{vmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{vmatrix} \\ &= (A_{11}B_{11} + A_{12}B_{21})(A_{21}B_{12} + A_{22}B_{22}) - (A_{11}B_{12} + A_{12}B_{22})(A_{21}B_{11} + A_{22}B_{21}) \\ &= A_{11}A_{21}B_{11}B_{12} + A_{11}A_{22}B_{11}B_{22} + A_{12}A_{21}B_{12}B_{21} + A_{12}A_{22}B_{21}B_{22} \\ &\quad - A_{11}A_{21}B_{11}B_{22} - A_{11}A_{22}B_{12}B_{21} - A_{12}A_{21}B_{11}B_{22} - A_{12}A_{22}B_{21}B_{22} \\ &= A_{11}A_{22}(B_{11}B_{22} - B_{12}B_{21}) + A_{12}A_{21}(B_{12}B_{21} - B_{11}B_{22}) \\ &= (A_{11}A_{22} - A_{12}A_{21})(B_{11}B_{22} - B_{12}B_{21}) = |\mathbf{A}||\mathbf{B}|. \end{aligned}$$

Exercise 1.6

Using properties (1)-(5) prove that in general

6. If any two rows (or columns) of a determinant are equal, the value of the determinant is zero.
7. $|\mathbf{A}^{-1}| = (|\mathbf{A}|)^{-1}$.
8. If $\mathbf{A}\mathbf{A}^\dagger = \mathbf{1}$, then $|\mathbf{A}|(|\mathbf{A}|)^* = 1$.
9. If $\mathbf{U}^\dagger\mathbf{O}\mathbf{U} = \mathbf{\Omega}$ and $\mathbf{U}^\dagger\mathbf{U} = \mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$, then $|\mathbf{O}| = |\mathbf{\Omega}|$.

Solution 1.6

6. According to the property 3, a single interchange of any two rows (or columns) of a determinant $|\mathbf{A}|$ generates a minus sign. However, a single interchange of any two equal rows (or columns) does not change $|\mathbf{A}|$, which means $|\mathbf{A}| = -|\mathbf{A}|$ and thus $|\mathbf{A}| = 0$.

7. For a general square matrix \mathbf{A} , it is evident that

$$1 = |\mathbf{1}| = |\mathbf{A}\mathbf{A}^{-1}| = |\mathbf{A}||\mathbf{A}^{-1}|.$$

Therefore,

$$|\mathbf{A}^{-1}| = (|\mathbf{A}|)^{-1}.$$

8. Using the property 4, we find that

$$1 = |\mathbf{1}| = |\mathbf{A}\mathbf{A}^\dagger| = |\mathbf{A}||\mathbf{A}^\dagger| = |\mathbf{A}|(|\mathbf{A}|)^*.$$

9. Note that

$$\mathbf{U}\mathbf{U}^\dagger = \mathbf{1} \Leftrightarrow \mathbf{U}^{-1} = \mathbf{U}^\dagger \Rightarrow |\mathbf{U}^{-1}| = |\mathbf{U}^\dagger| = |\mathbf{U}|^*,$$

and thus $1 = |\mathbf{1}| = |\mathbf{U}^\dagger\mathbf{U}| = |\mathbf{U}^\dagger||\mathbf{U}| = |\mathbf{U}|^*|\mathbf{U}|$. Therefore, we obtain that

$$|\mathbf{\Omega}| = |\mathbf{U}^\dagger\mathbf{O}\mathbf{U}| = |\mathbf{U}^\dagger||\mathbf{O}\mathbf{U}| = |\mathbf{U}|^*|\mathbf{O}||\mathbf{U}| = |\mathbf{O}||\mathbf{U}|^*|\mathbf{U}| = |\mathbf{O}|.$$

Exercise 1.7

Using Eq.(1.39), note that the inverse of a 2×2 matrix \mathbf{A} obtained in Exercise 1.4f can be written as

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$$

and thus \mathbf{A}^{-1} does not exist when $|\mathbf{A}| = 0$. This result holds in general for $N \times N$ matrices. Show that the equation

$$\mathbf{A}\mathbf{c} = \mathbf{0}$$

where \mathbf{A} is an $N \times N$ matrix and \mathbf{c} is a column matrix with elements c_i , $i = 1, 2, \dots, N$ can have a nontrivial solution ($\mathbf{c} \neq \mathbf{0}$) only when $|\mathbf{A}| = 0$.

Solution 1.7

This conclusion is a corollary of Cramer's rule, whose WIKIPEDIA's url is https://en.wikipedia.org/wiki/Cramer's_rule. Its proof is too lengthy to be omitted here.

1.1.4 N -Dimentional Complex Vector Spaces**1.1.5 Change of Basis****Exercise 1.8**

Show that the trace of a matrix is invariant under a unitary transformation, i.e., if $\mathbf{\Omega} = \mathbf{U}^\dagger\mathbf{O}\mathbf{U}$ then show that $\text{tr}(\mathbf{\Omega}) = \text{tr}(\mathbf{O})$.

Solution 1.8

Using the conclusion of Exercise 1.4(a), we find that

$$\text{tr}(\mathbf{\Omega}) = \text{tr}(\mathbf{U}^\dagger \mathbf{O} \mathbf{U}) = \text{tr}(\mathbf{O} \mathbf{U} \mathbf{U}^\dagger) = \text{tr}(\mathbf{O} \mathbf{1}) = \text{tr}(\mathbf{O}). \quad (1.8-1)$$

1.1.6 The Eigenvalue Problem**Exercise 1.9**

Show that Eq.(1.90) contains Eq.(1.87) for all $\alpha = 1, 2, \dots, N$.

Solution 1.9

It is evident that

$$\mathbf{O} \mathbf{U} = \mathbf{O}(\mathbf{c}^1, \mathbf{c}^2, \dots, \mathbf{c}^N) = (\mathbf{O} \mathbf{c}^1, \mathbf{O} \mathbf{c}^2, \dots, \mathbf{O} \mathbf{c}^N) = (\omega_1 \mathbf{c}^1, \omega_2 \mathbf{c}^2, \dots, \omega_N \mathbf{c}^N) = \mathbf{U} \boldsymbol{\omega}, \quad (1.9-1)$$

which equals (1.87) for all $\alpha = 1, 2, \dots, N$.

Exercise 1.10

Since the components of an eigenvector can be found from the eigenvalue equation only to within a multiplicative constant, which is later determined by the normalization, one can set $c_1 = 1$ and $c_2 = c$ in Eq.(1.94). If this is done, Eq.(1.94) becomes

$$\begin{aligned} O_{11} + O_{12}c &= \omega \\ O_{21} + O_{22}c &= \omega c. \end{aligned}$$

After eliminating c , find the two roots of the resulting equation and show that they are the same as those given in Eq.(1.96). This technique, which we shall use numerous times in the book for finding the lowest eigenvalue of a matrix, is basically the secular determinant approach *without* determinants. Thus one can use it to find the lowest eigenvalue of certain $N \times N$ matrices without having to evaluate an $N \times N$ determinant.

Solution 1.10

The proof should be discussed according to the value of $O_{12} = O_{21}$.

- When $O_{12} = O_{21} = 0$, assume that $c \neq 0$ and we find that

$$\omega_1 = O_{11}, \quad \omega_2 = O_{22}.$$

If $c = 0$, we find that

$$\begin{pmatrix} O_{11} & 0 \\ 0 & O_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} Q_{11} \\ 0 \end{pmatrix} = Q_{11} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

At this time, the only $\omega = O_{11}$.

- When $O_{12} = O_{21} \neq 0$, from the first equation, we get that

$$c = \frac{\omega - O_{11}}{O_{12}}.$$

Substitute it into the second equation, we obtain that

$$O_{21} = (\omega - O_{22})c = (\omega - O_{22}) \frac{\omega - O_{11}}{O_{12}},$$

which equals

$$(\omega - O_{11})(\omega - O_{22}) - O_{12}O_{21} = \omega^2 - (O_{11} + O_{22})\omega + (O_{11}O_{22} - O_{12}O_{21}) = 0.$$

The discriminant of the quadratic equation is

$$\Delta_\omega = (O_{11} + O_{22})^2 - 4 \times 1 \times (O_{11}O_{22} - O_{12}O_{21}) = (O_{11} - O_{22})^2 + 4O_{12}O_{21} > 0,$$

and the two roots are

$$\omega_{\pm} = \frac{1}{2} \left[O_{11} + O_{22} \pm \sqrt{(O_{11} - O_{22})^2 + 4O_{12}O_{21}} \right].$$

Note that

$$\begin{aligned} \lim_{O_{12} \rightarrow 0} \omega_{\pm} &= \lim_{O_{12} \rightarrow 0} \frac{1}{2} \left[O_{11} + O_{22} \pm \sqrt{(O_{11} - O_{22})^2 + 4O_{12}O_{21}} \right] \\ &= \frac{1}{2} (O_{11} + O_{22}) \pm \frac{1}{2} \lim_{O_{12} \rightarrow 0} \sqrt{(O_{11} - O_{22})^2 + 4O_{12}O_{21}} \\ &= \frac{1}{2} (O_{11} + O_{22}) \pm \frac{1}{2} |O_{11} - O_{22}|. \end{aligned}$$

When $O_{11} \geq O_{22}$,

$$\omega_+ = \frac{1}{2} (O_{11} + O_{22}) + \frac{1}{2} (O_{11} - O_{22}) = O_{11}, \quad \omega_- = \frac{1}{2} (O_{11} + O_{22}) - \frac{1}{2} (O_{11} - O_{22}) = O_{22}.$$

while $O_{11} < O_{22}$,

$$\omega_+ = \frac{1}{2} (O_{11} + O_{22}) + \frac{1}{2} (O_{22} - O_{11}) = O_{22}, \quad \omega_- = \frac{1}{2} (O_{11} + O_{22}) - \frac{1}{2} (O_{22} - O_{11}) = O_{11}.$$

In conclusion, we obtain that

$$\lim_{O_{12} \rightarrow 0} \omega_1 = O_{11}, \quad \lim_{O_{12} \rightarrow 0} \omega_2 = O_{22}.$$

Thus, the special case of $O_{12} = O_{21} = 0$ can be merged in the general case. And we conclude that two roots obtained by eliminating c are the same as those given in Eq.(1.96).

Exercise 1.11

Consider the matrices

$$\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}.$$

Find numerical values for the eigenvalues and corresponding eigenvectors of these matrices by a) the secular determinant approach; b) the unitary transformation approach. You will see that approach (b) is much easier.

Solution 1.11

Using the secular determinant approach, the intermediate results are as follows.

- For the matrix \mathbf{A} . Its eigenpolynomial is

$$|\mathbf{A} - \omega \mathbf{1}| = \begin{vmatrix} 3 - \omega & 1 \\ 1 & 3 - \omega \end{vmatrix} = (3 - \omega)^2 - 1 = (4 - \omega)(2 - \omega) = 0.$$

Its two roots are

$$\omega_1 = 2, \quad \omega_2 = 4.$$

For the eigenvalue ω_1 ,

$$\mathbf{A} - \omega_1 \mathbf{1} = \begin{pmatrix} 3 - 2 & 1 \\ 1 & 3 - 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

The eigenvector is $(1, -1)^T$.

For the eigenvalue ω_2 ,

$$\mathbf{A} - \omega_2 \mathbf{1} = \begin{pmatrix} 3 - 4 & 1 \\ 1 & 3 - 4 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}.$$

The eigenvector is $(1, 1)^T$.

- For the matrix **B**. Its eigenpolynomial is

$$|\mathbf{B} - \omega \mathbf{1}| = \begin{vmatrix} 3 - \omega & 1 \\ 1 & 2 - \omega \end{vmatrix} = (3 - \omega)(2 - \omega) - 1 = \omega^2 - 5\omega + 5 = 0.$$

Its two roots are

$$\omega_1 = \frac{5 + \sqrt{5}}{2}, \quad \omega_2 = \frac{5 - \sqrt{5}}{2}.$$

For the eigenvalue ω_1 ,

$$\mathbf{B} - \omega_1 \mathbf{1} = \begin{pmatrix} 3 - \frac{5 + \sqrt{5}}{2} & 1 \\ 1 & 2 - \frac{5 + \sqrt{5}}{2} \end{pmatrix} = \begin{pmatrix} \frac{1 - \sqrt{5}}{2} & 1 \\ 1 & \frac{-1 - \sqrt{5}}{2} \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -\frac{1 + \sqrt{5}}{2} \\ 0 & 0 \end{pmatrix}.$$

The eigenvector is $(\frac{1 + \sqrt{5}}{2}, 1)^T$.

For the eigenvalue ω_2 ,

$$\mathbf{B} - \omega_2 \mathbf{1} = \begin{pmatrix} 3 - \frac{5 - \sqrt{5}}{2} & 1 \\ 1 & 2 - \frac{5 - \sqrt{5}}{2} \end{pmatrix} = \begin{pmatrix} \frac{1 + \sqrt{5}}{2} & 1 \\ 1 & \frac{-1 + \sqrt{5}}{2} \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -\frac{1 + \sqrt{5}}{2} \\ 0 & 0 \end{pmatrix}.$$

The eigenvector is $(\frac{1 - \sqrt{5}}{2}, 1)^T$.

Using the unitary transformation approach, the intermediate results are as follows.

- For the matrix **A**. Its θ_0 is

$$\theta_0 = \frac{1}{2} \arctan \frac{2O_{12}}{O_{11} - O_{22}} = \frac{1}{2} \arctan \frac{2 \times 1}{3 - 3} = \frac{1}{2} \arctan \infty = \frac{1}{2} \times \frac{\pi}{2} = \frac{\pi}{4}.$$

Here note that ∞ means that both $+\infty$ and $-\infty$ are allowed but only $+\infty$ is used for the following calculation. The same results can be obtained by using $-\infty$ and this process is omitted for the sake of simplicity!

From (1.106a,b) and (1.107a,b), we obtain that the first eigenvalue ω_1 is

$$\omega_1 = O_{11} \cos^2 \theta_0 + O_{22} \sin^2 \theta_0 + O_{12} \sin 2\theta = 3 \cos^2 \frac{\pi}{4} + 3 \sin^2 \frac{\pi}{4} + 1 \sin \left(\frac{2\pi}{4} \right) = 4,$$

whose eigenvector is

$$\begin{pmatrix} c_1^1 \\ c_2^1 \end{pmatrix} = \begin{pmatrix} \cos \frac{\pi}{4} \\ \sin \frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix},$$

and the second eigenvalue ω_2 is

$$\omega_2 = O_{11} \sin^2 \theta_0 + O_{22} \cos^2 \theta_0 - O_{12} \sin 2\theta = 3 \sin^2 \frac{\pi}{4} + 3 \cos^2 \frac{\pi}{4} - 1 \sin \left(\frac{2\pi}{4} \right) = 2,$$

whose eigenvector is

$$\begin{pmatrix} c_1^2 \\ c_2^2 \end{pmatrix} = \begin{pmatrix} \sin \frac{\pi}{4} \\ -\cos \frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

- For the matrix **B**. Its θ_0 is

$$\theta_0 = \frac{1}{2} \arctan \frac{2O_{12}}{O_{11} - O_{22}} = \frac{1}{2} \arctan \frac{2 \times 1}{3 - 2} = \frac{1}{2} \arctan 2.$$

In other words,

$$2 = \tan 2\theta_0 = \frac{2 \tan \theta_0}{1 - \tan^2 \theta_0} \Rightarrow \tan(\theta_0)_{\pm} = \frac{-1 \pm \sqrt{5}}{2}.$$

We use $\tan \theta_0 = \frac{-1 + \sqrt{5}}{2}$ in the following calculation. Of course, with $\tan \theta_0 = \frac{-1 - \sqrt{5}}{2}$, the same results can be omitted and the derivation is omitted, too.

Thus, we know that from

$$\begin{aligned}\sin^2 \theta_0 + \cos^2 \theta_0 &= 1, \\ \frac{\sin \theta_0}{\cos \theta_0} &= \tan \theta_0 = \frac{-1 + \sqrt{5}}{2},\end{aligned}$$

we get

$$\begin{aligned}\cos^2 \theta_0 &= \frac{1}{1 + \left(\frac{-1+\sqrt{5}}{2}\right)^2} = \frac{4}{4 + 6 - 2\sqrt{5}} = \frac{4}{2\sqrt{5}(\sqrt{5} - 1)} = \frac{\sqrt{5} + 1}{2\sqrt{5}}, \\ \sin^2 \theta_0 &= 1 - \cos^2 \theta_0 = 1 - \frac{\sqrt{5} + 1}{2\sqrt{5}} = \frac{-1 + \sqrt{5}}{2\sqrt{5}}, \\ \sin 2\theta_0 &= 2 \sin \theta_0 \cos \theta_0 = 2 \frac{-1 + \sqrt{5}}{2} \cos \theta_0 \cos \theta_0 = (-1 + \sqrt{5}) \cos^2 \theta_0 = \frac{2}{\sqrt{5}}.\end{aligned}$$

From (1.106a,b) and (1.107a,b), we obtain that the first eigenvalue ω_1 is

$$\omega_1 = O_{11} \cos^2 \theta_0 + O_{22} \sin^2 \theta_0 + O_{12} \sin 2\theta = 3 \frac{\sqrt{5} + 1}{2\sqrt{5}} + 2 \frac{-1 + \sqrt{5}}{2\sqrt{5}} + 1 \frac{2}{\sqrt{5}} = \frac{5 + \sqrt{5}}{2}.$$

whose eigenvector is

$$\begin{pmatrix} c_1^1 \\ c_2^1 \end{pmatrix} = \begin{pmatrix} \cos \theta_0 \\ \sin \theta_0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} \cos^2 \theta_0 \\ \sin \theta \cos \theta_0 \end{pmatrix} = \begin{pmatrix} \cos^2 \theta_0 \\ \frac{1}{2} \sin 2\theta_0 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{5}+1}{2\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix} \Leftrightarrow \begin{pmatrix} \frac{\sqrt{5}+1}{2} \\ 1 \end{pmatrix}.$$

and the second eigenvalue ω_2 is

$$\omega_2 = O_{11} \sin^2 \theta_0 + O_{22} \cos^2 \theta_0 - O_{12} \sin 2\theta = 3 \frac{-1 + \sqrt{5}}{2\sqrt{5}} + 2 \frac{\sqrt{5} + 1}{2\sqrt{5}} - 1 \frac{2}{\sqrt{5}} = \frac{5 - \sqrt{5}}{2}.$$

whose eigenvector is

$$\begin{pmatrix} c_1^2 \\ c_2^2 \end{pmatrix} = \begin{pmatrix} \sin \theta_0 \\ -\cos \theta_0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \sin 2\theta_0 \\ -\cos^2 \theta_0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 \\ -\frac{\sqrt{5}-1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix}.$$

I think that the first approach is much easier.

1.1.7 Functions of Matrices

Exercise 1.12

Given that

$$\mathbf{U}^\dagger \mathbf{A} \mathbf{U} = \mathbf{a} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_N \end{pmatrix} \quad \text{or} \quad \mathbf{A} \mathbf{c}^\alpha = a_\alpha \mathbf{c}^\alpha, \quad \alpha = 1, 2, \dots, N.$$

Show that

a. $\det(\mathbf{A}^n) = a_1^n a_2^n \cdots a_N^n.$

b. $\text{tr}(\mathbf{A}^n) = \sum_{\alpha=1}^N a_\alpha^n.$

c. If $\mathbf{G}(\omega) = (\omega \mathbf{1} - \mathbf{A})^{-1}$, then

$$(\mathbf{G}(\omega))_{ij} = \sum_{\alpha=1}^N \frac{U_{i\alpha} U_{j\alpha}^*}{\omega - a_\alpha} = \sum_{\alpha=1}^N \frac{c_i^\alpha c_j^{\alpha*}}{\omega - a_\alpha}.$$

Show that using Dirac notation this can be rewritten as

$$(\mathbf{G}(\omega))_{ij} \equiv \langle i | \mathcal{G}(\omega) | j \rangle = \sum_{\alpha} \frac{\langle i | \alpha \rangle \langle \alpha | j \rangle}{\omega - a_{\alpha}}.$$

Solution 1.12

Before the formal derivation, note that

$$\mathbf{U}^{\dagger} \mathbf{A} \mathbf{U} = \mathbf{a} \Leftrightarrow \mathbf{A} = \mathbf{U} \mathbf{a} \mathbf{U}^{\dagger} \Rightarrow \mathbf{A}^n = (\mathbf{U} \mathbf{a} \mathbf{U}^{\dagger})^n = \mathbf{U} \mathbf{a}^n \mathbf{U}^{\dagger}.$$

a. Using $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}) = \det(\mathbf{B})\det(\mathbf{A})$ and $\det(\mathbf{1}) = 1$, we find that

$$\begin{aligned} \det(\mathbf{A}^n) &= \det(\mathbf{U} \mathbf{a}^n \mathbf{U}^{\dagger}) = \det(\mathbf{U})\det(\mathbf{a}^n \mathbf{U}^{\dagger}) = \det(\mathbf{a}^n \mathbf{U}^{\dagger})\det(\mathbf{U}) \\ &= \det(\mathbf{a}^n)\det(\mathbf{U}^{\dagger})\det(\mathbf{U}) = \det(\mathbf{a}^n)\det(\mathbf{U}^{\dagger} \mathbf{U}) = \det(\mathbf{a}^n)\det(\mathbf{1}) = \det(\mathbf{a}^n). \end{aligned}$$

b. Using $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$, we find that

$$\text{tr}(\mathbf{A}^n) = \text{tr}(\mathbf{U} \mathbf{a}^n \mathbf{U}^{\dagger}) = \text{tr}(\mathbf{a}^n \mathbf{U}^{\dagger} \mathbf{U}) = \text{tr}(\mathbf{a}^n).$$

c. From $\mathbf{A} = \mathbf{U} \mathbf{a} \mathbf{U}^{\dagger}$, we get that

$$\omega \mathbf{1} - \mathbf{A} = \omega \mathbf{U} \mathbf{1} \mathbf{U}^{\dagger} - \mathbf{U} \mathbf{a} \mathbf{U}^{\dagger} = \mathbf{U}(\omega \mathbf{1} - \mathbf{a}) \mathbf{U}^{\dagger}.$$

Thus,

$$\mathbf{G}(\omega) = (\omega \mathbf{1} - \mathbf{A})^{-1} = [\mathbf{U}(\omega \mathbf{1} - \mathbf{a}) \mathbf{U}^{\dagger}]^{-1} = (\mathbf{U}^{\dagger})^{-1}(\omega \mathbf{1} - \mathbf{a})^{-1} \mathbf{U}^{-1} = \mathbf{U}(\omega \mathbf{1} - \mathbf{a})^{-1} \mathbf{U}^{\dagger},$$

and then

$$(\mathbf{G}(\omega))_{ij} = [\mathbf{U}(\omega \mathbf{1} - \mathbf{a})^{-1} \mathbf{U}^{\dagger}]_{ij} = \sum_{\alpha=1}^N \sum_{\beta=1}^N (\mathbf{U})_{i\alpha} \frac{\delta_{\alpha\beta}}{\omega - a_{\alpha}} (\mathbf{U}^{\dagger})_{\beta j} = \sum_{\alpha=1}^N \frac{U_{i\alpha} U_{j\alpha}^*}{\omega - a_{\alpha}}. \quad (1.12-1)$$

Note that in the basis $\{|\alpha\rangle\}$, \mathbf{U} is diagonal, viz., $\mathbf{U} = \sum_{\beta} |\beta\rangle \langle \beta|$ and thus $\mathbf{U}|\alpha\rangle = |\alpha\rangle$. Using Dirac notation, the final result equals

$$\begin{aligned} (\mathbf{G}(\omega))_{ij} &= \langle i | \mathbf{U}(\omega \mathbf{1} - \mathbf{a})^{-1} \mathbf{U}^{\dagger} | j \rangle = \sum_{\alpha} \sum_{\beta} \langle i | \mathbf{U} | \alpha \rangle \langle \alpha | \frac{1}{\omega \mathbf{1} - \mathbf{a}} | \beta \rangle \langle \beta | \mathbf{U}^{\dagger} | j \rangle \\ &= \sum_{\alpha} \sum_{\beta} \langle i | \mathbf{U} | \alpha \rangle \frac{\delta_{\alpha\beta}}{\omega - a_{\alpha}} \langle \beta | \mathbf{U}^{\dagger} | j \rangle = \sum_{\alpha} \frac{\langle i | \mathbf{U} | \alpha \rangle \langle \alpha | \mathbf{U}^{\dagger} | j \rangle}{\omega - a_{\alpha}} = \sum_{\alpha} \frac{\langle i | \alpha \rangle \langle \alpha | j \rangle}{\omega - a_{\alpha}}. \end{aligned}$$

Exercise 1.13

If

$$\mathbf{A} = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

show that

$$f(\mathbf{A}) = \begin{pmatrix} \frac{1}{2}[f(a+b) + f(a-b)] & \frac{1}{2}[f(a+b) - f(a-b)] \\ \frac{1}{2}[f(a+b) - f(a-b)] & \frac{1}{2}[f(a+b) + f(a-b)] \end{pmatrix}.$$

Solution 1.13

The eigenpolynomial of \mathbf{A} is

$$\det(\mathbf{A} - \varepsilon \mathbf{1}) = \begin{vmatrix} a - \varepsilon & b \\ b & a - \varepsilon \end{vmatrix} = (a - \varepsilon)^2 - b^2 = (a + b - \varepsilon)(a - b - \varepsilon) = 0.$$

It has two roots,

$$\varepsilon_1 = a - b, \quad \varepsilon_2 = a + b.$$

The eigenvector of ε_1 is

$$\mathbf{A} - \varepsilon_1 \mathbf{1} = \begin{pmatrix} a - (a - b) & b \\ b & a - (a - b) \end{pmatrix} = \begin{pmatrix} b & b \\ 0 & b \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

The normalized eigenvector is $\frac{1}{\sqrt{2}}(1, -1)^T$.

Besides, the eigenvector of ε_2 is

$$\mathbf{A} - \varepsilon_2 \mathbf{1} = \begin{pmatrix} a - (a + b) & b \\ b & a - (a + b) \end{pmatrix} = \begin{pmatrix} -b & b \\ b & -b \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}.$$

The normalized eigenvector is $\frac{1}{\sqrt{2}}(1, 1)^T$.

Thus, we know that

$$\mathbf{A} = \mathbf{U} \mathbf{a} \mathbf{U}^\dagger = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} a + b & 0 \\ 0 & a - b \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

and thus

$$\begin{aligned} f(\mathbf{A}) &= \mathbf{U} f(\mathbf{a}) \mathbf{U}^\dagger = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} f(a + b) & 0 \\ 0 & f(a - b) \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}} f(a + b) & \frac{1}{\sqrt{2}} f(a - b) \\ \frac{1}{\sqrt{2}} f(a + b) & -\frac{1}{\sqrt{2}} f(a - b) \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} [f(a + b) + f(a - b)] & \frac{1}{2} [f(a + b) - f(a - b)] \\ \frac{1}{2} [f(a + b) - f(a - b)] & \frac{1}{2} [f(a + b) + f(a - b)] \end{pmatrix}. \end{aligned}$$

1.2 Orthogonal Functions, Eigenfunctions, and Operators

Exercise 1.14

Using the above representation of $\delta(x)$, show that

$$a(0) = \int_{-\infty}^{\infty} dx a(x) \delta(x).$$

Solution 1.14

Using first mean value theorem for definite integrals, we find that

$$\begin{aligned} \int_{-\infty}^{\infty} dx a(x) \delta(x) &= \int_{-\infty}^{\infty} dx a(x) \lim_{\varepsilon \rightarrow 0} \delta_\varepsilon(x) = \lim_{\varepsilon \rightarrow 0} \int_{-\varepsilon}^{\varepsilon} dx a(x) \delta_\varepsilon(x) \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} dx a(x) = \lim_{\substack{\varepsilon \rightarrow 0 \\ \zeta \in (-\varepsilon, \varepsilon)}} \frac{1}{2\varepsilon} a(\zeta) [\varepsilon - (-\varepsilon)] = a(0). \end{aligned}$$

More detailed description about first mean value theorem for definite integrals can be seen in https://en.wikipedia.org/wiki/Mean_value_theorem.

Exercise 1.15

As a further illustration of the consistency of our notation, consider the matrix representation of an operator \mathcal{O} in the basis $\{\psi_i(x)\}$. Starting with

$$\mathcal{O} \psi_i(x) = \sum_j \psi_j(x) O_{ji}. \quad (1)$$

Show that

$$O_{ji} = \int dx \psi_j^*(x) \mathcal{O} \psi_i(x).$$

Then using Eqs.(1.127a) and (1.138) rewrite (1) in the bra-ket notation and show that it is identical to Eq.(1.55).

Solution 1.15

From (1), we find that

$$\int dx \psi_j^*(x) \mathcal{O} \psi_i(x) = \int dx \psi_j^*(x) \sum_k \psi_k(x) O_{ki} = \sum_k \int dx \psi_j^*(x) \psi_k(x) O_{ki} = \sum_k \delta_{jk} O_{ki} = O_{ji}.$$

Therefore, using (1.127a) and (1.138), (1.55) equals

$$\mathcal{O}|i\rangle = \sum_j |j\rangle O_{ji}.$$

Exercise 1.16

Consider the eigenvalue problem

$$\mathcal{O}\phi(x) = \omega\phi(x). \quad (1.1)$$

By expanding ϕ in the complete set $\{\psi_i(x), i = 1, 2, \dots\}$ as

$$\phi(x) = \sum_{i=1}^{\infty} c_i \psi_i(x)$$

show that it becomes equivalent to the matrix eigenvalue problem

$$\mathbf{O}\mathbf{c} = \omega\mathbf{c}$$

where $(\mathbf{c})_i = c_i$ and $(\mathbf{O})_{ij} = \int dx \psi_i^*(x) \mathcal{O} \psi_j(x)$. Do this with and without using bra-ket notation. Note that \mathbf{O} is an infinite matrix. In practice, we cannot handle infinite matrices. To keep things manageable, one uses only a finite subset of the set $\{\psi_i(x)\}$, i.e., $\{\psi_i(x), i = 1, 2, \dots, N\}$. If the above analysis is repeated in this subspace, we obtain an $N \times N$ eigenvalue problem. As we shall see in Section 1.3, the corresponding N eigenvalues approximate the true eigenvalues. In particular, we shall prove that the lowest eigenvalue of the truncated eigenvalue problem is greater or equal to the exact lowest eigenvalue.

Solution 1.16

Without bra-ket notation, we start from

$$\mathcal{O}\phi(x) = \mathcal{O} \sum_j c_j \psi_j(x) = \sum_j \mathcal{O} \psi_j(x) c_j = \omega \sum_j c_j \psi_j(x).$$

Multiplying this by $\psi_i^*(x)$ and integrate over all x , we find that

$$\sum_j \int dx \psi_i^*(x) \mathcal{O} \psi_j(x) c_j = \sum_j O_{ij} c_j = \omega \sum_j c_j \int dx \psi_i^*(x) \psi_j(x) = \omega \sum_j c_j \delta_{ij} = \omega c_i,$$

which means $\mathbf{O}\mathbf{c} = \omega\mathbf{c}$.

With bra-ket notation, we start from

$$\sum_j \mathcal{O}|j\rangle c_j = \omega \sum_j |j\rangle c_j$$

Multiplying this by $\langle i|$ and we find that

$$\sum_j \langle i|\mathcal{O}|j\rangle c_j = \sum_j O_{ij} c_j = \omega \sum_j \langle i|j\rangle c_j = \omega \sum_j \delta_{ij} c_j = \omega c_i.$$

which means $\mathbf{O}\mathbf{c} = \omega\mathbf{c}$, too.

Exercise 1.17

In this subsection we have used a watered-down version of Dirac notation that is sufficient for our purposes but that oversimplifies the deep relationship between vectors and functions. The purpose of this exercise is to provide a glimpse at Dirac notation in its full glory. Consider a denumerably infinite set of complete orthonormal basis vectors, i.e.,

$$\sum_{i=1}^{\infty} |i\rangle\langle i| = \mathbf{1}, \quad (1a)$$

$$\langle i|j\rangle = \delta_{ij}. \quad (1b)$$

Let us introduce a continuously infinite complete set of basis vectors denoted by $|x\rangle$. The analogue of (1a) is

$$\int dx |x\rangle\langle x| = \mathbf{1}, \quad (2a)$$

that is, we have replaced the summation in (1a) by an integral. If we multiply (2a) on the left by $\langle a|$ and on the right by $|b\rangle$ we have

$$\int dx \langle a|x\rangle\langle x|b\rangle = \langle a|b\rangle.$$

Comparing this to Eq.(1.128) suggests that we identify $a^*(x)$ with $\langle a|x\rangle$ and $b(x)$ with $\langle x|b\rangle$. Recall that $\langle i|a\rangle$ is the component of $|a\rangle$ along the basis vector $|i\rangle$. Thus we can regard a function $b(x)$ as the x component of the abstract vector $|b\rangle$ in a coordinate system with a continuously infinite number of axes.

- a. Multiply (2a) on the left by $\langle i|$ and on the right by $|j\rangle$. Using (1b) show that the resulting equation is identical to (1.116) if

$$\psi_i^*(x) = \langle i|x\rangle, \quad \psi_j(x) = \langle x|j\rangle.$$

- b. Multiply (1a) by $\langle x|$ on the left and $|x'\rangle$ on the right. Show that the resulting equation is identical to (1.120) if

$$\langle x|x'\rangle = \delta(x - x'). \quad (2b)$$

This is just the continuous analogue of (1b).

- c. Multiply (2a) by $\langle x'|$ on the left and $|a\rangle$ on the right. Show that the resulting expression is identical to (1.121).
- d. Consider an abstract operator \mathcal{O} . Its matrix elements in the continuous basis $|x\rangle$ are

$$\langle x|\mathcal{O}|x'\rangle = O(x, x').$$

Starting with the relation $\mathcal{O}|a\rangle = |b\rangle$ and inserting unity, we have

$$\mathcal{O}|a\rangle = \mathcal{O}\mathbf{1}|a\rangle = \int dx \mathcal{O}|x\rangle\langle x|a\rangle = |b\rangle.$$

Multiply this equation by $\langle x'|$ and show that the result is identical to Eq.(1.133).

- e. If $O_{ij} = \langle i|\mathcal{O}|j\rangle$, show that

$$O(x, x') = \sum_{ij} \psi_i(x) O_{ij} \psi_j^*(x').$$

Solution 1.17

- a. Multiply (2a) on the left by $\langle i|$ and on the right by $|j\rangle$. Using (1b) show that the resulting equation is identical to

$$\int dx \psi_i^*(x) \psi_j(x) = \int dx \langle i|x\rangle\langle x|j\rangle = \langle i|j\rangle = \delta_{ij},$$

if $\psi_i^*(x) = \langle i|x\rangle$ and $\psi_j(x) = \langle x|j\rangle$.

- b. Multiply (1a) by $\langle x|$ on the left and $|x'\rangle$ on the right. We find that

$$\langle x|x'\rangle = \sum_{i=1}^{\infty} \langle x|i\rangle\langle i|x'\rangle = \sum_{i=1}^{\infty} \psi_i(x) \psi_i^*(x') = \delta(x - x').$$

This is just the continuous analogue of (1b).

c. Multiply (2a) by $\langle x'|$ on the left and $|a\rangle$ on the right. We find that

$$\langle x'|a\rangle = \int dx \langle x'|x\rangle \langle x|a\rangle = \int dx \delta(x - x') a(x) = a(x').$$

It is identical to (1.121).

d. From

$$\mathcal{O}|a\rangle = \mathcal{O}\mathbf{1}|a\rangle = \int dx \mathcal{O}|x\rangle \langle x|a\rangle = |b\rangle,$$

multiply this equation by $\langle x'|$ and we find that

$$b(x') = \langle x'|b\rangle = \int dx \langle x'| \mathcal{O}|x\rangle \langle x|a\rangle = \int dx \mathcal{O}(x', x) a(x),$$

which is identical to Eq.(1.133).

e. It is evident that

$$O(x, x') = \langle x|\mathcal{O}|x'\rangle = \sum_{ij} \langle x|i\rangle \langle i|\mathcal{O}|j\rangle \langle j|x'\rangle = \sum_{ij} \psi_i(x) O_{ij} \psi_j^*(x').$$

1.3 The Variation Method

1.3.1 The Variation Principle

Exercise 1.18

The Schrödinger equation (in atomic units) of an electron moving in one dimension under the influence of the potential $-\delta(x)$ is

$$\left(-\frac{1}{2} \frac{d^2}{dx^2} - \delta(x)\right) |\Phi\rangle = \mathcal{E} |\Phi\rangle$$

Use the variation method with the trial function

$$|\tilde{\Phi}\rangle = N e^{-\alpha x^2}$$

to show that $-\pi^{-1}$ is an upper bound to the exact ground state energy (which is -0.5). You will need the integral

$$\int_{-\infty}^{\infty} dx x^{2m} e^{-\alpha x^2} = \frac{(2m)! \pi^{1/2}}{2^{2m} m! \alpha^{m+1/2}}.$$

Solution 1.18

The norm of the trial function is

$$\langle \tilde{\Phi} | \tilde{\Phi} \rangle = \int dx \langle \tilde{\Phi} | x \rangle \langle x | \tilde{\Phi} \rangle = N^2 \int_{-\infty}^{\infty} dx e^{-\alpha x^2} e^{-\alpha x^2} = N^2 \int_{-\infty}^{\infty} dx e^{-2\alpha x^2} = N^2 \sqrt{\frac{\pi}{2\alpha}}.$$

Thus the normalized trial function is

$$|\Phi\rangle = \sqrt[4]{\frac{2\alpha}{\pi}} |\tilde{\Phi}\rangle = \sqrt[4]{\frac{2\alpha}{\pi}} e^{-\alpha x^2},$$

and its energy functional is

$$\begin{aligned} E(\alpha) &= \int dx \int dx' \langle \Phi | x \rangle \left\langle x \left| -\frac{1}{2} \frac{d^2}{dx^2} - \delta(x) \right| x' \right\rangle \langle x' | \Phi \rangle \\ &= \sqrt{\frac{2\alpha}{\pi}} \int_{-\infty}^{\infty} dx e^{-\alpha x^2} \left(-\frac{1}{2} \frac{d^2}{dx^2} - \delta(x) \right) e^{-\alpha x^2} \end{aligned}$$

$$= \sqrt{\frac{2\alpha}{\pi}} \left(-\frac{1}{2} \int_{-\infty}^{\infty} dx e^{-\alpha x^2} \frac{d^2}{dx^2} (e^{-\alpha x^2}) - \int_{-\infty}^{\infty} dx e^{-\alpha x^2} \delta(x) e^{-\alpha x^2} \right).$$

The first integral is

$$\begin{aligned} \int_{-\infty}^{\infty} dx e^{-\alpha x^2} \frac{d^2}{dx^2} (e^{-\alpha x^2}) &= \int_{-\infty}^{\infty} e^{-\alpha x^2} d \left[\frac{d}{dx} (e^{-\alpha x^2}) \right] = \int_{-\infty}^{\infty} e^{-\alpha x^2} d (-2\alpha x e^{-\alpha x^2}) \\ &= -2\alpha \int_{-\infty}^{\infty} e^{-\alpha x^2} d (x e^{-\alpha x^2}) = -2\alpha \left([x e^{-2\alpha x^2}]_{-\infty}^{+\infty} - \int_{-\infty}^{\infty} x e^{-\alpha x^2} d (e^{-\alpha x^2}) \right) \\ &= -2\alpha \left(0 - (-2\alpha) \int_{-\infty}^{\infty} x^2 e^{-2\alpha x^2} dx \right) = -4\alpha^2 \int_{-\infty}^{\infty} dx x^2 e^{-2\alpha x^2} = -\sqrt{\frac{\alpha\pi}{2}}, \end{aligned}$$

and the second integral is

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2} \delta(x) e^{-\alpha x^2} = [e^{-2\alpha x^2}]_{x=0} = 1.$$

Thus,

$$E(\alpha) = \sqrt{\frac{2\alpha}{\pi}} \left[-\frac{1}{2} \times \left(-\sqrt{\frac{\alpha\pi}{2}} \right) - 1 \right] = \frac{\alpha}{2} - \sqrt{\frac{2\alpha}{\pi}}.$$

And

$$\frac{dE(\alpha)}{d\alpha} = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{2}{\pi}} \alpha^{-\frac{1}{2}} = 0 \Rightarrow a_0 = \frac{2}{\pi}, \quad E(a_0) = -\frac{1}{\pi}.$$

We conclude that $-\pi^{-1}$ is an upper bound to the exact ground state energy.

Exercise 1.19

The Schrödinger equation (in atomic units) for the hydrogen atom is

$$\left(-\frac{1}{2} \nabla^2 - \frac{1}{r} \right) |\Phi\rangle = \mathcal{E} |\Phi\rangle.$$

Use the variation method with the trial function

$$|\tilde{\Phi}\rangle = N e^{-\alpha r^2}$$

to show that $-4/3\pi = -0.4244$ is an upper bound to the exact ground state energy (which is -0.5). You will need the relations

$$\begin{aligned} \nabla^2 f(r) &= r^{-2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) f(r), \\ \int_0^\infty dr r^{2m} e^{-\alpha r^2} &= \frac{(2m)! \pi^{1/2}}{2^{2m+1} m! \alpha^{m+1/2}}, \\ \int_0^\infty dr r^{2m+1} e^{-\alpha r^2} &= \frac{m!}{2\alpha^{m+1}}. \end{aligned}$$

Solution 1.19

The norm of the trial function is

$$\langle \tilde{\Phi} | \tilde{\Phi} \rangle = \int d\mathbf{r} \langle \tilde{\Phi} | \mathbf{r} \rangle \langle \mathbf{r} | \tilde{\Phi} \rangle = 4\pi N^2 \int_0^\infty r^2 dr e^{-\alpha r^2} e^{-\alpha r^2} = 4\pi N^2 \int_0^\infty dr r^2 e^{-2\alpha r^2} = N^2 \sqrt{\left(\frac{\pi}{2\alpha} \right)^3}.$$

Thus the normalized trial function is

$$|\Phi\rangle = \sqrt[4]{\left(\frac{2\alpha}{\pi} \right)^3} |\tilde{\Phi}\rangle = \sqrt[4]{\left(\frac{2\alpha}{\pi} \right)^3} e^{-\alpha r^2},$$

and its energy functional is

$$\begin{aligned} E(\alpha) &= \int d\mathbf{r} \int d\mathbf{r}' \langle \Phi | \mathbf{r} \rangle \left\langle \mathbf{r} \left| -\frac{1}{2} \nabla^2 - \frac{1}{r} \right| \mathbf{r}' \right\rangle \langle \mathbf{r}' | \Phi \rangle \\ &= 4\pi N^2 \int_0^\infty dr r^2 e^{-\alpha r^2} \left(-\frac{1}{2} \nabla^2 - \frac{1}{r} \right) e^{-\alpha r^2} \\ &= 4\pi \sqrt{\left(\frac{2\alpha}{\pi}\right)^3} \left[-\frac{1}{2} \int_0^\infty dr r^2 e^{-\alpha r^2} \nabla^2 (e^{-\alpha r^2}) - \int_0^\infty dr r e^{-2\alpha r^2} \right]. \end{aligned}$$

The first integral is

$$\begin{aligned} &\int_0^\infty dr r^2 e^{-\alpha r^2} \nabla^2 (e^{-\alpha r^2}) = \int_0^\infty dr e^{-\alpha r^2} \frac{d}{dr} \left[r^2 \frac{d}{dr} (e^{-\alpha r^2}) \right] \\ &= \int_0^\infty dr e^{-\alpha r^2} \frac{d}{dr} [r^2 (-2\alpha r e^{-\alpha r^2})] = -2\alpha \int_0^\infty dr e^{-\alpha r^2} \frac{d}{dr} (r^3 e^{-\alpha r^2}) = -2\alpha \int_0^\infty e^{-\alpha r^2} d(r^3 e^{-\alpha r^2}) \\ &= -2\alpha \left[r^3 e^{-2\alpha r^2} \right]_0^\infty - \int_0^\infty r^3 e^{-\alpha r^2} (-2\alpha r e^{-\alpha r^2} dr) = -2\alpha \left(0 + 2\alpha \int_0^\infty dr r^4 e^{-2\alpha r^2} \right) \\ &= -4\alpha^2 \int_0^\infty dr r^4 e^{-2\alpha r^2} = -4\alpha^2 \times \frac{3}{8} \sqrt{\frac{\pi}{(2\alpha)^5}} = -\frac{3}{2} \sqrt{\frac{\pi}{32\alpha}}, \end{aligned}$$

and the second integral is

$$\int_0^\infty dr r e^{-2\alpha r^2} = \frac{1}{2(2\alpha)} = \frac{1}{4\alpha}.$$

Thus,

$$E(\alpha) = 4\pi \sqrt{\left(\frac{2\alpha}{\pi}\right)^3} \left[-\frac{1}{2} \times \left(-\frac{3}{2} \sqrt{\frac{\pi}{32\alpha}} \right) - \frac{1}{4\alpha} \right] = \frac{3\alpha}{2} - 2\sqrt{\frac{2\alpha}{\pi}}.$$

And

$$\frac{dE(\alpha)}{d\alpha} = \frac{3}{2} - \sqrt{\frac{2}{\pi}} \alpha^{-\frac{1}{2}} = 0 \Rightarrow a_0 = \frac{8}{9\pi}, \quad E(a_0) = -\frac{4}{3\pi} = -0.4244.$$

We conclude that $-4/3\pi = -0.4244$ is an upper bound to the exact ground state energy.

Exercise 1.20

The variation principle as applied to matrix eigenvalue problems states that if \mathbf{c} is a normalized ($\mathbf{c}^\dagger \mathbf{c} = 1$) column vector, then $\mathbf{c}^\dagger \mathbf{O} \mathbf{c}$ is greater or equal to the lowest eigenvalue of \mathbf{O} . For the 2×2 symmetric matrix ($O_{12} = O_{21}$)

$$\mathbf{O} = \begin{pmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{pmatrix},$$

consider the trial vector

$$\mathbf{c} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$

which is normalized for any value of θ . Calculate

$$\omega(\theta) = \mathbf{c}^\dagger \mathbf{O} \mathbf{c}$$

and find the value of θ (i.e., θ_0) for which $\omega(\theta)$ is a minimum. Show that $\omega(\theta_0)$ is exactly equal to the lowest eigenvalue of \mathbf{O} (see Eqs.(1.105) and (1.106a)). Why should you have anticipated this result?

Solution 1.20

1.3.2 The Linear Variational Problem

Exercise 1.21

Consider a normalized trial function $|\tilde{\Phi}'\rangle$ that is orthogonal to the exact ground state wave function, i.e., $\langle\tilde{\Phi}'|\Phi_0\rangle = 0$.

- a. Generalize the proof of the variation principle of Subsection 1.3.1 to show that

$$\langle\tilde{\Phi}'|\mathcal{H}|\tilde{\Phi}'\rangle \geq \mathcal{E}_1.$$

- b. Consider the function

$$|\tilde{\Phi}'\rangle = x|\tilde{\Phi}_0\rangle + y|\tilde{\Phi}_1\rangle$$

where $|\tilde{\Phi}_\alpha\rangle$, $\alpha = 0, 1$ are given by Eq.(1.168). Show that if it is normalized, then

$$|x|^2 + |y|^2 = 1.$$

- c. When x and y are chosen so that $|\tilde{\Phi}'\rangle$ is normalized and so that $\langle\tilde{\Phi}'|\Phi_0\rangle = 0$, then from part (a) it follows that $\langle\tilde{\Phi}'|\mathcal{H}|\tilde{\Phi}'\rangle \geq \mathcal{E}_1$. Show that

$$\langle\tilde{\Phi}'|\mathcal{H}|\tilde{\Phi}'\rangle = E_1 - |x|^2(E_1 - E_0).$$

Since $E_1 \geq E_0$, conclude that $E_1 \geq \mathcal{E}_1$. The above argument can be generalized to show that $E_\alpha \geq \mathcal{E}_\alpha$, $\alpha = 2, 3, \dots$.

Solution 1.21

- a.

$$\langle\tilde{\Phi}'|\mathcal{H}|\tilde{\Phi}'\rangle = \sum_\alpha \sum_\beta \langle\tilde{\Phi}'|\tilde{\Phi}_\alpha\rangle \langle\tilde{\Phi}_\alpha|\mathcal{H}|\tilde{\Phi}_\beta\rangle \langle\tilde{\Phi}_\beta|\tilde{\Phi}'\rangle = \sum_\alpha \varepsilon_\alpha |\langle\tilde{\Phi}'|\tilde{\Phi}_\alpha\rangle|^2$$

Due to $\varepsilon_0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_{N-1}$ and $\langle\tilde{\Phi}'|\tilde{\Phi}_0\rangle = 0$,

$$\langle\tilde{\Phi}'|\mathcal{H}|\tilde{\Phi}'\rangle = \sum_\alpha \varepsilon_\alpha |\langle\tilde{\Phi}'|\tilde{\Phi}_\alpha\rangle|^2 \geq \sum_\alpha \varepsilon_1 |\langle\tilde{\Phi}'|\tilde{\Phi}_\alpha\rangle|^2 = \varepsilon_1. \quad (1.2)$$

- b.

$$\begin{aligned} 1 &= \langle\tilde{\Phi}'|\tilde{\Phi}'\rangle = (x^* \langle\tilde{\Phi}_0| + y^* \langle\tilde{\Phi}_1|)(x|\tilde{\Phi}_0\rangle + y|\tilde{\Phi}_1\rangle) \\ &= |x|^2 + |y|^2 + x^*y \langle\tilde{\Phi}_0|\tilde{\Phi}_1\rangle + xy^* \langle\tilde{\Phi}_0|\tilde{\Phi}_1\rangle = |x|^2 + |y|^2. \end{aligned} \quad (1.3)$$

- c.

$$\begin{aligned} \langle\tilde{\Phi}'|\mathcal{H}|\tilde{\Phi}'\rangle &= (x^* \langle\tilde{\Phi}_0| + y^* \langle\tilde{\Phi}_1|)\mathcal{H}(x|\tilde{\Phi}_0\rangle + y|\tilde{\Phi}_1\rangle) \\ &= |x|^2 E_0 + |y|^2 E_1 = (|x|^2 + |y|^2) E_1 - |x|^2 E_1 + |x|^2 E_0 \\ \langle\tilde{\Phi}'|\mathcal{H}|\tilde{\Phi}'\rangle &= E_1 - |x|^2(E_1 - E_0) \geq E_0. \end{aligned} \quad (1.4)$$

Exercise 1.22

The Schrödinger equation (in atomic units) for a hydrogen atom in a uniform electric field F in the z direction is

$$\left(-\frac{1}{2}\nabla^2 - \frac{1}{r} + Fr \cos \theta\right) |\Phi\rangle = (\mathcal{H}_0 + Fr \cos \theta) |\Phi\rangle = \mathcal{E}(F) |\Phi\rangle.$$

Use the trial function

$$|\tilde{\Phi}\rangle = c_1|1s\rangle + c_2|2p_z\rangle$$

where $|1s\rangle$ and $|2p_z\rangle$ are the normalized eigenfunctions of \mathcal{H}_0 , i.e.,

$$|1s\rangle = \pi^{-1/2} e^{-r},$$

$$|2p_z\rangle = (32\pi)^{-1/2} r e^{-r/2} \cos \theta$$

to find an upper bound to $\mathcal{E}(F)$. In constructing the matrix representation of \mathcal{H} , you can avoid a lot of work by noting that

$$\mathcal{H}_0|1s\rangle = -\frac{1}{2}|1s\rangle, \quad \mathcal{H}_0|2p_z\rangle = -\frac{1}{8}|2p_z\rangle.$$

Using $(1+x)^{1/2} \approx 1+x/2$, expand your answer in a Taylor series in F , i.e.,

$$E(F) = E(0) - \frac{1}{2}\alpha F^2 + \dots$$

Show that the coefficient α , which is the approximate dipole polarizability of the system, is equal to 2.96. The exact result is 4.5.

Solution 1.22

$$\begin{aligned} \langle 1s|\mathcal{H}_0|1s\rangle &= -\frac{1}{2}, \\ \langle 1s|\mathcal{H}_0|2p_z\rangle &= -\frac{1}{8}\langle 1s|2p_z\rangle = -\frac{1}{8}\int_0^{+\infty} r^3 e^{-\frac{3}{2}r} dr \int_0^\pi \sin\theta \cos\theta d\theta \int_0^{2\pi} d\varphi \cdot \frac{1}{\sqrt{\pi}} \cdot \frac{1}{\sqrt{32\pi}} = 0, \\ \langle 2p_z|\mathcal{H}_0|1s\rangle &= (\langle 1s|\mathcal{H}_0|2p_z\rangle)^* = 0, \\ \langle 2p_z|\mathcal{H}_0|2p_z\rangle &= -\frac{1}{8}\langle 2p_z|2p_z\rangle = -\frac{1}{8}. \end{aligned}$$