# Physics 4617/5617: Quantum Physics
# Course Lecture Notes

Dr. Donald G. Luttermoser
East Tennessee State University

**Edition 5.1**

**Abstract**

These class notes are designed for use of the instructor and students of the course Physics 4617/5617: Quantum Physics. This edition was last modified for the Fall 2006 semester.

# IV. Formulism and Techniques

## A. Linear Algebra.

**1.** In classical mechanics, **vectors** are typically defined in Cartesian coordinates as

$$\boldsymbol{\alpha} = \alpha_x \hat{\boldsymbol{x}} + \alpha_y \hat{\boldsymbol{y}} + \alpha_z \hat{\boldsymbol{z}} \ , \tag{IV-1}$$

with the "hat" unit vector notation or

$$\boldsymbol{\alpha} = \alpha_x i + \alpha_y j + \alpha_z k \ ,$$

in the $ijk$ unit vector notation (I prefer the use of the hat notation).

**a)** Vectors are added via the component method such that

$$\boldsymbol{\alpha} \pm \boldsymbol{\beta} = (\alpha_x \pm \beta_x)\hat{\boldsymbol{x}} + (\alpha_y \pm \beta_y)\hat{\boldsymbol{y}} + (\alpha_z \pm \beta_x)\hat{\boldsymbol{z}} \ . \tag{IV-2}$$

**b)** However in quantum mechanics, often we will have more than 3 coordinates to worry about — indeed, sometimes there may be an infinite amount of coordinates!

**c)** As such, we will introduce a new notation (the so-called **bra**-and-**ket** notation) to describe vectors:

$$\begin{aligned} \boldsymbol{\alpha} &\equiv |\alpha\rangle \qquad \text{(ket)} \\ \boldsymbol{\alpha}^* &\equiv \langle\alpha| \qquad \text{(bra)} \end{aligned} \tag{IV-3}$$

that was first introduced by Paul Dirac.

**d)** The Dirac bra-and-ket notation has the following meanings:

$$\langle\alpha|\beta\rangle = \sum_{n=1}^{N} \alpha_n^* \beta_n \tag{IV-4}$$

if <u>vectors</u> $\alpha$ and $\beta$ represent discrete (*i.e.*, bound) states and

$$\langle\alpha|\beta\rangle = \int_{-\infty}^{\infty} \alpha^* \beta \ d\tau \tag{IV-5}$$

for continuous (*i.e.*, free) states given by <u>functions</u> $\alpha$ and $\beta$, with $d\tau = dx$ in 1-D space and $d\tau = dx\,dy\,dz$ in 3-D Cartesian space.

**2.** A **vector space** consists of a set of **vectors** ($|\alpha\rangle$, $|\beta\rangle$, $|\gamma\rangle$, ...), together with a set of (real or complex) **scalars** ($a$, $b$, $c$, ...), which are subject to 2 operations:

**a)** **Vector addition:** The *sum* of any 2 vectors is another vector:

$$|\alpha\rangle + |\beta\rangle = |\gamma\rangle \; . \tag{IV-6}$$

**i)** Vector addition is **commutative**:

$$|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle \; . \tag{IV-7}$$

**ii)** Vector addition is **associative**:

$$|\alpha\rangle + (|\beta\rangle + |\gamma\rangle) = (|\alpha\rangle + |\beta\rangle) + |\gamma\rangle \; . \tag{IV-8}$$

**iii)** There exists a **zero** (or **null**) **vector**, $|0\rangle$, with the property that

$$|\alpha\rangle + |0\rangle = |\alpha\rangle \; , \tag{IV-9}$$

for every vector $|\alpha\rangle$.

**iv)** For every vector $|\alpha\rangle$ there is an associated **inverse vector** ($|-\alpha\rangle$) such that

$$|\alpha\rangle + |-\alpha\rangle = |0\rangle \; . \tag{IV-10}$$

**b)** **Scalar multiplication:** The *product* of any scalar with any vector is another vector:

$$a|\alpha\rangle = |\gamma\rangle \; . \tag{IV-11}$$

**i)** Scalar multiplication is **distributive** with respect to vector addition:

$$a(|\alpha\rangle + |\beta\rangle) = a|\alpha\rangle + a|\beta\rangle \ , \qquad \text{(IV-12)}$$

and with respect to scalar addition:

$$(a+b)|\alpha\rangle = a|\alpha\rangle + b|\alpha\rangle \ . \qquad \text{(IV-13)}$$

**ii)** It is also **associative**:

$$a(b|\alpha\rangle) = (ab)|\alpha\rangle \ . \qquad \text{(IV-14)}$$

**iii)** Multiplications by the **null** and **unit vector** are

$$0|\alpha\rangle = |0\rangle; \quad 1|\alpha\rangle = |\alpha\rangle \ . \qquad \text{(IV-15)}$$

(Note that $|-\alpha\rangle = (-1)|\alpha\rangle$.)

**c)** A **linear combination** of the vectors $|\alpha\rangle$, $|\beta\rangle$, $|\gamma\rangle$, ... is an expression of the form

$$a|\alpha\rangle + b|\beta\rangle + c|\gamma\rangle + \cdots .$$

**i)** A vector $|\lambda\rangle$ is said to be **linearly independent** of the set $|\alpha\rangle$, $|\beta\rangle$, $|\gamma\rangle$, ... if it cannot be written as a linear combination of them (*e.g.*, unit vectors $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$).

**ii)** A collection of vectors is said to **span** the space if *every* vector can be written as a linear combination of the members of this set.

**iii)** A set of *linearly independent* vectors that spans the space is called a **basis** $\Longrightarrow$ $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}$ define the Cartesian basis.

**iv)** The number of vectors in any basis is called the **dimension** of the space. Here we will introduce the *finite* bases (analogous to unit vectors),

$$|e_1\rangle, |e_2\rangle, ..., |e_n\rangle \ ,$$

of any given vector:

$$|\alpha\rangle = a_1|e_1\rangle + a_2|e_2\rangle + \cdots + a_n|e_n\rangle \ , \qquad \text{(IV-16)}$$

which is uniquely represented by the (ordered) $n$-tuple of its **components**:

$$|\alpha\rangle \leftrightarrow (a_1, a_2, ..., a_n) \ . \qquad \text{(IV-17)}$$

**v)** It is often easier to work with components than with the abstract vectors themselves. Use whatever method to which you are most comfortable.

**3.** In 3 dimensions we encounter 2 kinds of vector products: the *dot product* and the *cross product*. The latter does not generalize in any natural way to $n$-dimensional vector spaces, but the former *does* and is called the **inner product.**

**a)** The inner product of 2 vectors ($|\alpha\rangle$ and $|\beta\rangle$) is a complex number (which we write as $\langle\alpha|\beta\rangle$), with the following properties:

$$\langle\beta|\alpha\rangle = \langle\alpha|\beta\rangle^* \qquad \text{(IV-18)}$$

$$\langle\alpha|\alpha\rangle \geq 0, \ \& \ \langle\alpha|\alpha\rangle = 0 \Leftrightarrow |\alpha\rangle = |0\rangle \qquad \text{(IV-19)}$$

$$\langle\alpha|(b|\beta\rangle + c|\gamma\rangle) = b\langle\alpha|\beta\rangle + c\langle\alpha|\gamma\rangle \ . \qquad \text{(IV-20)}$$

**b)** A vector space with an inner product is called an **inner product space**.

**c)** Because the inner product of any vector with itself is a non-negative number (Eq. IV-19), its square root is *real*

— we call this the **norm** (think of this as the *length*) of the vector:

$$\|\alpha\| \equiv \sqrt{\langle\alpha|\alpha\rangle} \ . \qquad \text{(IV-21)}$$

**d)** A *unit* vector, whose norm is 1, is said to be **normalized**.

**e)** Two vectors whose inner product is zero are called **orthogonal** $\Longrightarrow$ a collection of mutually orthogonal normalized vectors,

$$\langle\alpha_i|\alpha_j\rangle = \delta_{ij} \ , \qquad \text{(IV-22)}$$

is called an **orthonormal set**, where $\delta_{ij}$ is the **Kronecker delta**.

**f)** Components of vectors can be written as

$$a_i = \langle e_i|\alpha\rangle \ . \qquad \text{(IV-23)}$$

**g)** The conditions given in Eqs. (IV-18,19,20) give rise to the **Schwarz inequality** which states that

$$|\langle\alpha|\beta\rangle|^2 \leq \langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \qquad \text{(IV-24)}$$

(see its proof in Example IV-1). Note that the Schwarz inequality holds only if the vectors $|\alpha\rangle$ and $|\beta\rangle$ are colinear (*i.e.*, proportional to each other: $|\alpha\rangle = c|\beta\rangle$).

**h)** We can define the (complex) angle between $|\alpha\rangle$ and $|\beta\rangle$ by the formula

$$\cos\theta = \sqrt{\frac{\langle\alpha|\beta\rangle\langle\beta|\alpha\rangle}{\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle}} \ . \qquad \text{(IV-25)}$$

**4.** A **linear transformation** ($\hat{T}$, the *hat* on an operator from this point forward will imply that the operator is a linear transformation — don't confuse it with the *hat* of a unit vector) takes each vector in a vector space and "transforms" it into some other vector ($|\alpha\rangle \rightarrow |\alpha'\rangle = \hat{T}|\alpha\rangle$), with the proviso that the operator is *linear*

$$\hat{T}(a|\alpha\rangle + b|\beta\rangle) = a(\hat{T}|\alpha\rangle) + b(\hat{T}|\beta\rangle) \ . \tag{IV-26}$$

**a)** We can write the linear transformation of basis vectors as

$$\hat{T}|e_j\rangle = \sum_{i=1}^{n} T_{ij}|e_i\rangle, \quad (j = 1, 2, ..., n) \ , \tag{IV-27}$$

hence the $\hat{T}$ operator is a **tensor**.

**b)** If $|\alpha\rangle$ is an arbitrary vector:

$$|\alpha\rangle = a_1|e_1\rangle + \cdots + a_n|e_n\rangle = \sum_{j=1}^{n} a_j|e_j\rangle \ , \tag{IV-28}$$

then

$$\hat{T}|\alpha\rangle = \sum_{j=1}^{n} a_j(\hat{T}|e_j\rangle) = \sum_{j=1}^{n}\sum_{i=1}^{n} a_j T_{ij}|e_i\rangle = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} T_{ij}a_j\right)|e_i\rangle \ . \tag{IV-29}$$

$\hat{T}$ takes a vector with components $a_1, a_2, ..., a_n$ into a vector with components

$$a_i' = \sum_{j=1}^{n} T_{ij}a_j \ . \tag{IV-30}$$

**c)** If the basis is orthonormal, it follows from Eq. (IV-27) that

$$T_{ij} = \langle e_i|\hat{T}|e_j\rangle \ , \tag{IV-31}$$

or in matrix notation

$$\mathbf{T} = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{21} & T_{22} & \cdots & T_{2n} \\ \vdots & \vdots & & \vdots \\ T_{n1} & T_{n2} & \cdots & T_{nn} \end{pmatrix} \ . \tag{IV-32}$$

**d)** The sum of 2 linear transformations is

$$(\hat{S} + \hat{T})|\alpha\rangle = \hat{S}|\alpha\rangle + \hat{T}|\alpha\rangle \ , \qquad \text{(IV-33)}$$

or, again, in matrix notation,

$$\mathbf{U} = \mathbf{S} + \mathbf{T} \Leftrightarrow U_{ij} = S_{ij} + T_{ij} \ . \qquad \text{(IV-34)}$$

**e)** The *product* of 2 linear transformations $(\hat{S}\hat{T})$ is the net effect of performing them in succession — first $\hat{T}$, the $\hat{S}$. In matrix notation:

$$\mathbf{U} = \mathbf{ST} \ \Leftrightarrow U_{ik} = \sum_{j=1}^{n} S_{ij} T_{jk} \ ; \qquad \text{(IV-35)}$$

this is the standard rule for matrix multiplication — to find the $ik^{\text{th}}$ element of the product, you look at the $i^{\text{th}}$ row of $\mathbf{S}$ and the $k^{\text{th}}$ column of $\mathbf{T}$, multiply corresponding entries, and add.

**f)** The **transpose** of a matrix $(\tilde{\mathbf{T}})$ is the same set of elements in $\mathbf{T}$, but with the rows and columns interchanged:

$$\tilde{\mathbf{T}} = \begin{pmatrix} T_{11} & T_{21} & \cdots & T_{n1} \\ T_{12} & T_{22} & \cdots & T_{n2} \\ \vdots & \vdots & & \vdots \\ T_{1n} & T_{2n} & \cdots & T_{nn} \end{pmatrix} . \qquad \text{(IV-36)}$$

Note that the transpose of a column matrix is a row matrix!

**g)** A square matrix is **symmetric** if it is equal to its transpose (reflection in the main diagonal — upper left to lower right — leaves it unchanged); it is **antisymmetric** if this operation reverses the sign:

$$\text{SYMMETRIC: } \tilde{\mathbf{T}} = \mathbf{T} \ ; \quad \text{ANTISYMMETRIC: } \tilde{\mathbf{T}} = -\mathbf{T} \ . \qquad \text{(IV-37)}$$

**h)** The (complex) **conjugate** ($\mathbf{T}^*$) is obtained by taking the complex conjugate of every element:

$$\mathbf{T}^* = \begin{pmatrix} T_{11}^* & T_{12}^* & \cdots & T_{1n}^* \\ T_{21}^* & T_{22}^* & \cdots & T_{2n}^* \\ \vdots & \vdots & & \vdots \\ T_{n1}^* & T_{n2}^* & \cdots & T_{nn}^* \end{pmatrix} \; ; \; \mathbf{a}^* = \begin{pmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_n^* \end{pmatrix} . \qquad \text{(IV-38)}$$

**i)** A matrix is **real** if all its elements are real and **imaginary** if they are all imaginary:

$$\text{REAL: } \mathbf{T}^* = \mathbf{T} \; ; \quad \text{IMAGINARY: } \mathbf{T}^* = -\mathbf{T} . \qquad \text{(IV-39)}$$

**j)** A square matrix is **Hermitian** (or **self-adjoint** as defined by $\mathbf{T}^\dagger \equiv \tilde{\mathbf{T}}^*$) if it is equal to its Hermitian conjugate; if Hermitian conjugation introduces a minus sign, the matrix is **skew Hermitian** (or **anti-Hermitian**):

$$\text{HERMITIAN: } \mathbf{T}^\dagger = \mathbf{T} \; ; \quad \text{SKEW HERMITIAN: } \mathbf{T}^\dagger = -\mathbf{T} . \qquad \text{(IV-40)}$$

**k)** With this notation, the inner product of 2 vectors (with respect to an orthonormal basis), can be written in matrix form:

$$\langle \alpha | \beta \rangle = \mathbf{a}^\dagger \mathbf{b} . \qquad \text{(IV-41)}$$

**l)** Matrix multiplication is not, in general, commutative ($\mathbf{ST} \neq \mathbf{TS}$) — the difference between 2 orderings is called the **commutator**:

$$[\mathbf{S}, \mathbf{T}] \equiv \mathbf{ST} - \mathbf{TS} . \qquad \text{(IV-42)}$$

It can also be shown that one can write the following commutator relation:

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}\,[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\,\hat{B} . \qquad \text{(IV-43)}$$

**Exercise:** Prove Eq. (IV-43).

**m)** The transpose of a product is the product of the transpose *in reverse order*:

$$\widetilde{(\mathbf{S}\mathbf{T})} = \tilde{\mathbf{T}}\tilde{\mathbf{S}} \; , \qquad \text{(IV-44)}$$

and the same goes for Hermitian conjugates:

$$(\mathbf{S}\mathbf{T})^\dagger = \mathbf{T}^\dagger\mathbf{S}^\dagger . \qquad \text{(IV-45)}$$

**n)** The **unit matrix** is defined by

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} . \qquad \text{(IV-46)}$$

In other words,

$$\mathbf{1}_{ij} = \delta_{ij} \; . \qquad \text{(IV-47)}$$

**o)** The **inverse** of a matrix (written $\mathbf{T}^{-1}$) is defined by

$$\mathbf{T}^{-1}\mathbf{T} = \mathbf{T}\mathbf{T}^{-1} = \mathbf{1} \; . \qquad \text{(IV-48)}$$

**i)** A matrix has an inverse if and only if its **determinant** is nonzero; in fact

$$\mathbf{T}^{-1} = \frac{1}{\det \mathbf{T}} \tilde{\mathbf{C}} = \frac{1}{|\mathbf{T}|} \tilde{\mathbf{C}} \; , \qquad \text{(IV-49)}$$

where $\mathbf{C}$ is the matrix of **cofactors**.

**ii)** The cofactor of element $T_{ij}$ is $(-1)^{i+j}$ times the determinant of the submatrix obtained from $\mathbf{T}$ by erasing the $i^{\text{th}}$ row by the $j^{\text{th}}$ column.

**iii)** As an example for taking the inverse of a matrix, let's assume that $\mathbf{T}$ is a 3x3 matrix of form

$$\mathbf{T} = \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} . \qquad \text{(IV-50)}$$

Its determinant is then

$$
\begin{aligned}
\det \mathbf{T} &= |\mathbf{T}| = \begin{vmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{vmatrix} \\
&= T_{11} \begin{vmatrix} T_{22} & T_{23} \\ T_{32} & T_{33} \end{vmatrix} - T_{12} \begin{vmatrix} T_{21} & T_{23} \\ T_{31} & T_{33} \end{vmatrix} \\
&\qquad + T_{13} \begin{vmatrix} T_{21} & T_{22} \\ T_{31} & T_{32} \end{vmatrix} \\
&= T_{11} \left( T_{22} T_{33} - T_{23} T_{32} \right) - T_{12} \left( T_{21} T_{33} - T_{23} T_{31} \right) \\
&\qquad + T_{13} \left( T_{21} T_{32} - T_{22} T_{31} \right) \ . \qquad \text{(IV-51)}
\end{aligned}
$$

**iv)** For this 3x3 matrix, the matrix of cofactors is given by

$$
\mathbf{C} = \begin{pmatrix}
\begin{vmatrix} T_{22} & T_{23} \\ T_{32} & T_{33} \end{vmatrix} & -\begin{vmatrix} T_{21} & T_{23} \\ T_{31} & T_{33} \end{vmatrix} & \begin{vmatrix} T_{21} & T_{22} \\ T_{31} & T_{32} \end{vmatrix} \\
-\begin{vmatrix} T_{12} & T_{13} \\ T_{32} & T_{33} \end{vmatrix} & \begin{vmatrix} T_{11} & T_{13} \\ T_{31} & T_{33} \end{vmatrix} & -\begin{vmatrix} T_{11} & T_{12} \\ T_{31} & T_{32} \end{vmatrix} \\
\begin{vmatrix} T_{12} & T_{13} \\ T_{22} & T_{23} \end{vmatrix} & -\begin{vmatrix} T_{11} & T_{13} \\ T_{21} & T_{23} \end{vmatrix} & \begin{vmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{vmatrix}
\end{pmatrix} . \qquad \text{(IV-52)}
$$

**v)** The transpose of this cofactor matrix is then (see Eq. IV-36)

$$
\tilde{\mathbf{C}} = \begin{pmatrix}
\begin{vmatrix} T_{22} & T_{32} \\ T_{23} & T_{33} \end{vmatrix} & -\begin{vmatrix} T_{12} & T_{32} \\ T_{13} & T_{33} \end{vmatrix} & \begin{vmatrix} T_{12} & T_{22} \\ T_{13} & T_{23} \end{vmatrix} \\
-\begin{vmatrix} T_{21} & T_{31} \\ T_{23} & T_{33} \end{vmatrix} & \begin{vmatrix} T_{11} & T_{31} \\ T_{13} & T_{33} \end{vmatrix} & -\begin{vmatrix} T_{11} & T_{21} \\ T_{13} & T_{23} \end{vmatrix} \\
\begin{vmatrix} T_{21} & T_{31} \\ T_{22} & T_{32} \end{vmatrix} & -\begin{vmatrix} T_{11} & T_{31} \\ T_{12} & T_{32} \end{vmatrix} & \begin{vmatrix} T_{11} & T_{21} \\ T_{12} & T_{22} \end{vmatrix}
\end{pmatrix} . \qquad \text{(IV-53)}
$$

**vi)** A matrix without an inverse is said to be **singular**.

**vii)** The inverse of a product (assuming it exists) is the product of the inverses *in reverse order*:

$$(\mathbf{ST})^{-1} = \mathbf{T^{-1}S^{-1}}. \qquad \text{(IV-54)}$$

**p)** A matrix is **unitary** if its inverse is equal to its Hermitian conjugate:

$$\text{UNITARY:} \quad \mathbf{U}^\dagger = \mathbf{U}^{-1}. \qquad \text{(IV-55)}$$

**q)** The **trace** of a matrix is the sum of the diagonal elements:

$$\text{Tr}(\mathbf{T}) \equiv \sum_{i=1}^{m} T_{ii}, \qquad \text{(IV-56)}$$

and has the property

$$\text{Tr}(\mathbf{T_1T_2}) = \text{Tr}(\mathbf{T_2T_1}). \qquad \text{(IV-57)}$$

**5.** A vector under a linear transformation that obeys the following equation:

$$\hat{T}|\alpha\rangle = \lambda|\alpha\rangle, \qquad \text{(IV-58)}$$

where $\hat{T}|\alpha\rangle$ is called the **eigenvector** of the transformation, and the (complex) number $\lambda$ is called the **eigenvalue**. Such an equation shows that a linear transformation creates a "scaled" duplicate (by a factor of $\lambda$) of the original vector $|\alpha\rangle$.

**a)** Notice that any (nonzero) multiple of an eigenvector is still an eigenvector with the same eigenvalue.

**b)** In matrix form, the eigenvector equation takes the form:

$$\mathbf{Ta} = \lambda\mathbf{a} \qquad \text{(IV-59)}$$

(for nonzero **a**), or

$$(\mathbf{T} - \lambda\mathbf{1})\mathbf{a} = \mathbf{0}. \qquad \text{(IV-60)}$$

(here **0** is the **zero matrix**, whose elements are all zero.)

**c)** If the matrix (**T** - $\lambda$**1**) had an *inverse*, we could multiply both sides of Eq. (IV-60) by $(\mathbf{T} - \lambda\mathbf{1})^{-1}$, and conclude that $\mathbf{a} = \mathbf{0}$. But by assumption, **a** is *not* zero, so the matrix (**T** - $\lambda$**1**) must in fact be singular, which means that its determinant vanishes:

$$\det(\mathbf{T}-\lambda\mathbf{1}) = \begin{vmatrix} (T_{11}-\lambda) & T_{12} & \cdots & T_{1n} \\ T_{21} & (T_{22}-\lambda) & \cdots & T_{2n} \\ \vdots & \vdots & & \vdots \\ T_{n1} & T_{n2} & \cdots & (T_{nn}-\lambda) \end{vmatrix} = 0. \tag{IV-61}$$

**d)** Expansion of the determinant yields an algebraic equation for $\lambda$:

$$C_n\lambda^n + C_{n-1}\lambda^{n-1} + \cdots + C_1\lambda + C_0 = 0, \tag{IV-62}$$

where the coefficients $C_i$ depend on the elements of **T**. This is called the **characteristic equation** for the matrix — its solutions determine the eigenvalues. Note that it is an $n^{\text{th}}$-order equation, so it has $n$ (complex) roots.

**i)** Some of these root may be duplicates, so all we can say for certain is that an $n$ x $n$ matrix has *at least one* and *at most* $n$ distinct eigenvalues.

**ii)** In the cases where duplicates exist, such states are said to be **degenerate**.

**iii)** To construct the corresponding eigenvectors, it is generally easiest simply to plug each $\lambda$ back into Eq. (IV-59) and solve (by hand) for the components of **a** (see Examples IV-3 and IV-4).

6. In many physical problems involving matrices in both classical mechanics and quantum mechanics it is desirable to carry out a (real) orthogonal similarity transformation or a unitary transformation to reduce the matrix to its diagonal form (*i.e.*, all nondiagonal elements equal to zero).

**a)** If eigenvectors span the space, we are free to use them as a basis

$$\begin{aligned} \hat{T}|f_1\rangle &= \lambda_1|f_1\rangle \\ \hat{T}|f_2\rangle &= \lambda_2|f_2\rangle \\ &\cdots \\ \hat{T}|f_n\rangle &= \lambda_n|f_n\rangle \end{aligned}$$

**b)** The matrix representing $\hat{T}$ takes on a very simple form in this basis, with the eigenvalues strung out along the main diagonal and all other elements zero:

$$\mathbf{T} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} . \qquad \text{(IV-63)}$$

**c)** The (normalized) eigenvectors are equally simple:

$$\mathbf{a}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{a}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{a}^{(n)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} . \qquad \text{(IV-64)}$$

**d)** A matrix that can be brought to **diagonal form** (Eq. IV-63) by change of basis is said to be **diagonalizable**.

**e)** In a geometrical sense, diagonalizing a matrix is equivalent to rotating the bases of a matrix about some point in the space until all of the off-diagonal elements go to zero. If $\mathbf{D}$ is the diagonalized matrix of matrix $\mathbf{M}$, the operation that diagonalizes $\mathbf{M}$ is

$$\mathbf{D} = \mathbf{SMS}^{-1} , \tag{IV-65}$$

where matrix $\mathbf{S}$ is called a similarity transformation. Note that the inverse of the **similarity matrix** can be constructed by using the eigenvectors (in the old basis) as the columns of $\mathbf{S}^{-1}$:

$$\left(\mathbf{S}^{-1}\right)_{ij} = \left(\mathbf{a}^{(j)}\right)_i . \tag{IV-66}$$

**f)** There is great advantage in bringing a matrix to diagonal form — it is much easier to work with. Unfortunately, not every matrix can be diagonalized — **the eigenvectors have to span the space for a matrix to be diagonalizable**.

**7.** The Hermitian conjugate of a linear transformation (called a **Hermitian transformation**) is that transformation $\hat{T}^\dagger$ which, when applied to the *first* member of an inner product, gives the same result as if $\hat{T}$ itself had been applied to the *second* vector:

$$\langle \hat{T}^\dagger \alpha | \beta \rangle = \langle \alpha | \hat{T} | \beta \rangle \tag{IV-67}$$

(for all vectors $|\alpha\rangle$ and $|\beta\rangle$).

**a)** Note that in the notation used in Eq. (IV-63), $\langle \hat{T}^\dagger \alpha | \beta \rangle$ means the inner product of the vector $\hat{T}^\dagger |\alpha\rangle$.

**b)** Note that we can also write

$$\langle \alpha | \hat{T} | \beta \rangle = \mathbf{a}^\dagger \mathbf{T} \mathbf{b} = (\mathbf{T}^\dagger \mathbf{a})^\dagger \mathbf{b} = \langle \hat{T}^\dagger \alpha | \beta \rangle. \tag{IV-68}$$

**c)** In quantum mechanics, a fundamental role is played by Hermitian transformations ($\hat{T}^\dagger = \hat{T}$). The eigenvectors and eigenvalues of a Hermitian transformation have 3 crucial properties (see Morrison §10.6 starting on page 464 for proofs to these theorems):

i) **The eigenvalues of a Hermitian transformation are real.**

ii) **The eigenvectors of a Hermitian transformation belonging to distinct eigenvalues are orthogonal.**

iii) **The eigenvectors of a Hermitian transformation span the space.**

---

**Example IV–1.** Prove the Schwartz inequality (Eq. IV-24). *Hint*: Define a new vector as a linear combination of $\alpha$ and $\beta$, then use Eqs. (IV-18), (IV-19), and (IV-20).

**Solution:**

**Step (a):**

Let's define the vector $|\gamma\rangle = |\beta\rangle + f|\alpha\rangle$, hence $|\gamma\rangle$ is a linear combination of $|\alpha\rangle$ and $|\beta\rangle$. But what is the value of the scale factor $f$? Its value is arbitrary here. There are two ways we can determine a specific value. In both cases for convenience, let's assume that $|\alpha\rangle$ and $|\beta\rangle$ are real vectors and $f$ is a real function.

**Method 1:** (This method is based on the method shown in Arfken's *Mathematical Methods for Physicists* on page 445.) Express Eq. (IV-19) in summation notation and take the minimum value of this equation: $\langle\gamma|\gamma\rangle = \sum \gamma_i^* \gamma_i = \sum \gamma_i^2 = 0$. Then,

$$\sum(\beta_i + f\alpha_i)^2 = \sum \alpha_i^2 \left(\frac{\beta_i}{\alpha_i} + f\right)^2 = 0.$$

If $\beta_i/\alpha_i$ is constant for all $i$, then $f = -\beta_i/\alpha_i$. But if $\beta_i/\alpha_i$ is <u>not</u> constant (which we assume here for a more general expression), then we must expand the polynomial out in the summation above and solve for $f$ using the quadratic formula:

$$\begin{aligned} 0 &= \sum(\beta_i^2 + 2f\alpha_i\beta_i + f^2\alpha_i^2) \\ &= \left(\sum\beta_i^2 + 2f\sum\alpha_i\beta_i + f^2\sum\alpha_i^2\right) \\ &= c + bf + af^2 \end{aligned}$$

where $c = \sum\beta_i^2$, $b = 2\sum\alpha_i\beta_i$, and $a = \sum\alpha_i^2$. Then the quadratic equation gives

$$\begin{aligned} f &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{-2\sum\alpha_i\beta_i \pm \sqrt{4\left(\sum\alpha_i\beta_i\right)^2 - 4\sum\alpha_i^2\sum\beta_i^2}}{2\sum\alpha_i^2} \\ &= \frac{-2\sum\alpha_i\beta_i \pm \sqrt{4\left(\sum\alpha_i\beta_i\right)^2 - 4\left(\sum\alpha_i\sum\beta_i\right)^2}}{2\sum\alpha_i^2} \\ &= \frac{-2\sum\alpha_i\beta_i \pm \sqrt{4\left(\sum\alpha_i\beta_i\right)^2 - 4\left(\sum\alpha_i\beta_i\right)^2}}{2\sum\alpha_i^2} \\ &= \frac{-2\sum\alpha_i\beta_i}{2\sum\alpha_i^2} \\ &= -\frac{\sum\alpha_i\beta_i}{\sum\alpha_i^2} \\ &= -\frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}. \end{aligned}$$

In the solution above, we can write $\left(\sum\alpha_i\sum\beta_i\right)^2$ as a single summation, $\left(\sum\alpha_i\beta_i\right)^2$, since we are summing over the same index.

**Method 2:** (This method is based upon the method described by Anderson's *Modern Physics and Quantum Mechanics* on page 217.) In vector notation,

$$\begin{aligned} 0 \leq & \langle\gamma|\gamma\rangle = \langle(\beta + f\alpha)|(\beta + f\alpha)\rangle \\ & \langle\beta|\beta\rangle + f\langle\beta|\alpha\rangle + f^*\langle\alpha|\beta\rangle + |f|^2\langle\alpha|\alpha\rangle \\ & \langle\beta|\beta\rangle + f\langle\alpha|\beta\rangle + f\langle\alpha|\beta\rangle + f^2\langle\alpha|\alpha\rangle \quad \text{(since } f, |\alpha\rangle, \text{ \& } |\beta\rangle \text{ are real)} \\ & \langle\beta|\beta\rangle + 2f\langle\alpha|\beta\rangle + f^2\langle\alpha|\alpha\rangle. \end{aligned}$$

Since $f$ is arbitrary, we can determine any value for it. Let's allow $f$ to have a value when it is at a minimum, or set $(\partial/\partial f)\langle\gamma|\gamma\rangle = 0$. Thus,

$$\begin{aligned} 0 &= \frac{\partial}{\partial f}\left[\langle\beta|\beta\rangle + 2f\langle\alpha|\beta\rangle + f^2\langle\alpha|\alpha\rangle\right] \\ &= 0 + 2\langle\alpha|\beta\rangle + 2f\langle\alpha|\alpha\rangle \\ 2f\langle\alpha|\alpha\rangle &= -2\langle\alpha|\beta\rangle \\ f &= -\frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}, \end{aligned}$$

which is identical to the value found from the summation approach.

**Step (b):**

As such, we can now write

$$|\gamma\rangle = |\beta\rangle - \left(\frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}\right)|\alpha\rangle.$$

Even though we assumed $f$, $\alpha\rangle$, and $\beta\rangle$ to be real in order to determine a functional form for $f$, keep in mind that it's functional form is arbitrary.

**Step (c):**

Continuing on now with the proof of the Schwartz inequality, we now will use the above functional form for $|\gamma\rangle$ whether or not our vectors are real or complex (again $f$ is just an arbitrary function).

With this form for $|\gamma\rangle$, use Eq. (IV-20) to show

$$\langle\gamma|\gamma\rangle = \langle\gamma\left(|\beta\rangle - \frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}|\alpha\rangle\right) = \langle\gamma|\beta\rangle - \frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}\langle\gamma|\alpha\rangle.$$

Now from Eq. (IV-18):

$$\begin{aligned} \langle\gamma|\beta\rangle^* = \langle\beta|\gamma\rangle &= \langle\beta\left(|\beta\rangle - \frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}|\alpha\rangle\right) = \langle\beta|\beta\rangle - \frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}\langle\beta|\alpha\rangle \\ &= \langle\beta|\beta\rangle - \frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle}, \end{aligned}$$

which is always real, hence, $\langle\gamma|\beta\rangle = \langle\gamma|\beta\rangle^*$. Doing the same thing with the $|\alpha\rangle$ vector gives

$$\langle\gamma|\alpha\rangle^* = \langle\alpha|\gamma\rangle = \langle\alpha\left(|\beta\rangle - \frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}|\alpha\rangle\right) = \langle\alpha|\beta\rangle - \frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}\langle\alpha|\alpha\rangle = |0\rangle,$$

therefore,

$$\langle\gamma|\alpha\rangle = |0\rangle.$$

Finally, plugging this back in to our original equation for $\langle\gamma|\gamma\rangle$ gives

$$\begin{aligned}\langle\gamma|\gamma\rangle &= \langle\beta|\beta\rangle - \frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle} - \frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}|0\rangle \\ &= \langle\beta|\beta\rangle - \frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle} \geq 0,\end{aligned}$$

and hence

$$\begin{aligned}\langle\beta|\beta\rangle &\geq \frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle} \\ \frac{|\langle\alpha|\beta\rangle|^2}{\langle\alpha|\alpha\rangle} &\leq \langle\beta|\beta\rangle \\ |\langle\alpha|\beta\rangle|^2 &\leq \langle\alpha|\alpha\rangle\langle\beta|\beta\rangle. \quad \text{Q.E.D.}\end{aligned}$$

Note that the minimum value is achieved

$$|\langle\alpha|\beta\rangle|^2 = \langle\alpha|\alpha\rangle\langle\beta|\beta\rangle,$$

if $|\alpha\rangle$ is proportional (hence parallel) to $|\beta\rangle$:

$$|\alpha\rangle = \lambda|\beta\rangle,$$

where $\lambda$ is some scalar.

---

**Example IV–2.** Given the following two matrices:

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & i \\ 2 & 0 & 3 \\ 2i & -2i & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 0 & -i \\ 0 & 1 & 0 \\ i & 3 & 2 \end{pmatrix},$$

compute (a) $\mathbf{A}+\mathbf{B}$, (b) $\mathbf{AB}$, (c) $[\mathbf{A},\mathbf{B}]$, (d) $\tilde{\mathbf{A}}$, (e) $\mathbf{A}^*$, (f) $\mathbf{A}^\dagger$, (g) $\mathrm{Tr}(\mathbf{B})$, (h) $\det(\mathbf{B})$, and (i) $\mathbf{B}^{-1}$. Check that $\mathbf{BB}^{-1}=\mathbf{1}$. Does $\mathbf{A}$ have an inverse?

**Solution (a):** Sum the respective elements of the matrix:

$$\mathbf{A}+\mathbf{B}=\begin{pmatrix}-1 & 1 & i\\ 2 & 0 & 3\\ 2i & -2i & 2\end{pmatrix}+\begin{pmatrix}2 & 0 & -i\\ 0 & 1 & 0\\ i & 3 & 2\end{pmatrix}=\boxed{\begin{pmatrix}1 & 1 & 0\\ 2 & 1 & 3\\ 3i & (3-2i) & 4\end{pmatrix}}.$$

**Solution (b):** Multiply rows of $\mathbf{A}$ by columns of $\mathbf{B}$:

$$\begin{aligned}\mathbf{AB} &= \begin{pmatrix}(-2+0-1) & (0+1+3i) & (i+0+2i)\\ (4+0+3i) & (0+0+9) & (-2i+0+6)\\ (4i+0+2i) & (0-2i+6) & (2+0+4)\end{pmatrix}\\ &= \boxed{\begin{pmatrix}-3 & (1+3i) & 3i\\ (4+3i) & 9 & (6-2i)\\ 6i & (6-2i) & 6\end{pmatrix}}.\end{aligned}$$

**Solution (c):** $[\mathbf{A},\mathbf{B}]=\mathbf{AB}-\mathbf{BA}$, we already have $\mathbf{AB}$,

$$\begin{aligned}\mathbf{BA} &= \begin{pmatrix}(-2+0+2) & (2+0-2) & (2i+0-2i)\\ (0+2+0) & (0+0+0) & (0+3+0)\\ (-i+6+4i) & (i+0-4i) & (-1+9+4)\end{pmatrix}\\ &= \begin{pmatrix}0 & 0 & 0\\ 2 & 0 & 3\\ (6+3i) & -3i & 12\end{pmatrix};\end{aligned}$$

$$\begin{aligned}[\mathbf{A},\mathbf{B}] &= \begin{pmatrix}-3 & (1+3i) & 3i\\ (4+3i) & 9 & (6-2i)\\ 6i & (6-2i) & 6\end{pmatrix}-\begin{pmatrix}0 & 0 & 0\\ 2 & 0 & 3\\ (6+3i) & -3i & 12\end{pmatrix}\\ &= \boxed{\begin{pmatrix}-3 & (1+3i) & 3i\\ (2+3i) & 9 & (3-2i)\\ (-6+3i) & (6+i) & -6\end{pmatrix}}.\end{aligned}$$

**Solution (d):** Transpose of **A** — flip **A** about the diagonal:

$$\tilde{\mathbf{A}} = \boxed{\begin{pmatrix} -1 & 2 & 2i \\ 1 & 0 & -2i \\ i & 3 & 2 \end{pmatrix}}.$$

**Solution (e):** Complex conjugate of **A** — multiply each $i$ term by –1 in **A**:

$$\mathbf{A}^* = \boxed{\begin{pmatrix} -1 & 1 & -i \\ 2 & 0 & 3 \\ -2i & 2i & 2 \end{pmatrix}}.$$

**Solution (f):** Hermitian of **A**:

$$\mathbf{A}^\dagger \equiv \tilde{\mathbf{A}}^* = \boxed{\begin{pmatrix} -1 & 2 & -2i \\ 1 & 0 & 2i \\ -i & 3 & 2 \end{pmatrix}}.$$

**Solution (g):** Trace of **B**:

$$\mathrm{Tr}(\mathbf{B}) = \sum_{i=1}^{3} B_{ii} = 2 + 1 + 2 = \boxed{5.}$$

**Solution (h):** Determinant of **B**:

$$\det(\mathbf{B}) = 2(2-0) - 0(0-0) - i(0-i) = 4 - 0 - 1 = \boxed{3.}$$

**Solution (i):** Inverse of **B**:

$$\mathbf{B}^{-1} = \frac{1}{\det(\mathbf{B})}\tilde{\mathbf{C}},$$

where

$$\mathbf{C} = \begin{pmatrix} \begin{vmatrix} 1 & 0 \\ 3 & 2 \end{vmatrix} & -\begin{vmatrix} 0 & 0 \\ i & 2 \end{vmatrix} & \begin{vmatrix} 0 & 1 \\ i & 3 \end{vmatrix} \\ -\begin{vmatrix} 0 & -i \\ 3 & 2 \end{vmatrix} & \begin{vmatrix} 2 & -i \\ i & 2 \end{vmatrix} & -\begin{vmatrix} 2 & 0 \\ i & 3 \end{vmatrix} \\ \begin{vmatrix} 0 & -i \\ 1 & 0 \end{vmatrix} & -\begin{vmatrix} 2 & -i \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 2 & 0 & -i \\ -3i & 3 & -6 \\ i & 0 & 2 \end{pmatrix},$$

then

$$\boxed{\mathbf{B}^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -3i & i \\ 0 & 3 & 0 \\ -i & -6 & 2 \end{pmatrix}}.$$

$$\begin{aligned} \mathbf{B}\mathbf{B}^{-1} &= \frac{1}{3}\begin{pmatrix} (4+0-1) & (-6i+0+6i) & (2i+0-2i) \\ (0+0+0) & (0+3+0) & (0+0+0) \\ (2i+0-2i) & (3+9-12) & (-1+0+4) \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad \surd \end{aligned}$$

If $\det(\mathbf{A}) \neq 0$, then $\mathbf{A}$ has an inverse:

$$\det(\mathbf{A}) = -1(0+6i) - 1(4-6i) + i(-4i-0) = -6i - 4 + 6i + 4 = 0.$$

As such, $\boxed{\mathbf{A} \text{ does } \underline{\text{not}} \text{ have an inverse.}}$

---

**Example IV–3.** Find the eigenvalues and normalized eigenvectors of the following matrix:

$$\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Can this matrix be diagonalized?

**Solution:**

$$\begin{aligned} \mathbf{0} = \det(\mathbf{M} - \lambda\mathbf{1}) &= \begin{vmatrix} (1-\lambda) & 1 \\ 0 & (1-\lambda) \end{vmatrix} \\ &= (1-\lambda)^2 \end{aligned}$$

$$\boxed{\lambda = 1} \quad \text{(only one eigenvalue).}$$

From Eq. (IV-59) we get

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 1 \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} .$$

We get two equations from this eigenvector equation:

$$\begin{aligned} a_1 + a_2 &= a_1 \\ a_2 &= a_2 . \end{aligned}$$

The second equation tells us nothing, but the first equation shows us that $a_2 = 0$. We still need to figure out the value for $a_1$. We can do this by normalizing our eigenvector $\mathbf{a} = |\alpha\rangle$:

$$\begin{aligned} 1 &= \langle\alpha|\alpha\rangle = \sum_{i=1}^{2} |a_i|^2 \\ &= |a_1|^2 + |a_2|^2 = |a_1|^2 \end{aligned}$$

or $a_1 = 1$. Hence our normalized eigenvector,

$$|\alpha\rangle = \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} ,$$

is

$$\boxed{\mathbf{a} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} .}$$

Since these eigenvectors do not span the space (as described on page IV-3, §A.2.c.ii.), this matrix cannot be diagonalized.

---

**Example IV–4.** Find the eigenvalues and eigenvectors of the following matrix:

$$\mathbf{M} = \begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix} .$$

**Solution:**

The characteristic equation is

$$\begin{aligned}
|\mathbf{M} - \lambda \mathbf{1}| &= \begin{vmatrix} (2-\lambda) & 0 & -2 \\ -2i & (i-\lambda) & 2i \\ 1 & 0 & (-1-\lambda) \end{vmatrix} \\
&= (2-\lambda) \begin{vmatrix} (i-\lambda) & 2i \\ 0 & (-1-\lambda) \end{vmatrix} - 0 - 2 \begin{vmatrix} -2i & (i-\lambda) \\ 1 & 0 \end{vmatrix} \\
&= (2-\lambda)[(i-\lambda)(-1-\lambda) - 0] - 2[0 - (i-\lambda)] \\
&= (2-\lambda)(-i - i\lambda + \lambda + \lambda^2) + 2i - 2\lambda) \\
&= -2i - 2i\lambda + 2\lambda + 2\lambda^2 + i\lambda + i\lambda^2 - \lambda^2 - \lambda^3 + 2i - 2\lambda \\
&= -\lambda^3 + (1+i)\lambda^2 - i\lambda = 0 \ .
\end{aligned}$$

To find the roots to this characteristic equation, factor out a $\lambda$ and use the quadratic formula solution equation:

$$\begin{aligned}
0 &= -\lambda^3 + (1+i)\lambda - i\lambda \\
&= [-\lambda^2 + (1+i)\lambda - i]\lambda \\
\lambda_1 &= 0 \\
\lambda_{2,3} &= \frac{-(1+i) \pm \sqrt{(1+i)^2 - 4i}}{-2} \\
&= \frac{-(1+i) \pm \sqrt{(1+2i-1) - 4i}}{-2} \\
&= \frac{-(1+i) \pm \sqrt{-2i}}{-2} \ .
\end{aligned}$$

However note that $(1-i)^2 = -2i$. As such, the equation above becomes

$$\begin{aligned}
\lambda_{2,3} &= \frac{-(1+i) \pm \sqrt{(1-i)^2}}{-2} \\
&= \frac{-(1+i) \pm (1-i)}{-2} \\
\lambda_2 &= \frac{-(1+i) - (1-i)}{-2} = \frac{-2}{-2} = 1 \\
\lambda_3 &= \frac{-(1+i) + (1-i)}{-2} = \frac{-2i}{-2} = i \ ,
\end{aligned}$$

so the roots of $\lambda$ (*i.e.*, the eigenvalues) are 0, 1, and $i$. Now, let's call the components of the first eigenvector $|\alpha\rangle$ $(a_1, a_2, a_3)$ which corresponds to eigenvalue $\lambda_1 = 0$. The eigenvector equation becomes

$$\begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = 0 \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} ,$$

which yield 3 equations:

$$\begin{aligned} 2a_1 - 2a_3 &= 0 \\ -2ia_1 + ia_2 + 2ia_3 &= 0 \\ a_1 - a_3 &= 0 \ . \end{aligned}$$

The first equation gives $a_3 = a_1$, the second gives $a_2 = 0$, and the third is redundant with the first equation. We can find the values for $a_1$ and $a_3$ by normalizing:

$$\begin{aligned} 1 &= \langle\alpha|\alpha\rangle = \sum_{i=1}^{3} |a_i|^2 \\ &= |a_1|^2 + |a_2|^2 + |a_3|^2 = |a_1|^2 + |a_1|^2 \\ &= 2\,|a_1|^2 \ , \end{aligned}$$

or $a_1 = a_3 = (1/\sqrt{2}) = \sqrt{2}/2$. Hence our eigenvector for $\lambda_1$ is

$$\boxed{|\alpha\rangle = \mathbf{a} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} , \ \text{for } \lambda_1 = 0 \ .}$$

For the second eigenvector, let's call it $|\beta\rangle = \mathbf{b}$, we have

$$\begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = 1 \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} ,$$

which yield the equations:

$$\begin{aligned} 2b_1 - 2b_3 &= b_1 \\ -2ib_1 + ib_2 + 2ib_3 &= b_2 \\ b_1 - b_3 &= b_3 \ , \end{aligned}$$

with the solutions $b_3 = (1/2)b_1$ and $b_2 = [(1-i)/2]b_1$. Normalizing gives

$$\begin{aligned}
1 &= \langle\beta|\beta\rangle = \sum_{i=1}^{3} |b_i|^2 \\
&= |b_1|^2 + |b_2|^2 + |b_3|^2 \\
&= |b_1|^2 + \left(\frac{1+i}{2}\right)\left(\frac{1-i}{2}\right)|b_1|^2 + \frac{1}{4}|b_1|^2 \\
&= |b_1|^2 + \left(\frac{1+i-i+1}{4}\right)|b_1|^2 + \frac{1}{4}|b_1|^2 \\
&= \frac{4}{4}|b_1|^2 + \frac{2}{4}|b_1|^2 + \frac{1}{4}|b_1|^2 \\
&= \frac{7}{4}|b_1|^2 \ ,
\end{aligned}$$

or $b_1 = (2/\sqrt{7})$. So $b_2 = [(1-i)/\sqrt{7}]$ and $b_3 = (1/\sqrt{7})$ giving our final eigenvector for $\lambda_2$ as

$$\boxed{|\beta\rangle = \mathbf{b} = \frac{\sqrt{7}}{7}\begin{pmatrix} 2 \\ (1-i) \\ 1 \end{pmatrix}, \text{ for } \lambda_2 = 1 \ .}$$

Finally, the third eigenvector (call it $|\gamma\rangle = \mathbf{c}$) is

$$\begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = i\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} ic_1 \\ ic_2 \\ ic_3 \end{pmatrix} \ ,$$

which gives the equations:

$$\begin{aligned}
2c_1 - 2c_3 &= ic_1 \\
-2ic_1 + ic_2 + 2ic_3 &= ic_2 \\
c_1 - c_3 &= ic_3 \ ,
\end{aligned}$$

with the solutions $c_3 = c_1 = 0$, with $c_2$ undetermined. Once again, we can normalize our eigenvector to determine this undetermined $c_2$ coefficient:

$$\begin{aligned}
1 &= \langle\gamma|\gamma\rangle = \sum_{i=1}^{3} |c_i|^2 \\
&= |c_1|^2 + |c_2|^2 + |c_3|^2 = |c_2|^2 \ ,
\end{aligned}$$

or $c_2 = 1$, which gives our third eigenvector:

$$\boxed{|\gamma\rangle = \mathbf{c} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \text{ for } \lambda_3 = i\ .}$$

---

## B. Function Spaces.

### 1. Functions as Vectors.

**a)** In quantum mechanics, we introduce the concept of a **function space** in which *vectors* are actually (complex) functions of $x$, inner products are integrals, and derivatives appear as linear transformations.

**b)** The **inner product of two functions** $[f(x)$ and $g(x)]$ is defined by the integral given in Eq. (IV-5):

$$\boxed{\langle f|g\rangle = \int f(x)^* g(x)\, dx\ ,} \qquad \text{(IV-69)}$$

where the limits of this integral will depend on the domain of the functions in question.

**i)** This integral may not *converge* $\Longrightarrow$ if we want a function space with an inner product, we must restrict the class of functions so as to ensure that $\langle f|g\rangle$ is always well defined.

**ii)** It is clearly *necessary* that every admissible function be **square integrable**:

$$\int |f(x)|^2\, dx < \infty, \qquad \text{(IV-70)}$$

otherwise the inner product of $f$ with *itself* wouldn't even exist.

**iii)** It turns out that this restriction is also *sufficient* — if $f$ and $g$ are both square integrable, then the integral in Eq. (IV-69) is necessarily finite.

---

**Example IV–5.** Let $T(N)$ be the set of all trigonometric functions of the form

$$f(x) = \sum_{n=0}^{N-1} [a_n \sin(n\pi x) + b_n \cos(n\pi x)], \qquad \text{(IV-71)}$$

on the interval $-1 \leq x \leq 1$. Show that

$$|e_n\rangle = \frac{1}{\sqrt{2}}\, e^{in\pi x}, \quad (n = 0, \pm 1, \pm 2, ..., \pm(N-1)) \qquad \text{(IV-72)}$$

constitute an orthonormal basis for this function. What is the dimension of this space?

**Solution:**

Rewrite the trig functions using Euler's relations,

$$\begin{aligned} f(x) &= \sum_{n=0}^{N-1} \left[ \left(\frac{a_n}{2i}\right) \left(e^{in\pi x} - e^{-in\pi x}\right) + \left(\frac{b_n}{2}\right) \left(e^{in\pi x} + e^{-in\pi x}\right) \right] \\ &= \sum_{n=0}^{N-1} \left[ \left(\frac{a_n}{2i} + \frac{b_n}{2}\right) e^{in\pi x} + \left(-\frac{a_n}{2i} + \frac{b_n}{2}\right) e^{-in\pi x} \right] \\ &= \sum_{n=-(N-1)}^{(N-1)} c_n\, e^{in\pi x}, \end{aligned}$$

where

$$\left\{ \begin{array}{lll} c_n &=& \frac{1}{2}(-ia_n + b_n), \quad \text{for } n = 1, 2, 3, ..., N-1 \\ c_0 &=& b_0 \\ c_n &=& \frac{1}{2}(ia_{-n} + b_{-n}), \quad \text{for } n = -1, -2, -3, ..., -(N-1) \end{array} \right\}$$

(see §IV.A.c.ii on page IV-3). So the set <u>does</u> span the space. Is it orthonormal?

$$\begin{aligned} \langle e_m | e_n \rangle &= \frac{1}{2} \int_{-1}^{1} e^{-im\pi x} e^{in\pi x}\, dx \\ &= \left\{ \begin{array}{ll} \frac{1}{2}\, \frac{e^{-i(m-n)\pi x}}{-i(m-n)\pi} \Big|_{-1}^{1} = 0, & \text{for } m \neq n \\ \frac{1}{2} \int_{-1}^{1} dx = 1, & \text{for } m = n \end{array} \right\} \\ &= \delta_{mn}. \end{aligned}$$

Yes it is orthonormal. So it's also a basis (*i.e.*, no "extra" functions included), since orthogonal vectors are necessarily linearly independent.

Looking at the $c_n$ coefficient equations above, the dimensions are

$$D = (N-1) + 1 + (N-1) = 2(N-1) + 1 = \boxed{2N-1} \ .$$

---

**2. Operators as Linear Transformations.**

**a)** In function spaces, *operators* (such as $d/dx$, $d^2/dx^2$, or simply $x$) behave as linear transformations, provided that they carry functions in the space into other functions in the space and satisfy the linearity condition (Eq. IV-26).

**i)** For example, in the polynomial space $P(N)$, the derivative operator ($\hat{D} \equiv d/dx$) is a linear transformation, since it takes $N$th-order polynomials into ($N$-1)th-order polynomials $\Longrightarrow$ hence still in the space.

**ii)** However, the operator $\hat{x}$ (multiplication by $x$) is *not*, for it takes ($N$-1)th-order polynomials into $N$th-order polynomials, which is no longer in the space.

**b)** In a function space, the eigenvectors of an operator $\hat{T}$ are called **eigenfunctions**:

$$\hat{T} f(x) = \lambda f(x). \qquad \text{(IV-73)}$$

**c)** A Hermitian operator is one that satisfies the defining condition (Eq. IV-63):

$$\langle f|\hat{T}|g\rangle = \langle \hat{T} f|g\rangle, \qquad \text{(IV-74)}$$

for all functions $f(x)$ and $g(x)$ in the space.

**d)** When dealing with operators you must always keep in mind the function space you're working in. An operator may not be a legitimate linear transformation because:

**i)** It carries functions out of the space.

**ii)** The eigenfunctions of an operator may not reside in the space.

**iii)** An operator that is Hermitian in one space may *not* be Hermitian in another.

**e)** One has to pay particular attention to transformations in infinite spaces:

**i)** Remember that $\hat{x}$ is *not* a linear transformation in the space $P(N)$ since multiplication by $x$ increases the order of the polynomial and hence takes functions outside the space.

**ii)** However, it *is* a linear transformation on $P(\infty)$, the space of *all* polynomials, on the interval $-1 \leq x \leq 1$.

**iii)** In fact, it's a *Hermitian* transformation, since

$$\int_{-1}^{1} [f(x)]^* \, [xg(x)] \; dx = \int_{-1}^{1} [xf(x)]^* \, [g(x)] \; dx .$$

---

**Example IV–6.** Show that $e^{-x^2/2}$ is an eigenfunction of the operator $\hat{Q} = (d^2/dx^2) - x^2$, and find its eigenvalues.

**Solution:**

The eigenfunction equation is

$$\hat{Q} f(x) = \lambda f(x).$$

So carrying out this operation, we get

$$\begin{aligned}
\hat{Q} f(x) &= \left( \frac{d^2}{dx^2} - x^2 \right) e^{-x^2/2} \\
&= \frac{d}{dx} \left( -x e^{-x^2/2} \right) - x^2 e^{-x^2/2} \\
&= -e^{-x^2/2} + x^2 e^{-x^2/2} - x^2 e^{-x^2/2} \\
&= -e^{-x^2/2} = -f(x).
\end{aligned}$$

So it is an eigenfunction with only one eigenvalue of $\boxed{\lambda = -1}$ .

---

## 3. Hilbert Space.

**a)** We will now start to talk about wave functions in 3-dimensional space. We have just been discussing the orthonormality of wave functions, now we define the **completeness** of a function.

**i)** We have shown that the generalized wave function is a linear combination of separable solutions (see Eq. III-21). In terms of the TISE, we can write this condition as

$$\psi = \sum_n a_n \, \psi_n(\boldsymbol{r}). \qquad \text{(IV-75)}$$

**ii)** The wave function components are orthonormal to each other if they satisfy the condition

$$\int_V \psi_m^*(\boldsymbol{r}) \, \psi_n(\boldsymbol{r}) \, dV = \delta_{mn}. \qquad \text{(IV-76)}$$

**iii)** In addition, the wave function components $\psi_n$ represent a **complete** system if it is impossible to find an additional function $\phi$ that is orthogonal to all of the $\psi_n$'s in the sense of Eq. (IV-76).

**iv)** If this is the case, the following completeness relation is valid:

$$\int_V \psi^* \, \psi \, dV = \int_V |\psi|^2 \, dV = \sum_n |a_n|^2, \qquad \text{(IV-77)}$$

where the $a_n$'s are the expansion coefficients of the arbitrary wave function as defined in Eq. (IV-75).

**b)** If completeness holds, the $\psi_n$'s constitute an orthonormal basis of a **Hilbert** space.

**i)** A Hilbert space is a finite or infinite complete vector space on the basic field of complex numbers.

**ii)** In this space a scalar product is defined such that it assigns a complex number to each pair of functions $\psi(x)$ and $\phi(x)$ out of a set of linear functions.

**iii)** This scalar product meets three requirements:

$$\begin{aligned}
(1) \qquad \langle\psi|\phi\rangle = \int \psi^* \phi \, dV &= \left(\int \phi^* \psi \, dV\right)^* = (\langle\phi|\psi\rangle)^*, \\
(2) \qquad \langle\psi|(a\phi_1 + b\phi_2)\rangle &= a\langle\psi|\phi_1\rangle + b\langle\psi|\phi_2\rangle \quad \text{or,} \\
\int \psi^* \left(a\phi_1 + b\phi_2\right) dV &= a \int \psi^* \phi_1 \, dV + b \int \psi^* \phi_2 \, dV, \\
(3) \qquad \langle\psi|\psi\rangle &= \int \psi^* \psi \, dV \geq 0.
\end{aligned}$$

Note for the last requirement, $\langle\psi|\psi\rangle = 0$ only if $\psi = 0$.

**c)** The state vectors (*i.e.*, wave functions) of a quantum mechanical system constitute a Hilbert space (hence, the Hilbert space is a function space).

d) Mathematicians refer to a complete inner product space as $L_2$. To physicists, $L_2$ is practically synonymous with Hilbert space.

4. We now recast the fundamental principles of quantum mechanics (as developed in §§II-III of the notes) in the more elegant language of linear algebra and function (*i.e.*, Hilbert) space. Remember that the state of a particle is represented by its wave function, $\Psi(x,t)$, whose absolute square is the probability density for finding the particle at point $x$ (or $\boldsymbol{r}$ in 3-D), at time $t$. It follows that $\Psi$ must be *normalized*, which is possible if and only if it is square integrable.

## C. The Generalized Statistical Interpretation.

1. **The state of a particle is represented by a normalized vector ($|\Psi\rangle$) in the Hilbert space $L_2$.**

   a) Classical dynamical quantities (such as position, velocity, momentum, and kinetic energy) can be expressed as functions of the "canonical" variables $x$ (or $\boldsymbol{r}$) and $p$ (and sometimes $t$): $Q(x,p,t)$. To each such classical observables we associate a quantum-mechanical *operator*, $\hat{Q}$, obtained from $Q$ by the substitution

$$p \rightarrow \frac{\hbar}{i}\frac{\partial}{\partial x}. \qquad \text{(IV-78)}$$

   b) The expectation value of $Q$, in the state $\Psi$, is

$$\langle Q\rangle = \int \Psi^*(x,t)\,\hat{Q}\,\Psi(x,t)\,dx, \qquad \text{(IV-79)}$$

   which we now write as an inner product:

$$\langle Q\rangle = \langle\Psi|\hat{Q}|\Psi\rangle = \langle\Psi|\hat{Q}\Psi\rangle. \qquad \text{(IV-80)}$$

   (Note that either notation, $\hat{Q}|\Psi$ or $\hat{Q}\Psi$, is considered acceptable.)

**c)** The expectation value of an observable quantity has got to be a *real* number, so

$$\langle\Psi|\hat{Q}|\Psi\rangle = \langle\Psi|\hat{Q}|\Psi\rangle^* = \langle\hat{Q}\Psi|\Psi\rangle, \qquad \text{(IV-81)}$$

for all vectors $|\Psi\rangle \Longrightarrow \hat{Q}$ must be a *Hermitian* operator.

**2. Observable quantities, $Q(x, p, t)$, are represented by Hermitian operators, $\hat{Q}(x, \frac{\hbar}{i}\frac{\partial}{\partial x}, t)$; the expectation value of $Q$, in the state $|\Psi\rangle$, is $\langle\Psi|\hat{Q}|\Psi\rangle$.**

**3. A measurement of the observable $Q$ on a particle in the state $|\Psi\rangle$ is *certain* to return the value $\lambda$ if and only if $|\Psi\rangle$ is an eigenvector of $\hat{Q}$, with eigenvalue $\lambda$.**

**a)** For example, the TISE (Eq. III-8) can be written in the form

$$\hat{H}\,\psi = E\,\psi. \qquad \text{(IV-82)}$$

**b)** Note, however, that this is nothing more than an eigenvalue equation for the Hamiltonian operator and the solutions are states of determinate energy $E$.

**c)** This third postulate can be rewritten in terms of a statistical argument $\Longrightarrow$ the **generalized statistical interpretation** (GSI) as given in the following postulate 4.

**4. If you measure an observable $Q$ on a particle in the state $|\Psi\rangle$, you are certain to get one of the eigenvalues of $\hat{Q}$. The probability of getting the particular eigenvalue $\lambda$ is equal to the absolute square of the $\lambda$ component of $|\Psi\rangle$, when expressed in the orthonormal basis of eigenvectors.**

**a)** To sustain this postulate, it is essential that the eigenfunctions of $\hat{Q}$ span the space.

**b)** This might not be possible however for infinite-dimensional cases. As such, *we shall take it as a restriction on the subset of Hermitian operators that are* **observable**, *that their eigenfunctions constitute a complete set* (though they need not fall inside $L_2$).

**c)** There are two kinds of eigenvectors which we need to treat separately. The first deals with systems whose spectra are *discrete* (with the discrete eigenvalues separated by finite gaps — *e.g.*, bound states in an atom).

**i)** We can label their eigenvectors with an integer $n$:

$$\hat{Q}|e_n\rangle = \lambda_n|e_n\rangle, \quad \text{with } n = 1, 2, 3, ... \qquad \text{(IV-83)}$$

**ii)** The eigenvectors are orthonormal (or rather, they can always be chosen so):

$$\langle e_m|e_n\rangle = \delta_{mn}. \qquad \text{(IV-84)}$$

**iii)** The completeness relation takes the form of a *sum*:

$$|\Psi\rangle = \sum_{n=1}^{\infty} c_n|e_n\rangle, \qquad \text{(IV-85)}$$

with the components given by

$$c_n = \langle e_n|\Psi\rangle, \qquad \text{(IV-86)}$$

and the probability of getting the particular eigenvalue $\lambda_n$ is

$$\boxed{|c_n|^2 = |\langle e_n|\Psi\rangle|^2.} \qquad \text{(IV-87)}$$

**d)** The second deals with systems whose spectra are *continuous* (*e.g.*, ionization states of an atom or scattering states).

**i)** The eigenvectors are labeled by a continuous variable $k$:

$$\hat{Q}|e_k\rangle = \lambda_k|e_k\rangle, \quad \text{with } -\infty < k < \infty. \qquad \text{(IV-88)}$$

**ii)** Here, the eigenfunctions are *not* normalizable, but they satisfy a sort of "orthonormality" condition:

$$\langle e_\ell|e_k\rangle = \delta(\ell - k) \qquad \text{(IV-89)}$$

(or rather, they can always be *chosen* so).

**iii)** The completeness relation takes the form of an *integral*:

$$|\Psi\rangle = \int_{-\infty}^{\infty} c_k|e_k\rangle\, dk, \qquad \text{(IV-90)}$$

with the components given by

$$c_k = \langle e_k|\Psi\rangle, \qquad \text{(IV-91)}$$

and the probability of getting an eigenvalue in the *range dk* about $\lambda_k$ is

$$\boxed{|c_k|^2\, dk = |\langle e_k|\Psi\rangle|^2\, dk.} \qquad \text{(IV-92)}$$

**e)** In the GSI, the "orthonormal" eigenfunctions of the position operator are

$$e_{x'}(x) = \delta(x - x'), \qquad \text{(IV-93)}$$

and the eigenvalue $(x')$ can take on any value between $-\infty$ and $\infty$.

**i)** The $x'$ "component" of $|\Psi\rangle$ is

$$c_{x'} = \langle e_{x'}|\Psi\rangle = \int_{-\infty}^{\infty} \delta(x - x')\Psi(x,t)\, dx = \Psi(x', t). \qquad \text{(IV-94)}$$

**ii)** Thus, the probability of finding the particle in the range $dx'$ about $x'$ is

$$|c_{x'}|^2 \, dx' = |\Psi(x', t)|^2 \, dx', \qquad \text{(IV-95)}$$

which is the original statistical interpretation of $\Psi$.

**f)** The momentum operator in the GSI is handled in the following manner:

**i)** Its "orthonormal" eigenfunctions are

$$e_p(x) = \frac{1}{\sqrt{2\pi\hbar}} \, e^{ipx/\hbar}, \qquad \text{(IV-96)}$$

and the eigenvalue ($p$) can take on any value in the range $-\infty < p < \infty$.

**ii)** The $p$ "component" of $|\Psi\rangle$ is

$$\boxed{c_p = \langle e_p|\Psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \, \Psi(x, t) \, dx \equiv \Phi(p, t).} \qquad \text{(IV-97)}$$

**iii)** Here, $\Phi(p, t)$ is called the **momentum-space wave function** — it is (apart from the factors of $\hbar$) the *Fourier transform* of the position-space wave function $\Psi(x, t)$.

**iv)** The probability of getting the momentum in the range $dp$ and $p$ is

$$P \, dp = |\Phi(p, t)|^2 \, dp \, . \qquad \text{(IV-98)}$$

**Example IV–7.** Problems concerning the generalized statistical interpretation of quantum mechanics:

**(a)** Show that $\sum |c_n|^2 = 1$ in Equation (IV-85).

**(b)** Show that $\int |c_k|^2 \, dk = 1$ in Equation (IV-90).

**(c)** From postulate 4 (*i.e.*, the generalized statistical interpretation) it follows that

$$\langle Q \rangle = \sum \lambda_n |c_n|^2, \quad \text{or} \quad \langle Q \rangle = \int \lambda_k |c_k|^2 \, dk,$$

for discrete and continuous spectra, respectively. Show that this is consistent with postulate 2: $\langle Q \rangle = \langle \Psi | \hat{Q} | \Psi \rangle$.

**Solution (a):**

$$\begin{aligned} 1 = \langle \Psi | \Psi \rangle &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m^* c_n \langle e_m | e_n \rangle \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m^* c_n \delta_{mn} \\ &= \sum_{n=1}^{\infty} |c_n|^2 \qquad \text{Q.E.D.} \end{aligned}$$

**Solution (b):**

$$\begin{aligned} 1 = \langle \Psi | \Psi \rangle &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c_\ell^* c_k \langle e_\ell | e_k \rangle \, d\ell \, dk \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c_\ell^* c_k \delta(\ell - k) \, d\ell \, dk \\ &= \int_{-\infty}^{\infty} |c_k|^2 \, dk \qquad \text{Q.E.D.} \end{aligned}$$

**Solution (c):** From Eq. (IV-85) and postulate 2:

$$\begin{aligned} \langle Q \rangle = \langle \Psi | \hat{Q} | \Psi \rangle &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m^* c_n \langle e_m | \hat{Q} | e_n \rangle \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m^* c_n \lambda_n \langle e_m | e_n \rangle \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m^* c_n \lambda_n \delta_{mn} \\ &= \sum_{n=1}^{\infty} \lambda_n |c_n|^2 \end{aligned}$$

From Eq. (IV-90) and postulate 2:

$$\begin{aligned}
\langle Q\rangle = \langle\Psi|\hat{Q}|\Psi\rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c_\ell^* c_k \langle e_\ell|\hat{Q}|e_k\rangle\, d\ell\, dk \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c_\ell^* c_k \lambda_k \langle e_\ell|e_k\rangle\, d\ell\, dk \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c_\ell^* c_k \lambda_k \delta(\ell - k)\, d\ell\, dk \\
&= \int_{-\infty}^{\infty} \lambda_k |c_k|^2\, dk
\end{aligned}$$

---

**Example IV–8.** Confirm that $e_p(x)$ (in Eq. IV-96) is the "orthonormal" eigenfunction of the momentum operator, with eigenvalue $p$.

**Solution:**

The momentum operator is $\hat{p} = (\hbar/i)d/dx$. Applying this operator on Eq. (IV-96) gives

$$\hat{p}\, e_p = \frac{\hbar}{i}\frac{d}{dx}\, e_p = \frac{\hbar}{i}\frac{1}{\sqrt{2\pi\hbar}}\frac{ip}{\hbar}\, e^{ipx/\hbar} = p\, e_p\ .$$

Hence $e_p$ is an eigenfunction of the momentum operator with an eigenvalue of $p$.

To prove that it is orthonormal, we need to take the inner product of this eigenfunction:

$$\begin{aligned}
\langle e_p|e_q\rangle &= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} e^{-ipx/\hbar} e^{iqx/\hbar}\, dx \\
&= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} e^{i(q-p)y}\hbar\, dy \quad (\text{with } y \equiv x/\hbar) \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(q-p)y}\, dy\ .
\end{aligned}$$

With the use of Plancherel's theorem (see Eq. III-142), let $f(x) = \delta(x)$ in Eq. (III-142), so the Fourier transform of the delta-function is

$$F(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \delta(x)\, e^{-ikx}\, dx = \frac{1}{\sqrt{2\pi}}\, e^{-ik\cdot 0} = \frac{1}{\sqrt{2\pi}}$$

and the inverse Fourier transform of $F(k) = 1/\sqrt{2\pi}$ is

$$f(k) = \delta(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{ikx}\, dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\, dx\ .$$

Comparing this equation with the equation derived for $\langle e_p | e_q \rangle$, we see that they are the same if $q - p = x$ and $y = k$. Making these variable substitutions in our equation above, we see that

$$\langle e_p | e_q \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iy(q-p)}\, dy = \delta(q - p) \ .$$

As such, $|e_p\rangle$ and $|e_q\rangle$ are "orthonormal," (hence are basis functions) in the sense of Eq. (IV-93).

## D. The Uncertainty Principle.

### 1. Proof of the Generalized Uncertainty Principle.

**a)** For any observable $A$, we can express the variance in the measurement of $A$ as

$$\sigma_A^2 = \langle (\hat{A} - \langle A \rangle)\Psi | (\hat{A} - \langle A \rangle)\Psi \rangle = \langle f | f \rangle, \qquad \text{(IV-99)}$$

where $|f\rangle \equiv (\hat{A} - \langle A \rangle)\Psi\rangle$.

**b)** Likewise, for any *other* observable $B$, we have

$$\sigma_B^2 = \langle g | g \rangle,$$

where $|g\rangle \equiv (\hat{B} - \langle B \rangle)\Psi\rangle$.

**c)** Invoking the Schwartz inequality (Eq. IV-24), we get

$$\sigma_A^2 \sigma_B^2 = \langle f | f \rangle \langle g | g \rangle \geq |\langle f | g \rangle|^2. \qquad \text{(IV-100)}$$

**i)** From the mathematics of complex variables, for any complex number:

$$|z|^2 = (\mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2 \geq (\mathrm{Im}(z))^2 = \left[\frac{1}{2i}(z - z^*)\right]^2, \qquad \text{(IV-101)}$$

where $z^*$ is the complex conjugate of $z$.

**ii)** Therefore, letting $z = \langle f | g \rangle$,

$$\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i}\left[\langle f | g \rangle - \langle g | f \rangle\right]\right)^2. \qquad \text{(IV-102)}$$

**iii)** But

$$\begin{aligned}
\langle f | g \rangle &= \langle (\hat{A} - \langle A \rangle)\Psi | (\hat{B} - \langle B \rangle)\Psi \rangle \\
&= \langle \Psi | (\hat{A} - \langle A \rangle)(\hat{B} - \langle B \rangle)\Psi \rangle \\
&= \langle \Psi | (\hat{A}\hat{B} - \hat{A}\langle B \rangle - \hat{B}\langle A \rangle + \langle A \rangle\langle B \rangle)\Psi \rangle
\end{aligned}$$

$$
\begin{aligned}
&= \langle\Psi|\hat{A}\hat{B}\Psi\rangle - \langle B\rangle\langle\Psi|\hat{A}\Psi\rangle - \langle A\rangle\langle\Psi|\hat{B}\Psi\rangle + \\
&\qquad \langle A\rangle\langle B\rangle\langle\Psi|\Psi\rangle \\
&= \langle\hat{A}\hat{B}\rangle - \langle B\rangle\langle A\rangle - \langle A\rangle\langle B\rangle + \langle A\rangle\langle B\rangle \\
&= \langle\hat{A}\hat{B}\rangle - \langle B\rangle\langle A\rangle = \langle\hat{A}\hat{B}\rangle - \langle A\rangle\langle B\rangle \ .
\end{aligned}
$$

**iv)** Similarly,

$$\langle g|f\rangle = \langle\hat{B}\hat{A}\rangle - \langle A\rangle\langle B\rangle,$$

so

$$\langle f|g\rangle - \langle g|f\rangle = \langle\hat{A}\hat{B}\rangle - \langle\hat{B}\hat{A}\rangle = \langle\,[\hat{A},\hat{B}]\,\rangle,$$

where

$$\langle\,[\hat{A},\hat{B}]\,\rangle \equiv \langle\hat{A}\hat{B}\rangle - \langle\hat{B}\hat{A}\rangle \qquad \text{(IV-103)}$$

is the commutator of the two operators.

**d)** As a result, Eq. (IV-102) becomes

$$\boxed{\sigma_A^2\sigma_B^2 \geq \left(\frac{1}{2i}\langle[\hat{A},\hat{B}]\rangle\right)^2 .} \qquad \text{(IV-104)}$$

This is the uncertainty principle in its most general form.

**i)** You may be asking, *isn't the right-hand side of the equation negative*? The answer is <u>no</u> since the commutator carries its own factor of $i$ and the two cancel out.

**ii)** For example, suppose the first observable is position ($\hat{A} = x$) and the second is momentum ($\hat{B} = (\hbar/i)d/dx$). To determine the commutator, we use an arbitrary "test" function $f(x)$:

$$
\begin{aligned}
[\hat{x},\hat{p}]\,f &= x\frac{\hbar}{i}\frac{d}{dx}(f) - \frac{\hbar}{i}\frac{d}{dx}(xf) \\
&= \frac{\hbar}{i}\left[x\frac{df}{dx} - \left(f + x\frac{df}{dx}\right)\right] = i\hbar f \ ,
\end{aligned}
$$

so

$$\boxed{[\hat{x}, \hat{p}] = i\hbar\ .} \tag{IV-105}$$

Accordingly,

$$\sigma_x^2 \sigma_p^2 \geq \left(\frac{1}{2i}\, i\hbar\right)^2 = \left(\frac{\hbar}{2}\right)^2 ,$$

or, since standard deviations are by their nature positive,

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}\ . \tag{IV-106}$$

**e)** That proves that the original Heisenberg uncertainty principle, but we now see that it is just one application of a far more general theorem:

**i)** There will be an "uncertainty principle" for *any pair of observables whose corresponding operators do not commute.*

**ii)** We call these **incompatible observables**.

**iii)** Incompatible observables do not have shared eigenvectors — at least, they cannot have a *complete set* of common eigenvectors. Matrices representing incompatible observables cannot be simultaneously diagonalized (that is, they cannot both be brought to diagonal form by the *same* similarity transformation.)

**iv)** On the other hand, *compatible* observables (whose operators *do* commute) share a complete set of eigenvectors, and the corresponding matrices *can* be simultaneously diagonalized.

**Example IV–9.** Prove the famous "Luttermoser uncertainty principle," relating the uncertainty in position ($A = x$) to the uncertainty in energy ($B = p^2/2m + V = H$):

$$\sigma_x \sigma_H \geq \frac{\hbar}{2m} |\langle p \rangle| .$$

For stationary states this does not tell you much — why not?

**Solution:**

$$\left[ x, \frac{p^2}{2m} + V \right] = \frac{1}{2m} [x, p^2] + [x, V] .$$

Then from Eq. (IV-42),

$$[x, p^2] = xp^2 - p^2 x = xp^2 - pxp + pxp - p^2 x = [x, p]p + p[x, p] \ .$$

By making use of Eq. (IV-105), we get

$$\begin{aligned} [x, p^2] &= (i\hbar)p + p(i\hbar) = 2i\hbar p \ , \quad \text{and} \\ [x, V] &= xV - Vx = xV - xV = 0 \ . \end{aligned}$$

So,

$$\left[ x, \frac{p^2}{2m} + V \right] = \frac{1}{2m} 2i\hbar p = i\hbar p/m .$$

Now from Eq. (IV-104),

$$\begin{aligned} \sigma_x^2 \sigma_H^2 &\geq \left( \frac{1}{2i} \langle [\hat{x}, \hat{H}] \rangle \right)^2 = \left( \frac{1}{2i} \left\langle \left[ x, \frac{p^2}{2m} + V \right] \right\rangle \right)^2 \\ &\geq \left( \frac{1}{2i} \frac{i\hbar}{m} \langle p \rangle \right)^2 = \left( \frac{\hbar}{2m} \langle p \rangle \right)^2 , \end{aligned}$$

or

$$\sigma_x \sigma_H \geq \frac{\hbar}{2m} |\langle p \rangle| . \quad \text{Q.E.D.}$$

For stationary states, $\sigma_H = 0$ and $\langle p \rangle = 0$, so the "Luttermoser uncertainty relation" just says $0 \geq 0$ for stationary states $\Longrightarrow$ hence tells us nothing.

**2. The Energy-Time Uncertainty Principle.**

**a)** Compute the time derivative of the expectation value of some observable $Q(x, p, t)$:

$$\frac{d}{dt}\langle Q\rangle = \frac{d}{dt}\langle\Psi|\hat{Q}|\Psi\rangle = \langle\frac{\partial\Psi}{\partial t}|\hat{Q}|\Psi\rangle + \langle\Psi|\frac{\partial\hat{Q}}{\partial t}|\Psi\rangle + \langle\Psi|\hat{Q}|\frac{\partial\Psi}{\partial t}\rangle \ .$$

**b)** Now the Schrödinger equation says

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi \ ,$$

where $H = p^2/2m + V$ is the Hamiltonian, so substituting $(1/i\hbar)\hat{H}\Psi$ for the time derivatives of the wave function in the equation above, our terms become

$$\begin{aligned}
\langle\frac{\partial\Psi}{\partial t}|\hat{Q}|\Psi\rangle &= \langle\frac{1}{i\hbar}\hat{H}\Psi|\hat{Q}|\Psi\rangle = -\frac{1}{i\hbar}\langle\hat{H}\Psi|\hat{Q}|\Psi\rangle \\
\langle\Psi|\frac{\partial\hat{Q}}{\partial t}|\Psi\rangle &= \langle\frac{\partial\hat{Q}}{\partial t}\rangle \\
\langle\Psi|\hat{Q}|\frac{\partial\Psi}{\partial t}\rangle &= \langle\Psi|\hat{Q}|\frac{1}{i\hbar}\hat{H}\Psi\rangle = \frac{1}{i\hbar}\langle\Psi|\hat{Q}|\hat{H}\Psi\rangle = \frac{1}{i\hbar}\langle\Psi|\hat{Q}\hat{H}|\Psi\rangle \ ,
\end{aligned}$$

so

$$\frac{d}{dt}\langle Q\rangle = -\frac{1}{i\hbar}\langle\hat{H}\Psi|\hat{Q}\Psi\rangle + \frac{1}{i\hbar}\langle\Psi|\hat{Q}\hat{H}|\Psi\rangle + \langle\frac{\partial\hat{Q}}{\partial t}\rangle \ .$$

**c)** But $\hat{H}$ is Hermitian, so $\langle\hat{H}\Psi|\hat{Q}|\Psi\rangle = \langle\Psi|\hat{H}\hat{Q}|\Psi\rangle$, thus

$$\begin{aligned}
\frac{d}{dt}\langle Q\rangle &= -\frac{1}{i\hbar}\langle\Psi|\hat{H}\hat{Q}|\Psi\rangle + \frac{1}{i\hbar}\langle\Psi|\hat{Q}\hat{H}|\Psi\rangle + \langle\frac{\partial\hat{Q}}{\partial t}\rangle \\
&= \frac{i}{\hbar}\langle\Psi|\hat{H}\hat{Q}|\Psi\rangle - \frac{i}{\hbar}\langle\Psi|\hat{Q}\hat{H}|\Psi\rangle + \langle\frac{\partial\hat{Q}}{\partial t}\rangle \\
&= \frac{i}{\hbar}\langle\Psi|\left[\hat{H},\hat{Q}\right]|\Psi\rangle + \langle\frac{\partial\hat{Q}}{\partial t}\rangle
\end{aligned}$$

or

$$\boxed{\frac{d}{dt}\langle Q\rangle = \frac{i}{\hbar}\langle[\hat{H},\hat{Q}]\rangle + \langle\frac{\partial\hat{Q}}{\partial t}\rangle \ .} \qquad \text{(IV-107)}$$

**i)** If the operator $\hat{Q}$ does not depend explicity upon $t$, the rate of change of the expectation value of the observable $Q$ is determined by the commutator of the operator with the Hamiltonian.

**ii)** If $\hat{Q}$ *commutes* with $\hat{H}$, then $\langle Q \rangle$ is constant in time, and in this sense, $Q$ is a *conserved* quantity.

**d)** Let us chose $A = H$ and $B = Q$ in the generalized uncertainty principle of Eq. (IV-104) and assume that $Q$ does not depend explicitly on $t$, then

$$\sigma_H^2 \sigma_Q^2 \geq \left( \frac{1}{2i} \langle [\hat{H}, \hat{Q}] \rangle \right)^2 = \left( \frac{1}{2i} \frac{\hbar}{i} \frac{d\langle Q \rangle}{dt} \right)^2 = \left( \frac{\hbar}{2} \right)^2 \left( \frac{d\langle Q \rangle}{dt} \right)^2 .$$

**e)** Taking the square root of both sides gives

$$\sigma_H \sigma_Q \geq \frac{\hbar}{2} \left| \frac{d\langle Q \rangle}{dt} \right| . \qquad \text{(IV-108)}$$

**i)** Let's define $\Delta E \equiv \sigma_H$ (with $\Delta$ as the usual sloppy notation for standard deviation), and

$$\Delta t \equiv \frac{\sigma_Q}{|d\langle Q \rangle / dt|} . \qquad \text{(IV-109)}$$

**ii)** Then

$$\Delta E \Delta t \geq \frac{\hbar}{2} , \qquad \text{(IV-110)}$$

which is the energy-time uncertainty relation we wrote down in Eq. (I-5) using the special theory of relativity.

**iii)** Note that the definition of $\Delta t$ here is the *amount of time it takes the expectation value of $Q$ to change by one standard deviation.*

**f)** In particular, $\Delta t$ depends entirely on what observable ($Q$) you care to look at — the change might be rapid for one observable and slow for another.

**g)** But if $\Delta E$ is small, then the rate of change of *all* observables must be very gradual, and conversely, if *any* observable changes rapidly, the "uncertainty" in the energy must be large.

**h)** In atomic physics, an electron will stay in the ground state forever $\Delta t \to \infty$ unless a passing photon interacts with it. As such, $\Delta E = 0$ for the ground state $\Longrightarrow$ the ground state is infinitely "sharp." Meanwhile, an electron will stay excited for a short time ($\Delta t = 10^{-8}$ s for a resonance transition). As such, such an excited state will have a "natural width" of at least $\Delta E = (\hbar/2)/\Delta t = 5.273 \times 10^{-27}$ J $= 3.29 \times 10^{-8}$ eV.