Bra, Ket, Dirac, and all that . . .

An informal self-study guide for beginning students of quantum mechanics

Evening MS in Physics Project Report

by

David DeBruyne

Table of Contents

Chapter 1

Who ever heard of rose petal soup? The soup de jour is rose petal soup. What kind of place is this? This is the place the client wanted the meet...so here we are...and no clam chowder on the menu...A faceless gray suit appeared in front of him. It sat down without a word and placed a slip of paper on the table. "Please enjoy your meal at the expense of our employer," were all the words from both men. He reached his hand to cover the paper and left it on the table under his palm while matching the gaze of his new non-acquaintance. The stare-down was worthy of two ten-year-old boys. The gray suit rose and disappeared. The uneasy feeling of unanticipated confrontation was balanced by the thought "Well..., I need the work." Lifting his hand unsteadily, he absorbed the large, bold words without picking the paper up. "What did de Broglie know?"

The Mathematics of Quantum Mechanics

Part 1, Objects and Arithmetic

The first part of chapter 1 introduces the objects and "arithmetic" of quantum mechanical calculation. You should finish this part with an appreciation of **scalars**, **vectors**, and **matrix operators**; and operations using these objects. You should understand the environment known as a **complex linear vector space** within which these objects and operations are described. You should leave part 1 with firm concepts of the process of **normalization**, and the properties of **linear independence** and **orthonormality**. The property of orthonormality is a practical necessity to any **basis** useful for quantum mechanical calculation. Finally, you should start part 2 with an understanding of **commutivity**. You should know which objects **commute** and which objects do not commute.

---

1. Given that $\alpha = 3+4i, \quad \beta = 2-3i\,,$ and $\gamma = 5e^{2+5i}\,,$ find

(a) $\alpha+\beta$ and $\alpha-\beta\,,$

(b) the complex conjugates of $\alpha,\ \beta\,,$ and $\gamma\,,$

(c) $\alpha\,\beta$ and $\beta\,\alpha\,,$ and compare the products. Then find

(d) $\alpha\,\alpha^*\,,\ \beta\,\beta^*\,,$ and $\gamma\,\gamma^*\,,$

(e) and, $\alpha\div\beta$ and $\beta\div\alpha\,.$

A brief review of operations with complex numbers and complex conjugates is appropriate because the components and elements in the vectors and operators used to describe quantum mechanical phenomena are intrinsically complex. Complex numbers are added or subtracted by adding or subtracting real parts and imaginary parts. A complex conjugate is formed by multiplying the imaginary part by -1. Usually, the symbol "*" denotes a complex conjugate, *e.g.*, α^* denotes the complex conjugate of α. Multiplication of complex numbers is a form of binomial multiplication. Part (c) is a numerical example of **commutivity**. If two objects **commute**, their product is the same regardless of the order of multiplication, *i.e.*, if $a\cdot b = b\cdot a\,,$ then a and b are said to commute. You will show that the product of complex conjugates is a real number in the next problem. This fact is used to divide complex numbers. “Division” of complex numbers is a form of rationalization. Multiplying the quotient desired by 1 in the form of the complex conjugate of the divisor divided by itself, *e.g.* $\dfrac{\alpha}{\beta} = \dfrac{\alpha}{\beta}\dfrac{\beta^*}{\beta^*}$, yields a scalar of the form $a + bi$, though the real and imaginary parts are often fractional.

(a)
$$\alpha + \beta = 3 + 4i + 2 - 3i = (3+2) + (4-3)i = 5 + i\,,$$
$$\alpha - \beta = 3 + 4i - (2 - 3i) = (3-2) + (4+3)i = 1 + 7i\,.$$

(b) $\alpha = 3+4i, \quad \beta = 2-3i, \quad \gamma = 5e^{2+5i} \quad \Rightarrow \quad \alpha^* = 3-4i, \quad \beta^* = 2+3i, \quad \text{and} \quad \gamma^* = 5e^{2-5i}.$

(c) $\alpha\,\beta = (3+4i)(2-3i) = 3{\cdot}2 + 3(-3i) + 4i{\cdot}2 + 4i(-3i) = 6 - 9i + 8i + 12 = 18 - i\,.$

$$\beta\,\alpha = (2-3i)(3+4i) = 2\cdot 3 + 2\cdot 4i - 3i\cdot 3 - 3i\cdot 4i = 6 + 8i - 9i + 12 = 18 - i\,.$$

These complex numbers commute. All complex numbers commute as you will show in problem 2.

(d) $\alpha\,\alpha^* = (3+4i)(3-4i) = 3{\cdot}3 + 3(-4i) + 4i{\cdot}3 + 4i(-4i) = 9 - 12i + 12i + 16 = 25\,.$

$$\beta\,\beta^* = (2-3i)(2+3i) = 2\cdot 2 + 2\cdot 3i - 3i\cdot 2 - 3i\cdot 3i = 4 + 6i - 6i + 9 = 13$$
$$\gamma\,\gamma^* = 5e^{2+5i}\cdot 5e^{2-5i} = 25e^{4+0i} = 25e^4\,.$$

These products of complex conjugates are real numbers. The products of complex conjugates are always real numbers which is another general result you will show in problem 2.

(e) $\dfrac{\alpha}{\beta} = \dfrac{3+4i}{2-3i} = \dfrac{3+4i}{2-3i}\dfrac{2+3i}{2+3i} = \dfrac{3\cdot 2 + 3\cdot 3i + 4i\cdot 2 + 4i\cdot 3i}{13} = \dfrac{6 - 12 + 9i + 8i}{13} = -\dfrac{6}{13} + \dfrac{17}{13}i\,.$

$$\frac{\beta}{\alpha} = \frac{2-3i}{3+4i} = \frac{2-3i}{3+4i}\frac{3-4i}{3-4i} = \frac{2\cdot 3 - 2\cdot 4i - 3i\cdot 3 + 3i\cdot 4i}{25} = \frac{6 - 12 - 8i - 9i}{25} = -\frac{6}{25} - \frac{17}{25}i\,.$$

Postscript: A **scalar** is defined as a tensor of zero order, and is invariant under rotation. An informal description is that any individual number is a scalar. A scalar may be real, imaginary, or complex. The first problem should reinforce the fact that the sum, difference, product, and quotient of two complex scalars is another scalar.

Observable and measurable quantities such as mass, charge, or energy are described using real scalars. Since the components and elements of vectors and operators used to describe quantum mechanical phenomena are intrinsically complex, calculations that predict measurable quantities must possess a mechanism that yields real numbers. This mechanism is embedded in the fact that the product of complex conjugates is a real number. Part (d) of the first problem provides three numerical examples of this fact.

---

2. Show in general that

(a) complex numbers commute under the operation of multiplication,

(b) the product of complex conjugates is the sum of the squares of the real and imaginary parts,

(c) and the product of complex conjugates is a real number.

Vectors and operators do not commute in general. Scalars do commute in general. This problem should fix this fact and two other general properties of complex numbers under the operation of multiplication. Assume general forms of complex numbers such as $a+bi$ and $c+di$ where a, b, c, and d are real numbers, then do the indicated multiplications.

(a) $$(a+bi)(c+di) = a\cdot c + a\cdot di + bi\cdot c + bi\cdot di = ac - bd + (ad+cb)i\,, \quad \text{and}$$

$$(c+di)(a+bi) = c\cdot a + c\cdot bi + di\cdot a + di\cdot bi = ac - bd + (ad+cb)i\,,$$

therefore $(a+bi)(c+di) = (c+di)(a+bi)$ and complex scalars commute in general. (b) $(a+bi)(a-bi) = a\cdot a + a(-bi) + bi\cdot a + bi(-bi) = a^2 - a\cancel{b}i + a\cancel{b}i + b^2 = a^2 + b^2$.

(c) Since a and b are real numbers, a^2 and b^2 are real numbers, and their sum is a real number.

---

3. If $\alpha = 3+4i$ and $|v> = \begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix}$, find $\alpha\,|v>$.

The first postulate of quantum mechanics states that a system is represented by a **state vector**. Operations with vectors are therefore fundamental to quantum mechanical calculations. Three types of multiplication are possible with vectors. This problem addresses **scalar multiplication** of a vector. Vector-vector, operator-vector, and vector-operator multiplication are the other types of vector multiplication that we will encounter in part 1. To multiply a vector by a scalar, multiply each component of the vector by the scalar. A scalar times a vector is another vector. Symbolically,

$$\alpha\,|v> = \alpha \begin{pmatrix} \beta \\ \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} \alpha\,\beta \\ \alpha\,\gamma \\ \alpha\,\delta \end{pmatrix}.$$

$$\begin{aligned} \alpha\,|v> &= (3+4i)\begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix} = \begin{pmatrix} (3+4i)(2+i) \\ (3+4i)(1-3i) \\ (3+4i)(3-2i) \end{pmatrix} \\ &= \begin{pmatrix} 3\cdot 2 + 3\cdot i + 4i\cdot 2 + 4i\cdot i \\ 3\cdot 1 + 3(-3i) + 4i\cdot 1 + 4i(-3i) \\ 3\cdot 3 + 3(-2i) + 4i\cdot 3 + 4i(-2i) \end{pmatrix} = \begin{pmatrix} 6+3i+8i-4 \\ 3-9i+4i+12 \\ 9-6i+12i+8 \end{pmatrix} = \begin{pmatrix} 2+11i \\ 15-5i \\ 17+6i \end{pmatrix}. \end{aligned}$$

---

4. If $\gamma = 2i$, and $\mathcal{B} = \begin{pmatrix} -2-2i & 3-i & 1+i \\ -3+2i & -2+i & 3-i \\ -2i & -3-3i & 4 \end{pmatrix}$, find $\gamma\,\mathcal{B}$.

An **operator** or a **matrix operator** contains the same information as a differential operator. Dirac demonstrated the equivalence of the Schrodinger differential equation and Heisenberg matrix formulations of quantum mechanics. Both types of operators have advantages in given circumstances, though you will see more matrix operators than differential operators in part 1 because matrix operators provide clarity to the **eigenvalue/eigenvector** problem and favor the economy contained in **Dirac notation**.

To multiply a matrix operator by a scalar, multiply each element of the operator by the scalar. A scalar times an operator is another operator.

$$\alpha\,\mathcal{A} = \alpha \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} \alpha\,a_{11} & \alpha\,a_{12} & \alpha\,a_{13} \\ \alpha\,a_{21} & \alpha\,a_{22} & \alpha\,a_{23} \\ \alpha\,a_{31} & \alpha\,a_{32} & \alpha\,a_{33} \end{pmatrix}.$$

$$\gamma\,\mathcal{B} = 2i \begin{pmatrix} -2-2i & 3-i & 1+i \\ -3+2i & -2+i & 3-i \\ -2i & -3-3i & 4 \end{pmatrix} = \begin{pmatrix} 2i(-2-2i) & 2i(3-i) & 2i(1+i) \\ 2i(-3+2i) & 2i(-2+i) & 2i(3-i) \\ 2i(-2i) & 2i(-3-3i) & 2i(4) \end{pmatrix}$$
$$= \begin{pmatrix} -4i+4 & 6i+2 & 2i-2 \\ -6i-4 & -4i-2 & 6i+2 \\ 4 & -6i+6 & 8i \end{pmatrix} = \begin{pmatrix} 4-4i & 2+6i & -2+2i \\ -4-6i & -2-4i & 2+6i \\ 4 & 6-6i & 8i \end{pmatrix}.$$

---

5. Find (a) $|\,v\!> + \,|\,w\!>$ and (b) $|\,v\!> - \,|\,w\!>$ given that

$$|\,v\!> = \begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix} \quad \text{and} \quad |\,w\!> = \begin{pmatrix} -1-2i \\ -2-2i \\ 1+i \end{pmatrix}.$$

Add or subtract two vectors by adding or subtracting corresponding components. Symbolically,

$$|\,a\!> = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{and} \quad |\,b\!> = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad \Rightarrow \quad |\,a\!> \pm \,|\,b\!> = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \pm \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 \pm b_1 \\ a_2 \pm b_2 \\ a_3 \pm b_3 \end{pmatrix}.$$

(a) $$\begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix} + \begin{pmatrix} -1-2i \\ -2-2i \\ 1+i \end{pmatrix} = \begin{pmatrix} 2+i-1-2i \\ 1-3i-2-2i \\ 3-2i+1+i \end{pmatrix} = \begin{pmatrix} 2-1+i-2i \\ 1-2-3i-2i \\ 3+1-2i+i \end{pmatrix} = \begin{pmatrix} 1-i \\ -1-5i \\ 4-i \end{pmatrix}.$$

(b) $$\begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix} - \begin{pmatrix} -1-2i \\ -2-2i \\ 1+i \end{pmatrix} = \begin{pmatrix} 2+i+1+2i \\ 1-3i+2+2i \\ 3-2i-1-i \end{pmatrix} = \begin{pmatrix} 2+1+i+2i \\ 1+2-3i+2i \\ 3-1-2i-i \end{pmatrix} = \begin{pmatrix} 3+3i \\ 3-i \\ 2-3i \end{pmatrix}.$$

Postscript: The two vectors being added need to have the same number of components. Addition for vectors possessing a dissimilar number of components is not defined. For instance, addition of a scalar and a vector is undefined. The sum of two vectors is another vector.

---

6. Find (a) $\mathcal{A} + \mathcal{B}$ and (b) $\mathcal{A} - \mathcal{B}$ given that

$$\mathcal{A} = \begin{pmatrix} -4+2i & 2+i & -1-3i \\ 6 & 2+3i & -3i \\ -3-i & 2-3i & -5 \end{pmatrix}, \quad \text{and} \quad \mathcal{B} = \begin{pmatrix} -2-2i & 3-i & 1+i \\ -3+2i & -2+i & 3-i \\ -2i & -3-3i & 4 \end{pmatrix}$$

The sum/difference of two matrix operators is the sum/difference of the corresponding elements.

(a) $$\begin{pmatrix} -4-2+2i-2i & 2+3+i-i & -1+1-3i+i \\ 6-3+2i & 2-2+3i+i & 3-3i-i \\ -3-i-2i & 2-3-3i-3i & -5+4 \end{pmatrix} = \begin{pmatrix} -6 & 5 & -2i \\ 3+2i & 4i & 3-4i \\ -3-3i & -1-6i & -1 \end{pmatrix}.$$

(b) $$\begin{pmatrix} -4+2+2i+2i & 2-3+i+i & -1-1-3i-i \\ 6+3-2i & 2+2+3i-i & -3-3i+i \\ -3-i+2i & 2+3-3i+3i & -5-4 \end{pmatrix} = \begin{pmatrix} -2+4i & -1+2i & -2-4i \\ 9-2i & 4+2i & -3-2i \\ -3+i & 5 & -9 \end{pmatrix}.$$

Postscript: Operators useful for quantum mechanical calculation are square matrices, *i.e.*, they have the same number of elements in columns as in rows. Operator addition is undefined unless the operators have identical configurations. You cannot add a 2 X 2 matrix to a 3 X 3 matrix, for instance. The sum of two matrix operators is another matrix operator. Addition of a scalar and an operator, and addition of a vector and an operator are undefined.

---

7. What are (a) $<v\,|$, (b) $<w\,|$, (c) $<v\,| + <w\,|$, and (d) $<v\,| - <w\,|$,

$$\text{given that} \quad |\,v> = \begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix} \quad \text{and} \quad |\,w> = \begin{pmatrix} -1-2i \\ -2-2i \\ 1+i \end{pmatrix}.$$

Dirac called a column vector a **ket**, denoted $|\,u>$, for instance. He called a row vector a **bra** and denoted it $<u\,|$, for example. The rationale behind these labels is explained in the next problem. A bra is "formed" from a ket by forming a row vector of the complex conjugates of the components of the column vector. There is a bra $<u\,|$ whose components correspond to the complex conjugates for every ket $|\,u>$ and a ket $|\,u>$ whose components correspond to the complex conjugates for every bra $<u\,|$. For example, to the ket

$$|\,u> = \begin{pmatrix} 2+5i \\ 7-3i \\ -4 \\ 6i \end{pmatrix} \quad \text{there corresponds a bra} \quad <u\,| = \big(2-5i,\ 7+3i,\ -4,\ -6i\big),$$

and to the bra $\quad < u \,| = \big(-4 - 3i,\ 8i,\ -3 \big)\quad$ there corresponds a ket $\quad | \, u > = \begin{pmatrix} -4 + 3i \\ -8i \\ -3 \end{pmatrix}.$

The **state vectors** of quantum mechanics are kets so kets are the more frequently encountered type of vector and we "form" the corresponding bra when it is needed. This practical procedure is technically incorrect. A technically correct picture is that the corresponding bra already exists among all other possible bras and we simply pick it from all other possible bras when it is needed. We will soon introduce the idea of **Hilbert space**, in which this technically correct concept is founded. There is no preference of kets over bras or bras over kets. You can form bras from kets or kets from bras as a practical matter, though we will have more occasion to form bras from kets because state vectors are kets.

Arithmetic operations previously described for kets apply to bras, where only the format is different since bras are vectors in a different format than kets. Adding a row vector to another row vector results in a row vector, for instance.

(a) $\quad < v \,| = \big(2 - i,\ 1 + 3i,\ 3 + 2i \big)$ $\qquad$ (b) $\quad < w \,| = \big(-1 + 2i,\ -2 + 2i,\ 1 - i \big)$

(c) $\quad < v \,| + < w \,| = \big(2 - 1 - i + 2i,\ \ 1 - 2 + 3i + 2i,\ \ 3 + 1 + 2i - i \big) = \big(1 + i,\ -1 + 5i,\ 4 + i \big)$

(d) $\quad < v \,| - < w \,| = \big(2 + 1 - i - 2i,\ \ 1 + 2 + 3i - 2i,\ \ 3 - 1 + 2i + i \big) = \big(3 - 3i,\ 3 + i,\ 2 + 3i \big)$

Postscript: A bra is the vector analog of the complex conjugate of a scalar.

---

8. Using $\,| \, v >$ and $\,| \, w >$ from problem 7, find $\quad$ (a) $< v \,| \, v >$, $\quad$ (b) $< w \,| \, w >$, $\quad$ (c) $< v \,| \, w >$, and $\quad$ (d) $< w \,| \, v >$.

Vector-vector multiplication is defined only for a row vector and a column vector with the same number of components. Vector-vector multiplication is not defined for a row vector and a column vector with different numbers of components, for two row vectors, or for two column vectors. The order of the vectors is important. Here we address only the vector-vector product where the row vector is on the left and the column vector is on the right. This order of multiplication is known as an **inner product**. An inner product is a generalization of the dot product of classical mechanics.

The procedure is to sum the product of corresponding components, meaning

$$\big(a_1,\ a_2 \big) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1 \cdot b_1 + a_2 \cdot b_2 \qquad \text{or} \qquad \big(a_1,\ a_2,\ a_3 \big) \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1 \cdot b_1 + a_2 \cdot b_2 + a_3 \cdot b_3 \,.$$

The procedure generalizes to arbitrary or infinite dimension. An inner product is a scalar.

$$\begin{aligned} \big(5 + 2i,\ -3,\ 2i,\ 1 - 4i \big) \begin{pmatrix} -3 - 2i \\ -4i \\ 3 \\ 3 + 3i \end{pmatrix} &= (5 + 2i)(-3 - 2i) + (-3)(-4i) + 2i \cdot 3 + (1 - 4i)(3 + 3i) \\ &= (-15 - 10i - 6i + 4) + (12i) + (6i) + (3 + 3i - 12i + 12) \\ &= -11 - 16i + 12i + 6i + 15 - 9i = 4 - 7i. \end{aligned}$$

An inner product is a bra times a ket. If we were to denote the vectors of the last example as

$$< v \,| \; = \big(5+2i,\; -3,\; 2i,\; 1-4i\big) \qquad \text{and} \qquad |\, w> \; = \begin{pmatrix} -3-2i \\ -4i \\ 3 \\ 3+3i \end{pmatrix}, \quad \text{then} \quad < v \,|\, w> \; = 4-7i\,.$$

An expression with a bra on the left and a ket on the right is known as a **bracket** or **braket**. The inner product is the first example of a braket. One vertical slash serves to separate the bra and the ket for an inner product. Parts (a) and (b) are easier if you remember that the product of complex conjugates is the sum of the squares of the real and the imaginary parts.

$$\text{(a)} \qquad < v \,|\, v> \; = \big(2-i,\; 1+3i,\; 3+2i\big) \begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix} = 4+1+1+9+9+4 = 28\,.$$

$$\text{(b)} \quad < w \,|\, w> \; = \big(-1+2i,\; -2+2i,\; 1-i\big) \begin{pmatrix} -1-2i \\ -2-2i \\ 1+i \end{pmatrix} = 1+4+4+4+1+1 = 15\,.$$

$$\text{(c)} \;\; < v \,|\, w> \, = \big(2-i,\; 1+3i,\; 3+2i\big) \begin{pmatrix} -1-2i \\ -2-2i \\ 1+i \end{pmatrix} = (2-i)(-1-2i) + (1+3i)(-2-2i) + (3+2i)(1+i)$$

$$= (-2-4i+i-2) + (-2-2i-6i+6) + (3+3i+2i-2) = -4-3i+4-8i+1+5i = 1-6i\,.$$

$$\text{(d)} \;\; < w \,|\, v> \, = \big(-1+2i,\; -2+2i,\; 1-i\big) \begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix} = (-1+2i)(2+i) + (-2+2i)(1-3i) + (1-i)(3-2i)$$

$$= (-2-i+4i-2) + (-2+6i+2i+6) + (3-2i-3i-2) = -4+3i+4+8i+1-5i = 1+6i\,.$$

Postscript: Notice that the inner product of the example and parts (a) through (d) are all scalars. You want to remember that an inner product is a scalar. Notice also that the results of parts (a) and (b) are real scalars. You may also want to notice that the results of parts (c) and (d) are complex conjugates. Also, the inner product is only the first example of a braket. Another object, for instance an operator, can be placed between the bra and the ket in a braket.

---

9. Find the norm of each ket given in problem 7.

The length of a classical vector, a vector in three dimensions with real components, is the square root of its dot product. The **norm** of a vector is a generalization of this concept of length to both an unspecified number of dimensions and complex components. The norm of a vector is the square root of the inner product of the ket and the corresponding bra. A norm is denoted by placing a vertical slash on each side of the vector symbol. For example,

$$|\,u\!> = \begin{pmatrix} -3-2i \\ -4i \\ 3 \\ 3+3i \end{pmatrix} \quad \Rightarrow \quad <\!u\,|\,u\!> = \big(-3+2i,\ 4i,\ 3,\ 3-3i\big)\begin{pmatrix} -3-2i \\ -4i \\ 3 \\ 3+3i \end{pmatrix}$$

$$= 9+4+16+9+9+9 = 56 \quad \Rightarrow \quad \Big|\,|\,u\!>\Big| = \sqrt{56} = 2\sqrt{14}\,.$$

$<\!v\,|\,v\!> = 28$ from problem 8, so $\big|\,|\,v\!>\big| = \sqrt{<\!v\,|\,v\!>} = \sqrt{28} = 2\sqrt{7}$. Similarly, $<\!w\,|\,w\!> = 15$ from problem 8, so $\big|\,|\,w\!>\big| = \sqrt{<\!w\,|\,w\!>} = \sqrt{15}$.

---

10. Normalize the vectors $|\,v\!>$ and $|\,w\!>$ given that

$$|\,v\!> = \begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix} \quad \text{and} \quad |\,w\!> = \begin{pmatrix} 4-i \\ -3+2i \end{pmatrix}.$$

A vector of norm 1 is said to be **normalized**. The process of normalization is geometrically equivalent to adjusting the norm of the vector to unit "length" while preserving the "direction." A vector is normalized by multiplying it by a scalar. The problem is to find this scale factor. If $|\,v'\!>$ is the vector to be normalized, attach an unknown scale factor A called a **normalization constant** to form $\mathrm{A}\,|\,v'\!>$. Form the corresponding bra remembering to conjugate the normalization constant. Set the inner product of these two vectors equal to one and solve for the normalization constant.

$$<\!v'\,|\,\mathrm{A}^*\,\mathrm{A}\,|\,v'\!> = 1 \quad \Rightarrow \quad |\,\mathrm{A}\,|^2 <\!v'\,|\,v'\!> = 1 \quad \Rightarrow \quad |\,\mathrm{A}\,|^2 = \frac{1}{<\!v'\,|\,v'\!>} \quad \Rightarrow \quad \mathrm{A} = \frac{1}{\sqrt{<\!v'\,|\,v'\!>}}\,.$$

The normalized vector is $|\,v\!> = \mathrm{A}\,|\,v'\!> = \dfrac{1}{\sqrt{<\!v'\,|\,v'\!>}}\,|\,v'\!>$. The convention is to use the principal value of the square root. Another convention is to refer to any vector with the same direction by the same symbol so the primes are not usually included. The reason for this is the probabilistic interpretation of the state vector that we will encounter in chapter 2. Any vector with the same "direction" is equivalent for the purposes of quantum mechanics. The process of normalization is similar to the process of forming unit vectors in introductory mechanics.

$$\begin{aligned} <\!v\,|\,\mathrm{A}^*\,\mathrm{A}\,|\,v\!> &= (2-i,\ 1+3i,\ 3+2i)\,\mathrm{A}^*\,\mathrm{A}\begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix} \\ &= |\,\mathrm{A}\,|^2\Big((4+1)+(1+9)+(9+4)\Big) \\ &= |\,\mathrm{A}\,|^2\Big(5+10+13\Big) = |\,\mathrm{A}\,|^2\,28 = 1 \end{aligned}$$

$$\Rightarrow \quad \mathrm{A} = \frac{1}{\sqrt{28}} \quad \Rightarrow \quad |\,w\!> = \frac{1}{\sqrt{28}} \begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix}$$

$$<\!w\,|\,\mathrm{A}^*\,\mathrm{A}\,|\,w\!> = \big(4+i,\ -3-2i\big)\,\mathrm{A}^*\,\mathrm{A} \begin{pmatrix} 4-i \\ -3+2i \end{pmatrix} = |\,\mathrm{A}\,|^2\big(16+1+9+4\big) = |\,\mathrm{A}\,|^2\, 30 = 1$$

$$\Rightarrow \quad \mathrm{A} = \frac{1}{\sqrt{30}} \quad \Rightarrow \quad |\,w\!> = \frac{1}{\sqrt{30}} \begin{pmatrix} 4-i \\ -3+2i \end{pmatrix}.$$

Postscript: The **normalization condition** is $<v\,|\,v> = 1$. If $<v\,|\,v> \neq 1$, it is wise to normalize it before attempting further calculations. A vector that satisfies the normalization condition is the same as a vector whose components have absorbed the normalization constant, *i.e.*,

$$|\,v\!> = \mathrm{A}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \mathrm{A}\,a \\ \mathrm{A}\,b \\ \mathrm{A}\,c \end{pmatrix}, \quad \text{for instance. You should notice that}$$

$$|\,v\!> = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \text{and} \quad |\,v\!> = \mathrm{A}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \mathrm{A}\,a \\ \mathrm{A}\,b \\ \mathrm{A}\,c \end{pmatrix} \quad \text{are equivalent for the purposes of quantum}$$

mechanics. The same symbol $|\,v\!>$ is used to denote both normalized and unnormalized vectors.

---

11. Argue that the set of integers constitutes a linear vector space and that the set of real numbers between -1 and 1 does not constitute a linear vector space.

This problem is an introduction to a **complex linear vector space**. It should also introduce you to some notation appropriate to vector spaces. The fundamental concept is that of a **linear vector space** which is defined by the two properties of closure and the eight algebraic properties described below. A complex linear vector space is a linear vector space where the components of the vectors are complex numbers. This problem uses real numbers because they are the simplest. Scalars can be considered one dimensional vectors and integers are simply real scalars.

A **linear vector space** *is a set of vectors which is closed with respect to vector addition and scalar multiplication and obeys the axioms of addition and scalar multiplication in equations (3) through (10).*

Symbolically, if the linear vector space is denoted $\mathbf{C}$ and $|\,v\!>$ and $|\,w\!>$ are any vectors in the space, the properties of closure with respect to vector addition can be expressed

$$|\,v\!>,\ |\,w\!> \in \mathbf{C} \quad \Rightarrow \quad |\,v\!> + \ |\,w\!> \in \mathbf{C}, \tag{1}$$

and if α is any scalar, the property of closure with respect to scalar multiplication is

$$|\,v\!> \in \mathbf{C} \quad \Rightarrow \quad \alpha\,|\,v\!> \in \mathbf{C}. \tag{2}$$

A linear vector space needs to obey the axioms of equations (3) through (10). The definition of addition must satisfy commutivity in addition, equation (3); associativity in addition, equation (4); the existence of an additive identity, equation (5); and the existence of an additive inverse, equation (6). Symbolically,

$$|\,v\!> + \;|\,w\!> \;=\; |\,w\!> + \;|\,v\!>, \tag{3}$$

$$\Big(\,|\,v\!> + \;|\,w\!>\Big) + |\,z\!> \;=\; |\,w\!> + \Big(\,|\,v\!> + \;|\,z\!>\Big), \tag{4}$$

$$|\,v\!> + \;|\,0\!> \;=\; |\,v\!>, \tag{5}$$

$$|\,v\!> + \;|\,-v\!> \;=\; |\,0\!>\,. \tag{6}$$

The definition of scalar multiplication must satisfy the existence of a multiplicative identity, equation (7); distribution in vectors, equation (8); distribution in scalars, equation (9); and associativity in scalar multiplication, equation (10). Symbolically,

$$1\cdot|\,v\!> \;=\; |\,v\!>, \tag{7}$$

$$\alpha\Big(\,|\,v\!> + \;|\,w\!>\Big) = \alpha\,|\,v\!> + \,\alpha\,|\,w\!>, \tag{8}$$

$$\big(\alpha+\beta\big)\,|\,v\!> = \alpha\,|\,v\!> + \,\beta\,|\,v\!>, \tag{9}$$

$$\alpha\Big(\beta\,|\,v\!>\Big) = \alpha\,\beta\,|\,v\!>, \tag{10}$$

The set of numbers that are the components of the vectors that constitute the space is called a **field**. The numbers of interest to quantum mechanics are complex numbers. Quantum mechanics requires a linear vector space in the field of complex numbers, or a complex linear vector space. Complex linear vector spaces are subsets of linear vector spaces, and real numbers are a subset of the complex numbers, so sets of one dimensional vectors whose components are solely real numbers can be considered as long as the set satisfies conditions (1) through (10).

Since scalars can be regarded as one dimensional vectors, a simple example of a linear vector space is the set of integers. If we add or multiply any two integers, we attain another integer so the set is closed under addition and multiplication. Integers are commutative, associative, the set contains 0 so possesses an additive identity, and contains the negative numbers so also possesses additive inverses, thus satisfy all the required properties of addition. The set also satisfies the multiplicative properties. Since 1 is a member of the set, it possesses a multiplicative identity, integers are distributive (so satisfy the properties of equations (8) and (9)), and associative with respect to multiplication. All ten conditions are satisfied, so the set of integers is a linear vector space.

The inclusive set of numbers between -1 and 1 also does not satisfy the conditions required for a linear vector space. It is not closed under addition. For instance, $0.75 + 0.75 = 1.5$ is not a member of the set.

Postscript: Complex numbers are the field required for quantum mechanics calculations. Since we have defined addition in kets and scalar multiplication of kets, and we know the properties of complex numbers, the two properties of closure and the eight algebraic properties follow for vector addition and scalar multiplication for vectors with complex components. Of course, all the kets in any complex linear vector space must have the same number of components. For instance, all two dimensional kets are a linear vector space over the field of complex numbers. The set of all three dimensional kets is similarly a linear vector space over the field of complex numbers. The set that is the union of all two and three dimensional kets does not form a linear vector space because addition of a two and three dimensional ket is not defined, therefore, none of the additive properties of a linear vector space can be satisfied.

---

12. Consider the set of all entities of the form $<v\,| = (a,\ b,\ c)$, where a, b, and c are real numbers, and where vector addition and scalar multiplication are defined as follows:

$$(a,\ b,\ c) + (d,\ e,\ f) = (a+d,\ b+e,\ c+f)$$

$$\alpha\,(a,\ b,\ c) = (\alpha\, a,\ \alpha\, b,\ \alpha\, c)$$

where α is also a real number. Verify that this constitutes a linear vector space.

This problem is designed to develop your understanding of the definition of a linear vector space encountered in the previous problem.

You need to show that the two properties of closure, and the eight algebraic axioms are true for the rules of vector addition and scalar multiplication that are defined in the statement of the problem. You should first verify the two properties of closure using the fact that the sum or product of two real numbers will be another real number. Then, for example, to verify the axiom that scalar multiplication is distributive over vector addition, equation (8) from problem 11, you might write:

$$\begin{aligned}
\alpha\Big(<v\,| + <w\,|\Big) &= \alpha\Big((a,\ b,\ c) + (d,\ e,\ f)\Big)\\
&= \alpha\,(a+d,\ b+e,\ c+f)\\
&= \big(\alpha\,(a+d),\ \alpha\,(b+e),\ \alpha\,(c+f)\big)\\
&= (\alpha\, a + \alpha\, d,\ \alpha\, b + \alpha\, e,\ \alpha\, c + \alpha\, f)\\
&= (\alpha\, a,\ \alpha\, b,\ \alpha\, c) + (\alpha\, d,\ \alpha\, e,\ \alpha\, f)\\
&= \alpha\,(a,\ b,\ c) + \alpha\,(d,\ e,\ f)\\
&= \alpha <v\,| + \alpha <w\,|\,.
\end{aligned}$$

You should use the definitions in the statement of the problem forwards and backwards to verify the two properties of closure and other seven algebraic axioms analogously. We use the symbol **R** to denote the set of real numbers. You should also know that $|\,0>$ means zero appropriate to the space. For instance, $<0\,| = (0,\ 0,\ 0)$ for this problem.

To show that this is a linear vector space, we must show that the ten properties described in problem 11 are satisfied individually.

- Closure under vector addition, equation (1): Here,

$$<v\,| + <w\,| \ = (a,\ b,\ c) + (d,\ e,\ f) = (a+d,\ b+e,\ c+f)\,.$$

If $a, b, c, d, e, f \in \mathbf{R}$, then $a+d,\ b+e,\ c+f \in \mathbf{R}$, and the addition operation results in other elements in the field of real numbers. The resulting vector has three component that are real numbers so is a member of the three dimensional space. Therefore, this space is closed under vector addition.

- Closure under scalar multiplication, equation (2): $\alpha < v\,| = \alpha\,(a,\ b,\ c) = (\alpha\, a,\ \alpha\, b,\ \alpha\, c)$. If $\alpha \in \mathbf{R}$, and $a,\ b,\ c \in \mathbf{R}$, then the products $\alpha\, a,\ \alpha\, b,$ and $\alpha\, c \in \mathbf{R}$, and the operation of scalar multiplication results in other elements in the field of real numbers. The resulting vector has three elements that are real numbers, so is a member of the three dimensional space. Therefore, this space is closed under scalar multiplication.

- Commutivity in addition, equation (3):

$$\begin{aligned} < v\,| + < w\,| &= (a,\ b,\ c) + (d,\ e,\ f) \\ &= (a + d,\ b + e,\ c + f) \\ &= (d + a,\ e + b,\ f + c) \\ &= (d,\ e,\ f) + (a,\ b,\ c) \\ &= < w\,| + < v\,|, \end{aligned}$$

for all real numbers. Therefore, vector addition satisfies the property of commutivity in addition.

- Associativity in addition, equation (4):

$$\begin{aligned} \Big(< v\,| + < w\,| \Big) + < z\,| &= \Big((a,\ b,\ c) + (d,\ e,\ f) \Big) + (g,\ h,\ i) \\ &= (a + d,\ b + e,\ c + f) + (g,\ h,\ i) \\ &= (a + d + g,\ b + e + h,\ c + f + i) \\ &= (a,\ b,\ c) + (d + g,\ e + h,\ f + i) \\ &= (a,\ b,\ c)\ + \Big((d,\ e,\ f) + (g,\ h,\ i) \Big) \\ &= < v\,| + \Big(< w\,| + < z\,| \Big). \end{aligned}$$

Therefore, vector addition is associative.

- Existence of an additive identity, equation (5):

$$< v\,| + < 0\,| = (a,\ b,\ c) + (0,\ 0,\ 0) = (a + 0,\ b + 0,\ c + 0) = (a,\ b,\ c) = < v\,|,$$

therefore, an additive identity does exist.

- Existence of an additive inverse, equation (6):

$$< v\,| + < -v\,| = (a,\ b,\ c) + (a',\ b',\ c') = (a + a',\ b + b',\ c + c') = (0,\ 0,\ 0) = < 0\,|$$

if and only if $a' = -a,\ \ b' = -b,\ \ c' = -c \ \ \Rightarrow\ \ < -v\,| = (-a,\ -b,\ -c)$. Therefore, there is a unique additive inverse.

- Existence of a multiplicative identity, equation (7):

$$1 \cdot < v\,| = 1 \cdot (a,\ b,\ c) = (1 \cdot a,\ 1 \cdot b,\ 1 \cdot c) = (a,\ b,\ c) = < v\,|.$$

Therefore, a multiplicative identity does exist for scalar multiplication.

- Distribution in vectors, equation (8):

$$\begin{aligned}\alpha\Big(<v\,| + <w\,|\Big) &= \alpha\Big((a,\ b,\ c) + (d,\ e,\ f)\Big)\\ &= \alpha\,(a+d,\ b+e,\ c+f)\\ &= \Big(\alpha\,(a+d),\ \alpha\,(b+e),\ \alpha\,(c+f)\Big)\\ &= (\alpha\, a + \alpha\, d,\ \alpha\, b + \alpha\, e,\ \alpha\, c + \alpha\, f)\\ &= (\alpha\, a,\ \alpha\, b,\ \alpha\, c) + (\alpha\, d,\ \alpha\, e,\ \alpha\, f)\\ &= \alpha\,(a,\ b,\ c) + \alpha\,(d,\ e,\ f)\\ &= \alpha <v\,| + \alpha <w\,|\,.\end{aligned}$$

Therefore, scalar multiplication is distributive over vector addition.

- Distribution in scalars, equation (9):

$$\begin{aligned}(\alpha+\beta) <v\,| &= (\alpha+\beta)\,(a,\ b,\ c)\\ &= \Big((\alpha+\beta)\,a,\ (\alpha+\beta)\,b,\ (\alpha+\beta)\,c\Big)\\ &= (\alpha\, a + \beta\, a,\ \alpha\, b + \beta\, b,\ \alpha\, c + \beta\, c)\\ &= (\alpha\, a,\ \alpha\, b,\ \alpha\, c) + (\beta\, a,\ \beta\, b,\ \beta\, c)\\ &= \alpha\,(a,\ b,\ c) + \beta\,(a,\ b,\ c) = \alpha <v\,| + \beta <v\,|\,.\end{aligned}$$

Therefore, vector multiplication is distributive over scalar addition.

- Associativity in scalar multiplication, equation (10):

$$\alpha\,(\beta <v\,|\,) = \alpha\Big(\beta\,(a,\ b,\ c)\Big) = \alpha\,(\beta a,\ \beta b,\ \beta c) = (\alpha\,\beta\, a,\ \alpha\,\beta\, b,\ \alpha\,\beta\, c) = \alpha\,\beta\,(a,\ b,\ c) = \alpha\,\beta <v\,|\,.$$

Therefore, scalar multiplication is associative.

Since the given definitions satisfy all ten conditions over the field of real numbers, we conclude the definitions of addition and scalar multiplication constitute a linear vector space in $\mathbf{R}^3$.

Postscript: The symbol $\mathbf{R}^3$ means the space of vectors with three real components. The symbol $\mathbf{C}^3$ means the space of vectors with three complex components. The symbol $\mathbf{C}^\infty$ means the space of vectors that have an infinite number of complex components.

---

13. Use the definitions of vector addition and scalar multiplication given in problem 12 to show that the set of all vectors of the form $<u\,| = (a,\ b,\ 1)$ does not constitute a linear vector space.

This problem is a counterexample of a system that is not a linear vector space. It is intended to provide contrast and context to those systems that do form linear vector spaces. The set and definitions of vector addition and scalar multiplication must satisfy all ten properties to constitute a linear vector space. Check to see if this system is closed under vector addition; add two vectors of the form given. Does their sum have the same form?

For vectors of the form $<u\,| = (a,\ b,\ 1)$,

$$<v\,| + <w\,| = (a,\ b,\ 1) + (c,\ d,\ 1) = (a+c,\ b+d,\ 1+1) = (e,\ f,\ 2)$$

which is not of the form $(a,\ b,\ 1)$ because of the third component, *i.e.*, $2 \neq 1$, so the sum $(e,\ f,\ 2)$ is not contained within the space. Therefore, the space is not closed under vector addition, and this is sufficient to show that all vectors of the form $<u\,| = (a,\ b,\ 1)$ do not form a linear vector space using common definitions of addition and scalar multiplication.

---

14. (a) Show that $|\,u_1> = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $|\,u_2> = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$, and $|\,u_3> = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ are linearly dependent.

(b) Show that $|\,v_1> = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $|\,v_2> = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, and $|\,v_3> = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$ are linearly dependent.

(c) Show that $|\,w_1> = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $|\,w_2> = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, and $|\,w_3> = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ are linearly independent.

A set of vectors is **linearly independent** if and only if $\displaystyle\sum_{i=1}^{n} \alpha_i\,|\,v_i> \neq\ |\,0>$ unless all $\alpha_i = 0$. Linear independence is very important to quantum mechanical calculation though it is rarely a focal issue because linear independence is usually assumed. In fact, we will see that linear independence is inherent in the first postulate of quantum mechanics.

Likely the easiest way to determine linear independence is proof by contradiction. For example, assume the set of vectors given in part (a) is linearly independent. Then

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + \alpha_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

if and only if $\alpha_1 = \alpha_2 = \alpha_3 = 0$. But the components form the simultaneous equations

$$\begin{aligned} \alpha_1 + \alpha_2 - \alpha_3 &= 0 \\ \alpha_1 + 3\alpha_2 &= 0 \\ 2\alpha_2 + \alpha_3 &= 0 \end{aligned}$$

for which $\alpha_1 = 3$, $\alpha_2 = -1$ and $\alpha_3 = 2$ is a solution. This contradicts our assumption. Therefore, the given vectors are not linearly independent, or are **linearly dependent**.

(b) Assume the vectors are linearly independent. Then

$$a_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + a_3 \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

where $a_1 = a_2 = a_3 = 0$ would be the only coefficients that could make this equation true. Performing the scalar multiplications, the above condition means

$$\begin{pmatrix} a_1 \\ a_1 \\ 0 \end{pmatrix} + \begin{pmatrix} a_2 \\ 0 \\ a_2 \end{pmatrix} + \begin{pmatrix} 3a_3 \\ 2a_3 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \Rightarrow \qquad \begin{aligned} a_1 + a_2 + 3a_3 &= 0 \\ a_1 + 2a_3 &= 0 \\ a_2 + a_3 &= 0 \end{aligned}$$

that is, the individual components must sum to zero. This particular system has an infinite number of solutions such that $a_i \neq 0$, for example, $a_1 = 1$, $a_2 = 1/2$, and $a_3 = -1/2$. Therefore, in contradiction to the assumption, $|\, v_1 \!>$, $|\, v_2 \!>$, and $|\, v_3 \!>$ are linearly dependent.

(c) However, consider

$$a_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + a_3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} a_1 \\ a_1 \\ 0 \end{pmatrix} + \begin{pmatrix} a_2 \\ 0 \\ a_2 \end{pmatrix} + \begin{pmatrix} 0 \\ a_3 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad \begin{aligned} a_1 + a_2 &= 0 \\ a_1 + a_3 &= 0 \\ a_2 + a_3 &= 0 \end{aligned} \qquad \Rightarrow \quad \begin{aligned} a_1 &= -a_2 \\ a_1 &= -a_3 \\ a_2 &= -a_3 \end{aligned} \qquad \Rightarrow \quad \begin{aligned} a_1 &= -a_2 \\ -a_2 &= -a_3 \\ a_3 &= -a_3 \end{aligned}$$

The last equation can be true if and only if $a_3 = 0 \;\Rightarrow\; a_2 = 0 \;\Rightarrow\; a_1 = 0$. Since all $a_i = 0$, $|\, w_1 \!>$, $|\, w_2 \!>$, and $|\, w_3 \!>$ are linearly independent.

Postscript: The definition of linear independence has an upper limit on the summation of n. The upper limit can be any number including ∞. In other words, though the vectors in this problem are in three dimensions, the concept generalizes to any including infinite dimensions.

The condition of linear independence means that you cannot form any of the other vectors in the set using a linear combination of two or more vectors in the set.

If the number of vectors in a linearly independent set is the same as the dimension of the linear vector space, all of the other vectors in the space can be formed using a linear combination of this basic set, or set of **basis** vectors. There are an infinite number of sets of vectors that will constitute a basis for any linear vector space. The choice of a basis set is often suggested by the context of the problem just as you might chose the unit vectors $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$; or $\hat{r}$, $\hat{\theta}$, $\hat{\phi}$, for a mechanics application based on the symmetry of the problem. When all of the other vectors in the space can be formed using a linear combination of a set of vectors, that set of vectors is said to **span** the space. All basis sets span the space for which they are a basis.

The number of components in a column or row vector effectively determines the dimension of the space for spaces composed of vectors as we have defined vectors. We will find that sets of objects other than vectors, operators for instance, can constitute a linear vector space. The dimension of a linear vector space is more generally defined by the maximum number of linearly independent objects in the space.

---

15. Which of the following sets of vectors are orthogonal?

$$\text{(a)} \quad |\, 1 \!> = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad |\, 2 \!> = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \qquad |\, 3 \!> = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$$(b) \quad |1> = \begin{pmatrix} 1 \\ i \\ i \end{pmatrix}, \quad |2> = \begin{pmatrix} 1 \\ 2 \\ i \end{pmatrix}, \quad |3> = \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix}$$

$$(c) \quad |1> = \begin{pmatrix} 1+i \\ i \\ -2-i \end{pmatrix}, \quad |2> = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \quad |3> = \begin{pmatrix} 1+i \\ 1 \\ -1 \end{pmatrix}$$

$$(d) \quad |1> = \begin{pmatrix} 1 \\ 0 \\ 1+i \end{pmatrix}, \quad |2> = \begin{pmatrix} -1+i \\ -i \\ 1 \end{pmatrix}, \quad |3> = \begin{pmatrix} 1+i \\ 3 \\ -i \end{pmatrix}$$

This exercise is designed to introduce the property of **orthogonality**. Orthogonality is a property that leads to **orthonormality**. Orthonormality is essential to any set of basis vectors that is used for quantum mechanical calculation.

The vectors $3\,\hat{\imath},\ 2\,\hat{\jmath},\ 4\,\hat{k}$ are mutually perpendicular. The fact that the dot product of any of these vectors with any other is zero demonstrates this property of mutual perpendicularity. The term "perpendicular" does not extend to more than three dimensions. The concept is generalized to higher dimensions using the higher dimensional analog of the dot product, the inner product. If the inner product of every vector in the set with every other vector in the set is zero, then that set of vectors is **orthogonal**.

To show that any of these sets of vectors are orthogonal, you must show that the inner product between any two members of the set is zero, *i.e.*, you must show that $<i\,|\,j> = 0$ for all $i \neq j$. If any of the $<i\,|\,j> \neq 0$, then the system is not orthogonal. For the three vectors given for part (c) for instance,

$$<1\,|\,2> = \begin{pmatrix} 1-i, & -i, & -2+i \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = (2-2i) \ + \ (i) \ + \ (-2+i) = 2-2-2i+i+i = 0\,,$$

so these two vectors are orthogonal. However,

$$<1\,|\,3> = \begin{pmatrix} 1-i, & -i, & -2+i \end{pmatrix} \begin{pmatrix} 1+i \\ 1 \\ -1 \end{pmatrix} = (1+1) + (-i) + (2-i) = 4-2i \neq 0\,,$$

so the set of three vectors given in part (c) is not orthogonal. Remember to conjugate the components when forming bras. By the way, if $<i\,|\,j> = 0$, then $<j\,|\,i> = 0$.

(a)
$$<1\,|\,2> = \begin{pmatrix} 1, & 1, & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} = 2-1-1 = 0$$

$$<1\,|\,3> = \begin{pmatrix} 1, & 1, & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = 0+1-1 = 0$$

$$<2\,|\,3> = \begin{pmatrix}2, & -1, & -1\end{pmatrix}\begin{pmatrix}0\\1\\-1\end{pmatrix} = 0 - 1 + 1 = 0$$

therefore, this set is orthogonal because $<i\,|\,j> = 0$.

(b)
$$<1\,|\,2> = \begin{pmatrix}1, & -i, & -i\end{pmatrix}\begin{pmatrix}1\\2\\i\end{pmatrix} = 1 - 2i + 1 = 2 - 2i \neq 0$$

therefore, this set is not orthogonal because $<1\,|\,2> \neq 0$.

(c) $<1\,|\,3> \neq 0$ and $<2\,|\,3> \neq 0$, therefore, this set is not orthogonal. Of course, either non-zero inner product is sufficient to show that the set is not orthogonal.

(d)
$$<1\,|\,2> = \begin{pmatrix}1, & 0, & 1-i\end{pmatrix}\begin{pmatrix}-1+i\\-i\\1\end{pmatrix} = -1 + i - 0 + 1 - i = 0$$

$$<1\,|\,3> = \begin{pmatrix}1, & 0, & 1-i\end{pmatrix}\begin{pmatrix}1+i\\3\\-i\end{pmatrix} = 1 + i + 0 - i - 1 = 0$$

$$<2\,|\,3> = \begin{pmatrix}-1-i, & i, & 1\end{pmatrix}\begin{pmatrix}1+i\\3\\-i\end{pmatrix} = -1 - i - i + 1 + 3i - i = 0$$

therefore, this set is orthogonal because $<i\,|\,j> = 0$.

Postscript: A set of basis vectors that is useful for quantum mechanical calculation does not include a zero vector. All basis vectors must be non-zero.

As indicated, if $<i\,|\,j> = 0$, then $<j\,|\,i> = 0$. But if $<i\,|\,j> \neq 0$, then generally $<i\,|\,j> \neq <j\,|\,i>$.

None of the vectors given in this problem are normalized. A set of vectors that is both orthogonal and normalized has the property of **orthonormality**.

---

16. Show that the two dimensional systems

(a) $\quad |\,1> = \dfrac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \qquad |\,2> = \dfrac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}, \qquad$ and

(b) $\quad |\,v> = \dfrac{1}{2}\begin{pmatrix}1-i\\\sqrt{2}\end{pmatrix} \qquad |\,w> = \dfrac{1}{2}\begin{pmatrix}-1+i\\\sqrt{2}\end{pmatrix} \qquad$ are orthonormal.

This problem is an introduction to **orthonormality**. If the inner product of dissimilar vectors is zero and the inner product of a ket with its bra is 1, then that set is **orthonormal**. In other words, an orthonormal set of vectors is both orthogonal and normalized. All orthogonal sets can be made orthonormal by the process of normalization. In fact, any linearly independent set of vectors can be made orthogonal and thus orthonormal. A basis must be orthonormal to be useful for quantum mechanics. The unit vectors $\hat{\imath}\,,\ \hat{\jmath}\,,\ \hat{k}\,,$ are orthonormal in three dimensions, for instance. Show that the inner product of each ket with its corresponding bra is 1, and that the inner product of the two vectors is zero. Symbolically, show that $<i\,|\,i> = 1$ and $<i\,|\,j> = 0\,.$

(a) The inner products of each vector with itself are

$$(1,\ 1)\,\frac{1}{\sqrt{2}}\,\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \left(\frac{1}{\sqrt{2}}\right)^2(1+1) = \frac{1}{2}\,(2) = 1\,,$$

$$(1,-1)\,\frac{1}{\sqrt{2}}\,\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} = \left(\frac{1}{\sqrt{2}}\right)^2(1+1) = \frac{1}{2}\,(2) = 1\,,$$

from which we conclude they are normalized, and the inner product of the two vectors is

$$(1,\ 1)\,\frac{1}{\sqrt{2}}\,\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} = \left(\frac{1}{\sqrt{2}}\right)^2(1-1) = \frac{1}{2}\,(0) = 0\,,\quad \text{so the system is orthonormal.}$$

(b) The inner products of each vector with itself are

$$<v\,|\,v> = \big(1+i,\ \sqrt{2}\,\big)\,\frac{1}{2}\,\frac{1}{2}\begin{pmatrix}1-i\\ \sqrt{2}\end{pmatrix} = \frac{1}{4}\,\big(1+1+2\big) = 1\,,$$

$$<w\,|\,w> = \big(-1-i,\ \sqrt{2}\,\big)\,\frac{1}{2}\,\frac{1}{2}\begin{pmatrix}-1+i\\ \sqrt{2}\end{pmatrix} = \frac{1}{4}\,\big(1+1+2\big) = 1\,.$$

The inner product of the two vectors is

$$<v\,|\,w> = \big(1+i,\ \sqrt{2}\,\big)\,\frac{1}{2}\,\frac{1}{2}\begin{pmatrix}-1+i\\ \sqrt{2}\end{pmatrix} = \frac{1}{4}\,\big(-1+i-i-1+2\big) = 0\,,\quad \text{so the system is orthonormal.}$$

Postscript: The **orthonormality condition** is often expressed using the **Kronecker delta**. The Kronecker delta is defined

$$\delta_{ij} = \begin{cases} 1\,, & \text{for } i = j\,,\\ 0\,, & \text{for } i \neq j\,. \end{cases}$$

Using Dirac notation and the Kronecker delta, orthonormality is denoted

$$<i\,|\,j> = \delta_{ij}\,.$$

This problem demonstrates how a bra is formed from a ket that has a normalization constant. The normalization constant is placed to the right of a bra. If the normalization constant is complex, it must be conjugated in forming the bra.

---

17. Orthonormalize the vectors

$$|\,I> = \begin{pmatrix}1\\-3\end{pmatrix} \quad \text{and} \quad |\,II> = \begin{pmatrix}3\\4\end{pmatrix},$$

and check to ensure that your resulting vectors are orthonormal.

You can make an orthonormal basis from any linearly independent set of vectors using the **Gram-Schmidt orthonormalization procedure**. Given a linearly independent basis, the Gram-Schmidt orthonormalization procedure has you

(1) choose a vector $|\,I>$ at random and normalize it to attain the unit vector $|\,1>$.

(2) Choose a second vector $|\,II>$ and subtract its projection on the first vector. An inner product times either unit vector is the projection in the direction of the unit vector, per figure 1–1. This means

$$|\,2> = |\,II> - |\,1><1\,|\,II> \,. \qquad \text{Normalize this vector.}$$

Figure 1 − 1. The dot or inner product can be interpreted geometrically as the projection of one vector along the other. Multiply the projection by either normalized vector and you have a vector that is the projection of the other vector along it.

(3) Choose another vector $|\,III>$ and subtract its projection along the first two vectors, or

$$|\,3> = |\,III> - |\,1><1\,|\,III> - |\,2><2\,|\,II> \,. \qquad \text{Normalize this vector.}$$

(4) Repeat the procedure until you exhaust all the vectors in the linearly independent basis.

The vectors

$$|\,I> = \begin{pmatrix} 1 \\ -3 \end{pmatrix} \quad \text{and} \quad |\,II> = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

are linearly independent. They can be made orthonormal using the Gram-Schmidt orthonormalization procedure. First, normalize one vector. Attach a normalization constant to $|\,I>$ and

$$<I\,|\,\mathrm{A}^*\mathrm{A}\,|\,I> = \big(1,\ -3\big)\,\mathrm{A}^*\,\mathrm{A}\begin{pmatrix} 1 \\ -3 \end{pmatrix} = |\,\mathrm{A}\,|^2\big(1+9\big) = |\,\mathrm{A}\,|^2\,10 = 1\,,$$

$$\Rightarrow \quad \mathrm{A} = \frac{1}{\sqrt{10}} \quad \Rightarrow \quad |\,1> = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ -3 \end{pmatrix}$$

is normalized. Next subtract the projection of the second along the first,

$$\begin{aligned} |\,2> &= |\,II> - |\,1><1\,|\,II> \\ &= \begin{pmatrix} 3 \\ 4 \end{pmatrix} - \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ -3 \end{pmatrix}\left[\big(1,\ -3\,\big)\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right]\frac{1}{\sqrt{10}} \\ &= \begin{pmatrix} 3 \\ 4 \end{pmatrix} - \frac{1}{10}\begin{pmatrix} 1 \\ -3 \end{pmatrix}(3-12) = \begin{pmatrix} 3 \\ 4 \end{pmatrix} + \frac{9}{10}\begin{pmatrix} 1 \\ -3 \end{pmatrix} \\ &= \begin{pmatrix} 3+9/10 \\ 4-27/10 \end{pmatrix} = \begin{pmatrix} 39/10 \\ 13/10 \end{pmatrix} = \frac{13}{10}\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \end{aligned}$$

where the square brackets in the second line are used only to indicate that we are going to do that vector multiplication first. Normalizing this vector,

$$\frac{13}{10}\left(3,\ 1\right)\mathrm{A}^*\,\frac{13}{10}\begin{pmatrix}3\\1\end{pmatrix} = \left|\,\mathrm{A}\,\right|^2\frac{169}{100}\left(9+1\right) = \left|\,\mathrm{A}\,\right|^2\frac{169}{10} = 1$$

$$\Rightarrow \quad \mathrm{A} = \frac{\sqrt{10}}{13} \quad \Rightarrow \quad |\,2> \,= \frac{\sqrt{10}}{13}\,\frac{13}{10}\begin{pmatrix}3\\1\end{pmatrix} = \frac{\sqrt{10}}{10}\begin{pmatrix}3\\1\end{pmatrix}$$

is the normalized vector. Here, we have exhausted all the vectors in the linearly independent, two dimensional basis, so have completed the Gram-Schmidt procedure.

Are these vectors now orthonormal? Checking

$$<1\,|\,2> \,= \left(1,\ -3\right)\frac{1}{\sqrt{10}}\,\frac{\sqrt{10}}{10}\begin{pmatrix}3\\1\end{pmatrix} = \frac{1}{10}\left(3-3\right) = \frac{1}{10}\left(0\right) = 0\,,$$

so they are orthogonal, and since they are normalized, they are also orthonormal.

Postscript: You should understand that any linearly independent set of vectors can be converted into an orthonormal basis, but we will never have occasion to employ the mathematical mechanics explicit in the Gram-Schmidt orthonormalization procedure. We will encounter linearly independent systems for which you now know that an orthonormal basis exists. The existence of the orthonormal basis provides the foundation for further calculation. The concept that any linearly independent set of vectors can be made into an orthonormal basis is the point of this problem, and the mathematical mechanics are a detail that is not particularly important to the physics.

---

18. (a) Show that the vectors

$$|\,I> \,= \begin{pmatrix}3\\0\\0\end{pmatrix}, \quad |\,II> \,= \begin{pmatrix}0\\1\\2\end{pmatrix}, \quad |\,III> \,= \begin{pmatrix}0\\2\\5\end{pmatrix} \quad \text{are linearly independent.}$$

(b) Show that the system of three vectors in part (a) is not orthogonal.

(c) Use the Gram-Schmidt process to form an orthonormal basis from the vectors in part (a).

(d) Show that the vectors resulting from part (c) are orthonormal.

This problem is an example of the fact that linear independence is a sufficient condition for orthonormality. The reason a student of quantum mechanics wants to understand linear independence is explicitly because it is a sufficient condition for orthonormality. An orthonormal basis is a practical necessity for quantum mechanical calculation. This problem also shows how a linearly independent system can be made orthonormal using the Gram-Schmidt procedure. Notice that orthogonality is not necessary (though an orthogonal system can be made orthonormal directly through the process of normalization).

Use the strategy of demonstrating linear independence seen in problem 14 for part (a). Part (b) requires you to form inner products. This system is not orthogonal, but only one of the nine possible inner products disqualifies it. Part (c) is the Gram-Schmidt process. You should find

$$|1> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |2> = \begin{pmatrix} 0 \\ 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}, \quad |3> = \begin{pmatrix} 0 \\ -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}.$$

Part (d) answers the question does this Gram-Schmidt process really work? Yes it does, and you really need only check that the inner product analogous to that which was non-zero in part (b) is zero after the Gram-Schmidt process has been applied. The other properties of orthonormality are satisfied because of the nature of normalized vectors and two of the results of part (b).

(a) The condition of linear independence is

$$\sum_i^n a_i \,|\, v_i> \neq |\, 0> \quad \text{unless all} \quad a_i = 0\,. \quad \text{Here this means}$$

$$a_1 \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + a_3 \begin{pmatrix} 0 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 3a_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ 2a_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 2a_3 \\ 5a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{array}{rl} 3a_1 & = 0 \\ a_2 + 2a_3 & = 0 \\ 2a_2 + 5a_3 & = 0 \end{array} \quad \Rightarrow \quad \begin{array}{rl} a_1 & = 0 \\ a_2 & = -2a_3\,. \\ 2a_2 & = -5a_3 \end{array}$$

The first line says a_1 is zero. Substituting the equation from the second line into the equation of the third line,

$$2a_2 = -5a_3 \quad \Rightarrow \quad 2\,(-2a_3) = -4a_3 = -5a_3\,,$$

which can be true only if $a_3 = 0$. Then since $a_2 = -2a_3$, a_2 is also zero. The unique solution for all coefficients is that all $a_i = 0$, therefore, the three vectors are linearly independent.

(b) To show the three vectors are not orthogonal, we need to show only that any inner product of different vectors does not vanish. In fact

$$<II\,|\,III> = \begin{pmatrix} 0, & 1, & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 5 \end{pmatrix} = 0 + 2 + 10 = 12 \neq 0\,,$$

therefore, this system of three vectors is not orthogonal. By the way, $<I\,|\,II> = <I\,|\,III> = 0$.

(c) First normalize $|\,1>$,

$$<1\,|\,\mathrm{A}^*\mathrm{A}\,|\,1> = \begin{pmatrix} 3, & 0, & 0 \end{pmatrix} \mathrm{A}^*\mathrm{A} \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} = |\,\mathrm{A}\,|^2 \big(9 + 0 + 0\big) = |\,\mathrm{A}\,|^2\, 9 = 1$$

$$\Rightarrow \quad \mathrm{A} = \frac{1}{3} \quad \Rightarrow \quad |\,1\!> = \frac{1}{3}\begin{pmatrix}3\\0\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix}.$$

Then construct $|\,2'\!>$,

$$|\,2'\!> = \begin{pmatrix}0\\1\\2\end{pmatrix} - \begin{pmatrix}1\\0\\0\end{pmatrix}\Big(1,0,0\Big)\begin{pmatrix}0\\1\\2\end{pmatrix} = \begin{pmatrix}0\\1\\2\end{pmatrix} - \begin{pmatrix}1\\0\\0\end{pmatrix}(0+0+0) = \begin{pmatrix}0\\1\\2\end{pmatrix}$$

and then normalize $|\,2'\!>$ to obtain $|\,2\!>$,

$$<\!2\,|\,\mathrm{A}^*\mathrm{A}\,|\,2\!> = \Big(0,\ 1,\ 2\Big)\,\mathrm{A}^*\mathrm{A}\begin{pmatrix}0\\1\\2\end{pmatrix} = |\,\mathrm{A}\,|^2\Big(0+1+4\Big) = |\,\mathrm{A}\,|^2\,5 = 1$$

$$\Rightarrow \quad \mathrm{A} = \frac{1}{\sqrt{5}} \quad \Rightarrow \quad |\,2\!> = \frac{1}{\sqrt{5}}\begin{pmatrix}0\\1\\2\end{pmatrix}.$$

Finally, construct $|\,3'\!>$,

$$\begin{aligned}
|\,3'\!> &= \begin{pmatrix}0\\2\\5\end{pmatrix} - \begin{pmatrix}1\\0\\0\end{pmatrix}\Big(1,0,0\Big)\begin{pmatrix}0\\2\\5\end{pmatrix} - \begin{pmatrix}0\\1/\sqrt{5}\\2/\sqrt{5}\end{pmatrix}\left(0,\ \frac{1}{\sqrt{5}},\ \frac{2}{\sqrt{5}}\right)\begin{pmatrix}0\\2\\5\end{pmatrix}\\
&= \begin{pmatrix}0\\2\\5\end{pmatrix} - \begin{pmatrix}1\\0\\0\end{pmatrix}(0+0+0) - \begin{pmatrix}0\\1/\sqrt{5}\\2/\sqrt{5}\end{pmatrix}\left(0+\frac{2}{\sqrt{5}}+\frac{10}{\sqrt{5}}\right)\\
&= \begin{pmatrix}0\\2\\5\end{pmatrix} - \begin{pmatrix}1\\0\\0\end{pmatrix}(0) - \begin{pmatrix}0\\1/\sqrt{5}\\2/\sqrt{5}\end{pmatrix}\frac{12}{\sqrt{5}}\\
&= \begin{pmatrix}0\\2\\5\end{pmatrix} - \begin{pmatrix}0\\12/5\\24/5\end{pmatrix} = \begin{pmatrix}0\\-2/5\\1/5\end{pmatrix}
\end{aligned}$$

and then normalize $|\,3'\!>$ to obtain $|\,3\!>$,

$$<\!3\,|\,\mathrm{A}^*\mathrm{A}\,|\,3\!> = \left(0,\ -\frac{2}{5},\ \frac{1}{5}\right)\,\mathrm{A}^*\mathrm{A}\begin{pmatrix}0\\-2/5\\1/5\end{pmatrix} = |\,\mathrm{A}\,|^2\left(0+\frac{4}{25}+\frac{1}{25}\right) = |\,\mathrm{A}\,|^2\,\frac{1}{5} = 1$$

$$\Rightarrow \quad \mathrm{A} = \sqrt{5} \quad \Rightarrow \quad |\,3\!> = \sqrt{5}\begin{pmatrix}0\\-2/5\\1/5\end{pmatrix} = \begin{pmatrix}0\\-2/\sqrt{5}\\1/\sqrt{5}\end{pmatrix}.$$

Summarizing, the application of the Gram-Schmidt procedure yields the three vectors

$$|\,1\!> = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad |\,2\!> = \begin{pmatrix}0\\1/\sqrt{5}\\2/\sqrt{5}\end{pmatrix}, \quad |\,3\!> = \begin{pmatrix}0\\-2/\sqrt{5}\\1/\sqrt{5}\end{pmatrix}.$$

(d) The vectors are normalized, so each inner product with itself is 1. $<I\,|\,II> = <I\,|\,III> = 0$, as before, because all inner products of the form

$$\begin{pmatrix} a, & 0, & 0 \end{pmatrix} \begin{pmatrix} 0 \\ b \\ c \end{pmatrix} = 0 + 0 + 0 = 0.$$

However, $<II\,|\,III>$ was non-zero, but

$$<2\,|\,3> = \begin{pmatrix} 0, & 1/\sqrt{5}, & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 0 \\ -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} = -\frac{2}{5} + \frac{2}{5} = 0,$$

so the three vectors in the summary of part (c) are orthonormal.

Postscript: Again, this problem is an example that linear independence is a sufficient condition for orthonormality. The practical reason for this is that if linearly independence can be demonstrated, an orthonormal basis is assumed. Our interest is in the orthonormal basis. Linear independence establishes the existence of the orthonormal basis. The validity of this assumption can be based on the Gram-Schmidt procedure, though we will never have reason to explicitly employ the Gram-Schmidt process.

---

19. Find the adjoint of the operator

$$\mathcal{A} = \begin{pmatrix} -4+2i & 2+i & -1-3i \\ 6 & 2+3i & -3i \\ -3-i & 2-3i & -5 \end{pmatrix}.$$

An adjoint is a **transpose conjugate**. Symbolically,

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \Rightarrow \quad \mathcal{A}^{\dagger} = \begin{pmatrix} a_{11}^* & a_{21}^* & a_{31}^* \\ a_{12}^* & a_{22}^* & a_{32}^* \\ a_{13}^* & a_{23}^* & a_{33}^* \end{pmatrix},$$

where the dagger symbol, †, located at the upper right of the operator indicates adjoint. Simply make the rows into columns and the columns into rows and complex conjugate all of the elements. The mechanics of forming an adjoint operator is easy. The reason to form an adjoint is deeper, and we will attempt to explain the deeper significance in the postscript.

$$\mathcal{A} = \begin{pmatrix} -4+2i & 2+i & -1-3i \\ 6 & 2+3i & -3i \\ -3-i & 2-3i & -5 \end{pmatrix} \quad \Rightarrow \quad \mathcal{A}^{\dagger} = \begin{pmatrix} -4-2i & 6 & -3+i \\ 2-i & 2-3i & 2+3i \\ -1+3i & 3i & -5 \end{pmatrix}.$$

Postscript: A **dual space** contains complex conjugates or complex conjugate analogies to the objects that comprise the original space. The concept is that the complex conjugate or analogy to a complex conjugate of any scalar, vector, or matrix operator exists in the dual space, and when necessary, the appropriate object is simply picked from all of the other objects in the dual space. A complex conjugate is in the dual space of the corresponding scalar, a bra is in the dual space of the corresponding ket, and an adjoint operator is in the dual space of the corresponding operator.

$$\text{scalars} \quad \overbrace{\alpha \longleftrightarrow \alpha^*}^{\text{complex conjugate}} \qquad \text{vectors} \quad \overbrace{|\,v> \longleftrightarrow <v\,|}^{\text{adjoint}} \qquad \text{operators} \quad \overbrace{\mathcal{A} \longleftrightarrow \mathcal{A}^\dagger}^{\text{adjoint}}$$

Notice that a bra formed from a ket satisfies the definition of an adjoint. "Transpose" means to make the rows into columns and the columns into rows. A bra is a transpose conjugate of a ket.

Among the reasons to consider a dual space is that a complex conjugate or complex conjugate analogy exists for any object that can be considered for quantum mechanical calculation. The complex linear vector space of actual quantum mechanical calculation is often of infinite dimension. It is impossible to form an infinity of complex conjugates or adjoints, yet, we need to know that these objects also exist. It is straightforward to show that they exist in a dual space comprised of complex conjugates or objects analogous to complex conjugates.

Secondly, the physical measurements of any branch of physics including quantum mechanics must be real quantities. The product of any scalar and its complex conjugate is a real number. The product of any ket and its corresponding bra is a real number. We will find that the product of any matrix operator and its adjoint is a matrix operator that has elements that are real. The dual space comprised of complex conjugates/complex conjugate analogies provides a mathematically tractable mechanism for obtaining real quantities from intrinsically complex objects. A significant byproduct of the dual space is that there is a geometric analog to the length of a vector, the norm, regardless of the number of dimensions.

The concept of a space and a corresponding dual space is a precursor to the **Hilbert space**. A Hilbert space is defined by three properties, namely: 1) linearity, 2) orthonormality, and 3) completeness. Physical phenomena that can be described by quantum mechanics appears to be linear, you should understand orthonormality as it applies to kets and bras at this point, and completeness means that the set spans the space. The physicist's view of the Hilbert space is not in concert with the mathematician's view of a Hilbert space. There are an infinite number of spaces that satisfy these three properties. Physicists collect every possible set that meets the definition of a Hilbert space and call that the Hilbert space. Mathematics discusses a specific Hilbert space because there are an infinite number, while physics discusses the Hilbert space because the physicist sees only one. Infinite and finite dimensional operators and their adjoints, infinite and finite dimensional kets and their bras, all scalars and their complex conjugates, as well as orthogonal functions likes sines and cosines, and orthogonal polynomials like Hermite, Laguerre, and Legendre polynomials, and many other objects are included in the physicist's view of the Hilbert space. Most problems can be addressed using a subset or subspace of the Hilbert space. We will say more about the Hilbert space. Hilbert space is discussed in most books on functional analysis such as Shilov[1] and Riesz and Nagy[2]. Chapter 5 of Byron and Fuller[3] is an

[1] Shilov *Elementary Functional Analysis,* Dover Publications, 1974, pp. 54–68.

[2] Riesz and Nagy *Functional Analysis,* Dover Publications, 1990, pp. 197–199.

[3] Byron and Fuller *The Mathematics of Classical and Quantum Physics*, Dover Publications, 1969.

excellent discussion of Hilbert space. These references provide more extensive information than needed for our presentation. For our purposes, think of the Hilbert space as the collection of all orthonormal sets.

---

20. Given that

$$\mathcal{A} = \begin{pmatrix} -4+2i & 2+i & -1-3i \\ 6 & 2+3i & -3i \\ -3-i & 2-3i & -5 \end{pmatrix} \quad \text{and} \quad \mathcal{B} = \begin{pmatrix} 3-i & 2-2i & -4+2i \\ 1-3i & 2+4i & 5-2i \\ -1-6i & 7 & -4+i \end{pmatrix}, \quad \text{find}$$

(a) $\mathcal{A}+\mathcal{B}$,

(b) $\mathcal{A}-\mathcal{B}$,

(c) $\mathcal{A}^\dagger + \mathcal{B}^\dagger$, and

(d) $\left(\mathcal{A}+\mathcal{B}\right)^\dagger$.

Operators are added/subtracted by adding/subtracting corresponding elements. Symbolically,

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \\ a_{31}+b_{31} & a_{32}+b_{32} & a_{33}+b_{33} \end{pmatrix}$$

for addition. Subtract corresponding elements if the indicated operation is subtraction. You know how to find adjoints from problem 19. You should find that the answers to parts (c) and (d) are identical—the sum of the adjoints is the same as the adjoint of the sum.

(a)
$$\begin{aligned} \mathcal{A}+\mathcal{B} &= \begin{pmatrix} -4+2i+(3-i) & 2+i+(2-2i) & -1-3i+(-4+2i) \\ 6+(1-3i) & 2+3i+(2+4i) & -3i+(5-2i) \\ -3-i+(-1-6i) & 2-3i+(7) & -5+(-4+i) \end{pmatrix} \\ &= \begin{pmatrix} -1+i & 4-i & -5-i \\ 7-3i & 4+7i & 5-5i \\ -4-7i & 9-3i & -9+i \end{pmatrix}. \end{aligned}$$

(b)
$$\begin{aligned} \mathcal{A}-\mathcal{B} &= \begin{pmatrix} -4+2i-(3-i) & 2+i-(2-2i) & -1-3i-(-4+2i) \\ 6-(1-3i) & 2+3i-(2+4i) & -3i-(5-2i) \\ -3-i-(-1-6i) & 2-3i-(7) & -5-(-4+i) \end{pmatrix} \\ &= \begin{pmatrix} -7+3i & 3i & 3-5i \\ 5+3i & -i & -5-i \\ -2+5i & -5-3i & -1-i \end{pmatrix}. \end{aligned}$$

(c)
$$\begin{pmatrix} -4-2i & 6 & -3+i \\ 2-i & 2-3i & 2+3i \\ -1+3i & 3i & -5 \end{pmatrix} + \begin{pmatrix} 3+i & 1+3i & -1+6i \\ 2+2i & 2-4i & 7 \\ -4-2i & 5+2i & -4-i \end{pmatrix} = \begin{pmatrix} -1-i & 7+3i & -4+7i \\ 4+i & 4-7i & 9+3i \\ -5+i & 5+5i & -9-i \end{pmatrix}.$$

(d) Using the result of part (a),

$$\left(\mathcal{A}+\mathcal{B}\right)^{\dagger} = \begin{pmatrix} -1+i & 4-i & -5-i \\ 7-3i & 4+7i & 5-5i \\ -4-7i & 9-3i & -9+i \end{pmatrix}^{\dagger} = \begin{pmatrix} -1-i & 7+3i & -4+7i \\ 4+i & 4-7i & 9+3i \\ -5+i & 5+5i & -9-i \end{pmatrix}.$$

---

21. Given that

$$\mathcal{A} = \begin{pmatrix} -4+2i & 2+i & -1-3i \\ 6 & 2+3i & -3i \\ -3-i & 2-3i & -5 \end{pmatrix} \qquad \text{and} \qquad |\,v> = \begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix}, \qquad \text{find } \mathcal{A}\,|\,v>\ .$$

An operator times a vector is a vector. A symbolic example of $\mathcal{A}\,|\,v>$ in 3 dimensions is

$$\mathcal{A}\,|\,v> = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + a_{12}b_2 + a_{13}b_3 \\ a_{21}b_1 + a_{22}b_2 + a_{23}b_3 \\ a_{31}b_1 + a_{32}b_2 + a_{33}b_3 \end{pmatrix} = |\,w>\ .$$

In layman's terms, multiply each row of the operator by the column vector as if it was a row vector. The sum of the products is the component of the new vector $|\,w>$ at the same level as the horizontal row of the operator that was used.

$$\mathcal{A}\,|\,v> = \begin{pmatrix} -4+2i & 2+i & -1-3i \\ 6 & 2+3i & -3i \\ -3-i & 2-3i & -5 \end{pmatrix} \begin{pmatrix} 2+i \\ 1-3i \\ 3-2i \end{pmatrix}$$

$$= \begin{pmatrix} (-4+2i)(2+i) + (2+i)(1-3i) + (-1-3i)(3-2i) \\ 6(2+i) + (2+3i)(1-3i) + (-3i)(3-2i) \\ (-3-i)(2+i) + (2-3i)(1-3i) + (-5)(3-2i) \end{pmatrix}$$

$$= \begin{pmatrix} (-8-4i+4i-2) + (2-6i+i+3) + (-3+2i-9i-6) \\ (12+6i) + (2-6i+3i+9) + (-9i-6) \\ (-6-3i-2i+1) + (2-6i-3i-9) + (-15+10i) \end{pmatrix}$$

$$= \begin{pmatrix} -10+5-5i-9-7i \\ 12+6i+11-3i-9i-6 \\ -5-5i-7-9i-15+10i \end{pmatrix} = \begin{pmatrix} -14-12i \\ 17-6i \\ -27-4i \end{pmatrix}.$$

Postscript: This is called matrix-vector multiplication in the language of linear algebra, but physics uses the terminology $\mathcal{A}$ **operates** on $|\,v>$. Since an operator operating on a ket is another ket, we can write

$$|\,w> = \mathcal{A}\,|\,v> = |\mathcal{A}\,v>,$$

where the last combinations of symbols is new but is clear because we know the product of an operator and a ket is a ket.

---

22. Find the vector-operator product $<v\,|\,\mathcal{A}^\dagger$ using $\mathcal{A}$ and $|\,v>$ from problem 21.

The inverse order of the two objects from problem 21, $|\,v>\mathcal{A}$, by itself is undefined. However, if the object on the left is a row vector, a bra, then each column of the operator can be treated as a ket for multiple multiplications like vector-vector multiplication. The result is a new bra vector. Symbolically,

$$\begin{aligned} <v\,|\,\mathcal{A} &= (\,b_1,\ b_2,\ b_3\,)\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ &= (\,b_1a_{11}+b_2a_{21}+b_3a_{31},\quad b_1a_{12}+b_2a_{22}+b_3a_{32},\quad b_1a_{13}+b_2a_{23}+b_3a_{33}\,) = \,<w\,|\,. \end{aligned}$$

Notice that this problem asks for the product of the bra of $|\,v>$ and the adjoint of $\mathcal{A}$.

$$\begin{aligned} <v\,|\,\mathcal{A}^\dagger &= (\,2-i,\ 1+3i,\ 3+2i\,)\begin{pmatrix} -4-2i & 6 & -3+i \\ 2-i & 2-3i & 2+3i \\ -1+3i & 3i & -5 \end{pmatrix} \\ &= \Big((2-i)(-4-2i)+(1+3i)(2-i)+(3+2i)(-1+3i), \\ &\qquad (2-i)(6)+(1+3i)(2-3i)+(3+2i)(3i), \\ &\qquad\qquad (2-i)(-3+i)+(1+3i)(2+3i)+(3+2i)(-5)\Big) \\ &= \Big((-8-4i+4i-2)+(2-i+6i+3)+(-3+9i-2i-6), \\ &\qquad (12-6i)+(2-3i+6i+9)+(9i-6), \\ &\qquad\qquad (-6+2i+3i+1)+(2+3i+6i-9)+(-15-10i)\Big) \\ &= \Big(-10+5+5i-9+7i,\quad 12-6i+11+3i+9i-6,\quad -5+5i-7+9i-15-10i\Big) \\ &= \Big(-14+12i,\quad 17+6i,\quad -27+4i\Big)\,. \end{aligned}$$

Postscript: An operator can operate to the right on a ket or to the left on a bra. Just as $|\,v>\mathcal{A}$ is undefined, the combination $\mathcal{A}<v\,|$ is also undefined.

We have chosen objects that are adjoints to those in problem 21. Notice that the result of problem 22 is the adjoint of the result of problem 21. Analogous operations with adjoints result in adjoint objects in the dual space.

Since an operator operating on a bra is another bra, we can write

$$<w\,| \,=\, <v\,|\,\mathcal{A}^\dagger = \,<v\,\mathcal{A}\,|\,,$$

where the last combinations of symbols is new. Notice that "†" is not included when the operator is inside the bra. An adjoint is understood for any object inside the "<" and the "|" that denote the bra.

---

23. Find the product $\mathcal{A}\mathcal{B}$ given that

$$\mathcal{A} = \begin{pmatrix} 2+5i & 3-2i \\ -1-4i & 4-i \end{pmatrix} \qquad \mathcal{B} = \begin{pmatrix} -3-2i & -1-6i \\ 4+5i & -2+3i \end{pmatrix}.$$

The product of two operators is another operator. For general $\mathcal{A}$ and $\mathcal{B}$ in three dimensions,

$$\mathcal{A}\mathcal{B} \;=\; \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} & a_{11}b_{13}+a_{12}b_{23}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} & a_{21}b_{13}+a_{22}b_{23}+a_{23}b_{33} \\ a_{31}b_{11}+a_{32}b_{21}+a_{33}b_{31} & a_{31}b_{12}+a_{32}b_{22}+a_{33}b_{32} & a_{31}b_{13}+a_{32}b_{23}+a_{33}b_{33} \end{pmatrix}.$$

This is essentially a generalization of operator-vector multiplication. It is like the second operator is composed of three column vectors. The operator-vector like product of each row of the first operator and the first column of the second operator becomes the first column of the new operator, the operator-vector like product of each row of the first operator and the second column of the second operator becomes the second column of the new operator, and so on. Operator-operator multiplication can be generalized to any including infinite dimension.

$$\mathcal{A}\mathcal{B} = \begin{pmatrix} 2+5i & 3-2i \\ -1-4i & 4-i \end{pmatrix} \begin{pmatrix} -3-2i & -1-6i \\ 4+5i & -2+3i \end{pmatrix}$$

$$= \begin{pmatrix} (2+5i)(-3-2i)+(3-2i)(4+5i) & (2+5i)(-1-6i)+(3-2i)(-2+3i) \\ (-1-4i)(-3-2i)+(4-i)(4+5i) & (-1-4i)(-1-6i)+(4-i)(-2+3i) \end{pmatrix}$$

$$= \begin{pmatrix} (-6-4i-15i+10)+(12+15i-8i+10) & (-2-12i-5i+30)+(-6+9i+4i+6) \\ (3+2i+12i-8)+(16+20i-4i+5) & (1+6i+4i-24)+(-8+12i+2i+3) \end{pmatrix}$$

$$= \begin{pmatrix} 4-19i+22+7i & 28-17i+13i \\ -5+14i+21+16i & -23+10i-5+14i \end{pmatrix} = \begin{pmatrix} 26-12i & 28-4i \\ 16+30i & -28+24i \end{pmatrix}.$$

---

24. For
$$|v> = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \qquad |w> = \begin{pmatrix} 1-i \\ 2+3i \end{pmatrix}, \qquad \text{and}$$
$$\mathcal{A} = \begin{pmatrix} 2 & -1 \\ -3 & 1 \end{pmatrix}, \qquad \mathcal{B} = \begin{pmatrix} 1+2i & -2-i \\ -2+i & 1-3i \end{pmatrix},$$

(a) find $<v\,|\,w>$ and $<w\,|\,v>$, then compare these two inner products.

(b) Find $\mathcal{A}\,|\,v>$ and $<v\,|\,\mathcal{A}^\dagger$. Compare these products.

(c) Find $\mathcal{A}^\dagger\,|\,v>$ and $<v\,|\,\mathcal{A}$. Compare these products. Compare the results of part (c) with those of part (b).

(d) Find $\mathcal{B}\,|\,w>$, and $<w\,|\,\mathcal{B}^\dagger$, and compare these products.

(e) Find $\mathcal{A}\,\mathcal{B}$ and $\mathcal{B}\,\mathcal{A}$, and compare these products.

(f) Find $\mathcal{B}^\dagger\,\mathcal{A}^\dagger$ and $\mathcal{A}^\dagger\,\mathcal{B}^\dagger$. Compare these products. Compare the results of part (f) with those of part (e).

This problem summarizes the different types of multiplication appropriate to a complex linear vector space. More importantly, this problem is designed to highlight some features of adjoint objects and dual spaces, and introduce the concept of **non-commutivity**. Commutivity is the algebraic property of multiplication that indicates the order of multiplication does not matter. If $a \cdot b = b \cdot a$, then a and b are said to commute. Scalars commute. Matrix operators do not generally commute. Order matters in matrix multiplication. Symbolically, $\mathcal{A}\,\mathcal{B} \neq \mathcal{B}\,\mathcal{A}$ in general.

You should find that $<v\,|\,w>$ and $<w\,|\,v>$ are complex conjugates in part (a). The operator-vector and vector-operator products of parts (b), (c), and (d) will be adjoints. You should see that there are no simple connections between the results of parts (b) and (c). There are no equalities nor are there any adjoints in spite of the similarity of objects being multiplied. Though there are two sets of adjoints in the results of parts (e) and (f), you should notice that $\mathcal{A}\,\mathcal{B} \neq \mathcal{B}\,\mathcal{A}$ and $\mathcal{A}^\dagger\,\mathcal{B}^\dagger \neq \mathcal{B}^\dagger\,\mathcal{A}^\dagger$. These operators do not commute. Finally, notice that for any product formed, if we take the adjoints of each factor and reverse the order of the multiplication, we attain the adjoint of the initial product. This is true for vector-vector, vector-operator, operator-vector, and operator-operator multiplication.

(a) An inner product is a scalar, and
$$<v\,|\,w> = \big(2,\ -1\big)\begin{pmatrix} 1-i \\ 2+3i \end{pmatrix} = 2\big(1-i\big) - 1\big(2+3i\big) = 2 - 2i - 2 - 3i = -5i\,.$$
$$<w\,|\,v> = \big(1+i,\ 2-3i\big)\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \big(1+i\big)2 + \big(2-3i\big)(-1) = 2 + 2i - 2 + 3i = 5i\,.$$
These two inner products are scalars that are complex conjugates.

(b) The operator-vector product indicated is
$$\mathcal{A}\,|\,v> = \begin{pmatrix} 2 & -1 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 2\cdot 2 + (-1)(-1) \\ (-3)2 + 1(-1) \end{pmatrix} = \begin{pmatrix} 4+1 \\ -6-1 \end{pmatrix} = \begin{pmatrix} 5 \\ -7 \end{pmatrix}.$$

The vector-operator product indicated is

$$<v\,|\,\mathcal{A}^\dagger = \Big(2,\ -1\Big)\begin{pmatrix} 2 & -3 \\ -1 & 1 \end{pmatrix} = \Big(2{\cdot}2{+}(-1)(-1),\ 2(-3){+}(-1)1\Big) = \Big(4{+}1,\ -6{-}1\Big) = \Big(5,\ -7\Big).$$

The results are adjoints, a ket and its corresponding bra.

(c) The operator-vector product is

$$\mathcal{A}^\dagger\,|\,v> = \begin{pmatrix} 2 & -3 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 2\cdot 2+(-3)(-1) \\ (-1)2+1(-1) \end{pmatrix} = \begin{pmatrix} 4+3 \\ -2-1 \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \end{pmatrix}.$$

The vector-operator product indicated is

$$<v\,|\,\mathcal{A} = \Big(2,\ -1\Big)\begin{pmatrix} 2 & -1 \\ -3 & 1 \end{pmatrix} = \Big(2{\cdot}2{+}(-1)(-3),\ 2(-1){+}(-1)1\Big) = \Big(4{+}3,\ -2{-}1\Big) = \Big(7,\ -3\Big).$$

The results are adjoints, a ket and its corresponding bra. Besides the fact that these are kets and bras in a two dimensional space, we see no simple connection between the vectors attained in parts (b) and (c). There are no results of parts (b) and (c) that are equal, nor are there any adjoints.

(d) The next operator-vector product is

$$\begin{aligned}
\mathcal{B}\,|\,w> &= \begin{pmatrix} 1+2i & -2-i \\ -2+i & 1-3i \end{pmatrix}\begin{pmatrix} 1-i \\ 2+3i \end{pmatrix} = \begin{pmatrix} (1+2i)(1-i)+(-2-i)(2+3i) \\ (-2+i)(1-i)+(1-3i)(2+3i) \end{pmatrix} \\
&= \begin{pmatrix} 1-i+2i+2-4-6i-2i+3 \\ -2+2i+i+1+2+3i-6i+9 \end{pmatrix} \\
&= \begin{pmatrix} 1+2-4+3+(-1+2-6-2)i \\ -2+1+2+9+(2+1+3-6)i \end{pmatrix} = \begin{pmatrix} 2-7i \\ 10 \end{pmatrix}.
\end{aligned}$$

The corresponding vector-operator product is

$$\begin{aligned}
<w\,|\,\mathcal{B}^\dagger &= \Big(1+i,\ 2-3i\Big)\begin{pmatrix} 1-2i & -2-i \\ -2+i & 1+3i \end{pmatrix} \\
&= \Big((1+i)(1-2i)+(2-3i)(-2+i),\ (1+i)(-2-i)+(2-3i)(1+3i)\Big) \\
&= \Big(1-2i+i+2-4+2i+6i+3,\ -2-i-2i+1+2+6i-3i+9\Big) \\
&= \Big(1+2-4+3+(-2+1+2+6)i,\ -2+1+2+9+(-1-2+6-3)i\Big) \\
&= \Big(2+7i,\ 10\Big),
\end{aligned}$$

and the two products are adjoint vectors, a ket and its corresponding bra.

(e) The indicated operator-operator product is

$$\begin{aligned}
\mathcal{A}\,\mathcal{B} &= \begin{pmatrix} 2 & -1 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 1+2i & -2-i \\ -2+i & 1-3i \end{pmatrix} \\
&= \begin{pmatrix} 2(1+2i)-1(-2+i) & 2(-2-i)-1(1-3i) \\ -3(1+2i)+1(-2+i) & -3(-2-i)+1(1-3i) \end{pmatrix} \\
&= \begin{pmatrix} 2+4i+2-i & -4-2i-1+3i \\ -3-6i-2+i & 6+3i+1-3i \end{pmatrix} = \begin{pmatrix} 4+3i & -5+i \\ -5-5i & 7 \end{pmatrix}.
\end{aligned}$$

$$\mathcal{B}\,\mathcal{A} = \begin{pmatrix} 1+2i & -2-i \\ -2+i & 1-3i \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -3 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} (1+2i)2 + (-2-i)(-3) & (1+2i)(-1) + (-2-i)1 \\ (-2+i)2 + (1-3i)(-3) & (-2+i)(-1) + (1-3i)1 \end{pmatrix}$$
$$= \begin{pmatrix} 2+4i+6+3i & -1-2i-2-i \\ -4+2i-3+9i & 2-i+1-3i \end{pmatrix} = \begin{pmatrix} 8+7i & -3-3i \\ -7+11i & 3-4i \end{pmatrix}.$$

As indicated earlier, you should notice that these operators do not commute, *i.e.*, $\mathcal{A}\mathcal{B} \neq \mathcal{B}\,\mathcal{A}$.

(f) The operator-operator product of the adjoints in inverse order is

$$\mathcal{B}^\dagger\,\mathcal{A}^\dagger = \begin{pmatrix} 1-2i & -2-i \\ -2+i & 1+3i \end{pmatrix}\begin{pmatrix} 2 & -3 \\ -1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} (1-2i)2 + (-2-i)(-1) & (1-2i)(-3) + (-2-i)1 \\ (-2+i)2 + (1+3i)(-1) & (-2+i)(-3) + (1+3i)1 \end{pmatrix}$$
$$= \begin{pmatrix} 2-4i+2+i & -3+6i-2-i \\ -4+2i-1-3i & 6-3i+1+3i \end{pmatrix} = \begin{pmatrix} 4-3i & -5+5i \\ -5-i & 7 \end{pmatrix}.$$

$$\mathcal{A}^\dagger\,\mathcal{B}^\dagger = \begin{pmatrix} 2 & -3 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1-2i & -2-i \\ -2+i & 1+3i \end{pmatrix}$$
$$= \begin{pmatrix} 2(1-2i) - 3(-2+i) & 2(-2-i) - 3(1+3i) \\ -1(1-2i) + 1(-2+i) & -1(-2-i) + 1(1+3i) \end{pmatrix}$$
$$= \begin{pmatrix} 2-4i+6-3i & -4-2i-3-9i \\ -1+2i-2+i & 2+i+1+3i \end{pmatrix} = \begin{pmatrix} 8-7i & -7-11i \\ -3+3i & 3+4i \end{pmatrix}.$$

Notice that $\mathcal{A}^\dagger\,\mathcal{B}^\dagger \neq \mathcal{B}^\dagger\,\mathcal{A}^\dagger$. These operators do not commute. $\mathcal{B}^\dagger\,\mathcal{A}^\dagger$ from part (f) is the adjoint of $\mathcal{A}\mathcal{B}$ from part (e), and $\mathcal{A}^\dagger\,\mathcal{B}^\dagger$ from part (f) is the adjoint of $\mathcal{B}\,\mathcal{A}$ from part (e).

Postscript: Some operators and even some classes of operators do commute. Operators that do commute are particularly important to quantum mechanical calculation.

Supplementary Problems

25. Show that $<v\,|\,w> = <w\,|\,v>^*$ in $\mathbf{C}^3$.

This problem is intended to convince you that the complex conjugate of an inner product is the same as the inner product in inverse order. If you can prove this in $\mathbf{C}^3$, you can see that the process is easily extended to larger dimension. Form two general vectors in $\mathbf{C}^3$, such as

$$|\,v> = \begin{pmatrix} a_1 + b_1 i \\ a_2 + b_2 i \\ a_3 + b_3 i \end{pmatrix} \qquad \text{and} \qquad |\,w> = \begin{pmatrix} c_1 + d_1 i \\ c_2 + d_2 i \\ c_3 + d_3 i \end{pmatrix}.$$

Form the inner products $<v\,|\,w>$ and $<w\,|\,v>^*$. Compare them. They will be identical.

---

26. If $\alpha = 3\,,\quad \beta = 5 - 2i\,,\quad <v\,| = \big(2-4i,\ -6,\ 2+3i\big)\,,$ and $|\,w> = \begin{pmatrix} 3+i \\ 4-3i \\ -2 \end{pmatrix},$ find

(a) $\alpha\,|\,v>\,,$

(b) $<w\,\beta\,|\,,$

(c) $\beta\,|\,v> + \,\alpha\,|\,w>\,,$ and

(d) $<w\,|\,v>$.

Remember to conjugate the components when forming a bra from a ket or a ket from a bra. Remember also that if a symbol is between the $<$ and $|$ of a bra, it is understood to be conjugated or adjoint, so $<w\,\beta\,| \; = <w\,|\,\beta^*$ in part (b).

---

27. Normalize

(a) $|\,1> = \begin{pmatrix} 1 \\ 0 \\ 1+i \end{pmatrix},$

(b) $|\,2> = \begin{pmatrix} -1+i \\ -i \\ 1 \end{pmatrix},$ and

(c) $|\,3> = \begin{pmatrix} 1+i \\ 3 \\ -i \end{pmatrix}.$

(d) Do the three normalized vectors constitute an orthonormal basis in $\mathbf{C}^3$?

Normalize each ket using the procedures of problem 10. Problem 15 may be helpful for part (d).

---

28. Given that the components of two-dimensional vectors are complex numbers, the definitions of vector addition and scalar multiplication

$$(a,\ b) + (c,\ d) = (2a + 2c,\ 2b + 2d) \qquad \text{and} \qquad \alpha\,(a,\ b) = (2\alpha a,\ 2\alpha b)$$

do not satisfy the requirements for a complex linear vector space. Which properties of a linear vector space are satisfied by these definitions, and which properties are not satisfied?

The definitions are known as double addition and double multiplication. Each component is a complex number, so you can employ the algebraic properties of complex numbers for component addition and multiplication. As an example of what is intended,

$$\Big[\,(a,\ b) + (c,\ d)\,\Big] + (e,\ f\,) = (2a + 2c,\ 2b + 2d) + (e,\ f\,) = (4a + 4c + 2e,\ 4b + 4d + 2f\,)\,, \quad \text{but}$$

$$(a,\ b)+\Big[(c,\ d)+(e,\ f)\Big]=(a,\ b)+(2c+2e,\ 2d+2f)=(2a+4c+4e,\ 2b+4d+4f),$$

and the results are not the same so double addition is not associative. You should find the definitions meet four of the remaining nine properties of a linear vector space. See problems 11, 12, and 13.

---

29. Which of the following sets of vectors are linearly independent?

(a) $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

(b) $\begin{pmatrix} i \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ -i \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix}$

(c) $\begin{pmatrix} i \\ i \\ 0 \end{pmatrix}, \quad \begin{pmatrix} i \\ i \\ i \end{pmatrix}, \quad \begin{pmatrix} i \\ 0 \\ i \end{pmatrix}$

(d) $\begin{pmatrix} -i \\ 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} i \\ -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

Use the procedures of problem 14. The vectors of parts (b) and (c) are linearly independent. The vectors of parts (a) and (d) are linearly dependent.

---

30. Show that the Pauli spin matrices are linearly independent.

The Pauli spin matrices exist in the subspace $\mathbf{C}^2$ of the general complex linear vector space $\mathbf{C}^\infty$. They are

$$\sigma_x=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y=\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_z=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

If they are linearly independent, they must satisfy the condition of linear independence,

$$a\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}+b\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}+c\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

if and only if $a=b=c=0$. Examine the last matrix equation and you will find that these conditions are satisfied. The Pauli spin matrices are focal in the study of spin angular momentum.

---

31. Are the following sets of three 2×2 matrices linearly independent?

(a) $\mathcal{A}=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \mathcal{B}=\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \mathcal{C}=\begin{pmatrix} -2 & -1 \\ 0 & -2 \end{pmatrix}$

(b) $\mathcal{D}=\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \quad \mathcal{E}=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \mathcal{F}=\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$

Problems 30 and 31 suggest that the concept of a vector space extends to objects not commonly identified as vectors. The ten properties of a linear vector space are all satisfied by 2×2 matrices with real elements using the procedures of matrix addition and scalar-matrix multiplication that have been defined. As such, 2×2 matrices with real elements form a linear vector space.

Assume linear independence. This means

$$a\,\mathcal{A} + \ b\,\mathcal{B} + \ c\,\mathcal{C} = |\,0> = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

for part (a). Solve the resulting simultaneous equations. The system is linearly independent if and only if the coefficients are all zero, *i.e.*, $a = b = c = 0$. If any of the coefficients is non-zero, you have contradicted the assumption and can conclude that the set is linearly dependent. Notice that the 2×2 zero matrix plays the role of zero in the space composed of 2×2 matrices.

---

32. Orthonormalize the following sets of vectors:

$$|\,1> = \begin{pmatrix} 1 \\ 0 \\ 1+i \end{pmatrix}, \quad |\,2> = \begin{pmatrix} -1+i \\ -i \\ 1 \end{pmatrix}, \quad |\,3> = \begin{pmatrix} 1+i \\ 3 \\ -i \end{pmatrix}.$$

You found in problem 15 that these vectors are orthogonal. Therefore, you need only to normalize each vector to attain an orthonormal basis.

---

33. (a) Show that the vectors

$$|\,I> = \begin{pmatrix} 0 \\ -i \\ 0 \end{pmatrix}, \quad |\,II> = \begin{pmatrix} i \\ 1+i \\ 0 \end{pmatrix}, \quad |\,III> = \begin{pmatrix} 1-i \\ 0 \\ 2+i \end{pmatrix} \quad \text{are linearly independent.}$$

(b) Show that the system of three vectors in part (a) is not orthogonal.

(c) Use the Gram-Schmidt process to form an orthonormal basis from the vectors in part (a).

(d) Show the resulting vectors are orthonormal.

This problem is the complex number analogy to problem 18. See problem 18 for procedures. The actual value of this problem and problem 18 is to reinforce that linear independence is a sufficient condition for orthonormality. The Gram-Schmidt procedure is a mathematical mechanism through which an orthonormal set can be realized given a linearly independent set. Orthonormality is a practical necessity for quantum mechanical calculations. Linear independence is the precondition.

---

34. Given that

$$\mathcal{A} = \begin{pmatrix} -3+4i & 5+2i & -2-i \\ 2 & 1+4i & -6i \\ -1-2i & 4-2i & -8 \end{pmatrix} \quad \text{and} \quad \mathcal{B} = \begin{pmatrix} 2-2i & 3-i & -4+5i \\ 3-2i & 5+i & 3-i \\ -4-3i & 5 & -1+5i \end{pmatrix}, \quad \text{find}$$

(a) $\mathcal{A}+\mathcal{B}$,

(b) $\mathcal{A}-\mathcal{B}$,

(c) $\mathcal{A}^\dagger+\mathcal{B}^\dagger$, and

(d) $\big(\mathcal{A}+\mathcal{B}\big)^\dagger$. Compare the results from parts (c) and (d).

See problem 20 for both intent and procedures.

---

35. Find the product $\mathcal{A}\mathcal{B}$ given that

$$\mathcal{A}=\begin{pmatrix} 3+2i & 4-i \\ -5-2i & 3-2i \end{pmatrix} \qquad \mathcal{B}=\begin{pmatrix} -2-7i & -3-2i \\ 2+4i & -3+6i \end{pmatrix}.$$

See problem 23 for both intent and procedures.

---

36. For

$$|\,v> = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \qquad |\,w> = \begin{pmatrix} 3-i \\ 2+4i \end{pmatrix}, \qquad \text{and}$$

$$\mathcal{A}=\begin{pmatrix} 3 & -4 \\ -1 & 6 \end{pmatrix}, \qquad \mathcal{B}=\begin{pmatrix} 3+i & -5-3i \\ -3+2i & 4-i \end{pmatrix},$$

(a) find $<v\,|\,w>$ and $<w\,|\,v>$, then compare these two inner products.

(b) Find $\mathcal{A}\,|\,v>$ and $<v\,|\,\mathcal{A}^\dagger$. Compare these products.

(c) Find $\mathcal{A}^\dagger\,|\,v>$ and $<v\,|\,\mathcal{A}$. Compare these products. Compare the results of part (c) with those of part (b).

(d) Find $\mathcal{B}\,|\,w>$, and $<w\,|\,\mathcal{B}^\dagger$, and compare these products.

(e) Find $\mathcal{A}\mathcal{B}$ and $\mathcal{B}\mathcal{A}$, and compare these products.

(f) Find $\mathcal{B}^\dagger\mathcal{A}^\dagger$ and $\mathcal{A}^\dagger\mathcal{B}^\dagger$. Compare these products. Compare the results of part (f) with those of part (e).

See problem 24 for intent and procedures.

---

37. Show that $<v\,|\,\mathcal{A}^\dagger$ is the adjoint of $\mathcal{A}\,|\,v>$ in $\mathbf{C}^2$.

Problem 24 indicates that the adjoints multiplied in the inverse order yields the adjoint of the original product. Problems 24 and 36 provide numerical examples. Problem 25 illustrates this fact in general for inner products. This problem is intended to illustrate this fact for operator-vector and vector-operator products. If you can do this in two dimensions, you can easily extend the process to higher or infinite dimension.

Use an arbitrary vector and operator like

$$|\,v\!> = \begin{pmatrix} a_1 + b_1 i \\ a_2 + b_2 i \end{pmatrix} \qquad \text{and} \qquad \mathcal{A} = \begin{pmatrix} c_1 + d_1 i & e_1 + f_1 i \\ c_2 + d_2 i & e_2 + f_2 i \end{pmatrix}.$$

Form $\mathcal{A}\,|\,v\!>$ and $<\!v\,|\,\mathcal{A}^\dagger$ and you will find that they are the same.

Operators are a primary subject of part 2. The operator-operator analogy of adjoints multiplied in the inverse order yielding the adjoint of the original product, $\big(\mathcal{A}\,\mathcal{B}\big)^\dagger = \mathcal{B}^\dagger\,\mathcal{A}^\dagger$, is addressed in part 2.

Tapestries from India, oriental vases, paintings by Dali, is that a Fabrege egg? Cases and stands made of ebony that matched the woodwork of the door frames, all intricately carved and highly polished. Is this a house or a gallery? He did not recognize most of the other guests; he did recognize a couple of ball players, a judge, and that beautiful lady who was a news reporter on channel 9. The guest list appeared to be affluent business people, successful entertainers, elected officials, appointed dignitaries...and a private eye who felt like a pair of unmatched socks in the silver drawer. Where can I get a beer? A petite maid in a black and white uniform appeared holding a tray of wine glasses and asked "Champagne?" Will have to do until I can find a beer..."Thank you." She added while turning toward adjacent guests, "The real values are those of Hermite."

The Mathematics of Quantum Mechanics, Part 2

Operators dominate part 2. When you finish part 2, you should understand the meaning of a **linear operator**, the **identity operator**, **inverse operators**, **Hermitian operators**, **unitary operators**, the **commutator**, a **diagonal** operator, and the **determinant** of an operator. You need to understand the meaning of **eigenvalues** and **eigenvectors**. You need to know how to solve the **eigenvalue/eigenvector** problem. You want to know how to **diagonalize** an operator. You want to know how to **simultaneously diagonalize** two operators. You want to understand the meaning of **degenerate operators**. The reason to understand simultaneous diagonalization is degenerate operators.

---

1. Show that $\mathcal{L}_y = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$ is linear.

A state vector, or wave function, will describe a quantum mechanical system. An observable quantity, or simply an **observable** is described by an operator. Examples of observable quantities are energy, momentum, and position. *All* the information about the *possible* outcomes we *could* observe is contained in the operator, without reference to a specific state vector or system.

Quantum mechanics appears to be a linear theory. In particular, we will informally state that quantum mechanics is dominated by **linear operators**. Specifically, *a linear operator commutes with scalars, and will distribute over vectors.* Symbolically this means

$$\Omega \Big[\alpha \, | \, v> \Big] = \alpha \, \Omega \, | \, v> , \tag{1}$$

$$\Omega \Big[\alpha \, | \, v_1> + \beta \, | \, v_2> \Big] = \alpha \, \Omega \, | \, v_1> + \beta \, \Omega \, | \, v_2> , \tag{2}$$

where the last equation is focal because it includes the operation described in the first equation.

Show that $\mathcal{L}_y$ satisfies these two equations. Since $\mathcal{L}_y$ is a three dimensional operator, pick generic three dimensional vectors like

$$| \, v_1> = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} \qquad \text{and} \qquad | \, v_2> = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix},$$

and do the appropriate multiplications and additions to show that equations (1) and (2) are true. It may be wise to start on both ends and meet in the middle, which is how most proofs are actually completed.

$$\mathcal{L}_y\Big[\,\alpha\,|\,v\!>\Big] = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \left[\alpha \begin{pmatrix} a \\ b \\ c \end{pmatrix}\right] = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \begin{pmatrix} \alpha a \\ \alpha b \\ \alpha c \end{pmatrix} = \begin{pmatrix} -i\alpha b \\ i\alpha a - i\alpha c \\ i\alpha b \end{pmatrix}$$

$$= \begin{pmatrix} \alpha(-ib) \\ \alpha(ia) + \alpha(-ic) \\ \alpha(ib) \end{pmatrix} = \alpha \begin{pmatrix} -ib \\ ia - ic \\ ib \end{pmatrix} = \alpha \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \alpha\,\mathcal{L}_y\,|\,v\!>,$$

so $\mathcal{L}_y$ satisfies equation (1). Seeing a three dimensional matrix in a vector to arrive at the next to last step is somewhat subtle, and is why we advised you to start at both ends. Equation (2) is

$$\mathcal{L}_y\Big[\,\alpha\,|\,v_1\!> + \,\beta\,|\,v_2\!>\Big] = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \left[\alpha \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} + \beta \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}\right]$$

$$= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \left[\begin{pmatrix} \alpha a_1 \\ \alpha b_1 \\ \alpha c_1 \end{pmatrix} + \begin{pmatrix} \beta a_2 \\ \beta b_2 \\ \beta c_2 \end{pmatrix}\right] = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \begin{pmatrix} \alpha a_1 \\ \alpha b_1 \\ \alpha c_1 \end{pmatrix} + \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \begin{pmatrix} \beta a_2 \\ \beta b_2 \\ \beta c_2 \end{pmatrix}$$

and this is apparently two applications of the reduction for equation (1). We will continue for the purpose of completeness, but realize that the rest of this reduction duplicates the reduction of equation (1) twice. Proceeding,

$$\mathcal{L}_y\Big[\,\alpha\,|\,v_1\!> + \,\beta\,|\,v_2\!>\Big] = \begin{pmatrix} -i\alpha b_1 \\ i\alpha a_1 - i\alpha c_1 \\ i\alpha b_1 \end{pmatrix} + \begin{pmatrix} -i\beta b_2 \\ i\beta a_2 - i\beta c_2 \\ i\beta b_2 \end{pmatrix}$$

$$= \begin{pmatrix} \alpha(-ib_1) \\ \alpha(ia_1) + \alpha(-ic_1) \\ \alpha(ib_1) \end{pmatrix} + \begin{pmatrix} \beta(-ib_2) \\ \beta(ia_2) + \beta(-ic_2) \\ \beta(ib_2) \end{pmatrix} = \alpha \begin{pmatrix} -ib_1 \\ ia_1 - ic_1 \\ ib_1 \end{pmatrix} + \beta \begin{pmatrix} -ib_2 \\ ia_2 - ic_2 \\ ib_2 \end{pmatrix}$$

$$= \alpha \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} + \beta \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} = \alpha\,\mathcal{L}_y\,|\,v_1\!> + \,\beta\,\mathcal{L}_y\,|\,v_2\!>\ .$$

---

2. Show by explicit multiplication in $\mathbf{C}^3$ that an identity operator multiplying

(a) a ket,

(b) a bra,

(c) a matrix operator from the left, and

(d) a matrix operator from the right result in the original ket, bra, or matrix operator.

The identity operator is the analogy of "1" in the real number system. One times a number is the original number. This problem introduces the identity operator but also intends to persuade you that any legitimate multiplication by the identity operator results in the original object. In $\mathbf{C}^3$ the identity operator is

$$\mathcal{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

An identity operator is a matrix with 1's on the principal diagonal and zeros elsewhere. Assume arbitrary objects like

$$|\, v > = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad \text{and} \quad \mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

You will find that $\mathcal{I}\,|\, v> = \,|\, v>, \quad <v\,|\,\mathcal{I} = \,<v\,|, \quad \mathcal{I}\,\mathcal{A} = \mathcal{A}, \text{ and } \mathcal{A}\mathcal{I} = \mathcal{A}.$

(a) $$\mathcal{I}\,|\, v> = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1+0+0 \\ 0+b_2+0 \\ 0+0+b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \,|\, v> .$$

(b) $$<v\,|\,\mathcal{I} = \big(b_1,\ b_2,\ b_3\big) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \big(b_1{+}0{+}0,\ 0{+}b_2{+}0,\ 0{+}0{+}b_3\big) = \big(b_1,\ b_2,\ b_3\big) = \,<v\,|\,.$$

(c) $$\mathcal{I}\,\mathcal{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}+0+0 & a_{12}+0+0 & a_{13}+0+0 \\ 0+a_{21}+0 & 0+a_{22}+0 & 0+a_{23}+0 \\ 0+0+a_{31} & 0+0+a_{32} & 0+0+a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \mathcal{A}\,.$$

(d) $$\mathcal{A}\mathcal{I} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}+0+0 & 0+a_{12}+0 & 0+0+a_{13} \\ a_{21}+0+0 & 0+a_{22}+0 & 0+0+a_{23} \\ a_{31}+0+0 & 0+a_{32}+0 & 0+0+a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \mathcal{A}\,.$$

Postscript: A scalar times an operator, from either side, is another operator. Symbolically,

$$\alpha\,\mathcal{I} = \alpha \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix},$$

the product is an operator with the scalar on the principal diagonal and zeros elsewhere. Operating on a vector or operator with this yields the same result as multiplication of the vector or operator by the scalar.

The scalar times the identity operator is what is intended when an author sets a scalar equal to an operator. For instance,

$$\mathcal{A}\mathcal{B} = \alpha \quad \text{means} \quad \mathcal{A}\mathcal{B} = \alpha\mathcal{I}\,.$$

---

3. Calculate the determinants of

$$\text{(a)} \qquad \mathcal{A} = \begin{pmatrix} 1-2i & 2+3i \\ 3-i & 4+2i \end{pmatrix}, \qquad \text{and} \qquad \text{(b)} \qquad \mathcal{C} = \begin{pmatrix} 2 & -4 & 7 \\ 8 & -3 & -5 \\ -4 & 9 & 1 \end{pmatrix}.$$

An integral with limits is a scalar, (though it may take considerable effort to calculate that scalar). Similarly, a **determinant** is a scalar. $\det \mathcal{A}$ is a scalar associated with a matrix operator. A determinant is a scalar associated with a square matrix. Specifically,

a determinant is a function of a matrix that is the sum of the products of the elements of any row or column and their respective cofactors.

$$\text{Symbolically} \qquad \mathcal{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \Rightarrow \quad \det \mathcal{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}, \qquad \text{and}$$

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \Rightarrow \quad \det \mathcal{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

$$= a_{11}\big(a_{22}a_{33} - a_{32}a_{23}\big) - a_{21}\big(a_{12}a_{33} - a_{32}a_{13}\big) + a_{31}\big(a_{12}a_{23} - a_{22}a_{13}\big)$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} + a_{21}a_{32}a_{13} - a_{21}a_{12}a_{33} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}.$$

This can be extended to arbitrary or infinite dimension. For evaluation of determinants beyond dimension 3, refer to Arfken[1], Boas[2], or your favorite linear algebra text.

$$\text{(a)} \qquad \mathcal{A} = \begin{pmatrix} 1-2i & 2+3i \\ 3-i & 4+2i \end{pmatrix} \quad \Rightarrow \quad \det \mathcal{A} = \begin{vmatrix} 1-2i & 2+3i \\ 3-i & 4+2i \end{vmatrix} = (1-2i)(4+2i) - (3-i)(2+3i)$$

$$= 4 + 2i - 8i + 4 - (6 + 9i - 2i + 3) = 8 - 6i - 6 - 9i + 2i - 3 = -1 - 13i\,,$$

which is a scalar in $\mathbf{C}^2$.

$$\text{(b)} \qquad \mathcal{C} = \begin{pmatrix} 2 & -4 & 7 \\ 8 & -3 & -5 \\ -4 & 9 & 1 \end{pmatrix} \quad \Rightarrow \quad \det \mathcal{C} = \begin{vmatrix} 2 & -4 & 7 \\ 8 & -3 & -5 \\ -4 & 9 & 1 \end{vmatrix}$$

---

[1] Arfken *Mathematical Methods for Physicists*, Academic Press, 19–, chap 4.
[2] Boas *Mathematical Methods in the Physical Sciences*, John Wiley & Sons, 1983, pp. 87–94.

$$= (2)(-3)(1) - (2)(9)(-5) + (8)(9)(7) - (8)(-4)(1) + (-4)(-4)(-5) - (-4)(-3)(7)$$
$$= -6 + 90 + 504 + 32 - 80 - 84 = 456 \qquad \text{which is a scalar in } \mathbf{R}^3.$$

Postscript: A determinant is a scalar associated only with a square matrix. A determinant is not defined for a matrix that does not have an equal number of rows and columns.

A **singular** matrix is one for which $\det \mathcal{A} = 0$. A singular operator does not have an **inverse** operator. We will address inverse operators directly.

Evaluating determinants is a skill that is central to solving the eigenvalue/eigenvector problem.

---

4. For an arbitrary matrix operator in $\mathbf{C}^2$,

(a) Show that $\det \mathcal{A}^T = \det \mathcal{A}$.

(b) Show that $\det \big(\mathcal{A}\mathcal{A} \big) = \det \mathcal{A} \det \mathcal{A}$.

(c) Show that $\det \big(\mathcal{B}^\dagger \big) = \big(\det \mathcal{B}^T \big)^*$ using explicitly complex matrix elements.

This problem is designed to increase your familiarity with determinants and some of their properties. Use a two-dimensional matrix like $\mathcal{A} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. The condition $\in \mathbf{C}^2$ means that α, β, γ, and δ are arbitrary complex numbers. For parts (a) and (b), just calculate the necessary operators and their determinants. In part (c), explicitly complex elements means $\mathcal{B} = \begin{pmatrix} \alpha + ia & \beta + ib \\ \gamma + ic & \delta + id \end{pmatrix}$, for instance. Part (c) is straightforward using matrix elements with real and imaginary parts; it may be less accessible using implicitly complex operator elements.

(a) $\det \mathcal{A} = \det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \alpha\delta - \beta\gamma$ and $\det \mathcal{A}^T = \det \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} = \alpha\delta - \gamma\beta = \alpha\delta - \beta\gamma$

$$\text{therefore,} \quad \det \mathcal{A}^T = \det \mathcal{A}.$$

(b)
$$\mathcal{A}\mathcal{A} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \alpha^2 + \beta\gamma & \alpha\beta + \beta\delta \\ \alpha\gamma + \gamma\delta & \beta\gamma + \delta^2 \end{pmatrix}$$

$$\begin{aligned} \det \big(\mathcal{A}\mathcal{A} \big) &= (\alpha^2 + \beta\gamma)(\beta\gamma + \delta^2) - (\alpha\beta + \beta\delta)(\alpha\gamma + \gamma\delta) \\ &= \cancel{\alpha^2\beta\gamma} + \alpha^2\delta^2 + \beta^2\gamma^2 + \cancel{\beta\gamma\delta^2} - \cancel{\alpha^2\beta\gamma} - \alpha\beta\gamma\delta - \alpha\beta\gamma\delta - \cancel{\beta\gamma\delta^2} \\ &= \alpha^2\delta^2 + \beta^2\gamma^2 - 2\alpha\beta\gamma\delta. \end{aligned}$$

$$\begin{aligned} \text{Then } \det \mathcal{A} \det \mathcal{A} &= (\alpha\delta - \beta\gamma)(\alpha\delta - \beta\gamma) \\ &= \alpha^2\delta^2 - \alpha\beta\gamma\delta - \alpha\beta\gamma\delta + \beta^2\gamma^2 \\ &= \alpha^2\delta^2 + \beta^2\gamma^2 - 2\alpha\beta\gamma\delta, \end{aligned}$$

therefore, $\det \big(\mathcal{A}\mathcal{A} \big) = \det \mathcal{A} \det \mathcal{A}$. In general, $\det \, (\mathcal{A}\mathcal{B}) = \det \mathcal{A} \det \mathcal{B}$.

(c)
$$\mathcal{B} = \begin{pmatrix} \alpha + ia & \beta + ib \\ \gamma + ic & \delta + id \end{pmatrix} \quad \Rightarrow \quad \mathcal{B}^T = \begin{pmatrix} \alpha + ia & \gamma + ic \\ \beta + ib & \delta + id \end{pmatrix}$$
$$\begin{aligned} \text{so} \quad \det \mathcal{B}^T &= (\alpha + ia)(\delta + id) - (\gamma + ic)(\beta + ib) \\ &= \alpha\delta + i\alpha d + ia\delta - ad - \beta\gamma - ib\gamma - i\beta c + bc \\ &= \alpha\delta - ad - \beta\gamma + bc + i(\alpha d + a\delta - b\gamma - \beta c) \\ \Rightarrow \quad (\det \mathcal{B}^T)^* &= \alpha\delta - ad - \beta\gamma + bc - i(\alpha d + a\delta - b\gamma - \beta c)\,. \end{aligned}$$

$$\begin{aligned} \det \mathcal{B}^\dagger &= \det \begin{pmatrix} \alpha - ia & \gamma - ic \\ \beta - ib & \delta - id \end{pmatrix} \\ &= (\alpha - ia)(\delta - id) - (\gamma - ic)(\beta - ib) \\ &= \alpha\delta - i\alpha d - ia\delta - ad - \beta\gamma + ib\gamma + i\beta c + bc \\ &= \alpha\delta - ad - \beta\gamma + bc - i(\alpha d + a\delta - b\gamma - \beta c)\,, \text{ so } \det \mathcal{B}^\dagger = (\det \mathcal{B}^T)^*\,. \end{aligned}$$

Postscript: A determinant is a scalar so has the properties of scalar. For instance, determinants commute because scalars commute.

You have demonstrated that the determinant of an operator and its transpose are the same, that the determinant of a product is the same as the product of the determinants, and that the determinant of an adjoint is the conjugate of the determinant of the transpose. Parts (a) and (c) together imply that $\det\left(\mathcal{B}^\dagger\right) = \left(\det \mathcal{B}\right)^*$, the determinant of the adjoint is the same as the conjugate of the determinant. These are not central results, but they can be useful (see problem 13). Note that these results can be extended to operators of dimension larger than 2×2.

---

5. Find the inverse of $\mathcal{D} = \begin{pmatrix} 1 & -2 \\ 4 & -2 \end{pmatrix}$, and then find $\mathcal{D}\,\mathcal{D}^{-1}$ and $\mathcal{D}^{-1}\,\mathcal{D}$ to verify your result.

The product of an operator and its inverse is the identity operator. Symbolically, the **inverse operator** of $\mathcal{A}$ is denoted $\mathcal{A}^{-1}$, and
$$\mathcal{A}\,\mathcal{A}^{-1} = \mathcal{A}^{-1}\,\mathcal{A} = \mathcal{I}\,.$$
Multiplication by an inverse is analogous to division. Calculating an inverse is an exercise in solving simultaneous equations. For example, if $\mathcal{A} = \begin{pmatrix} 5 & 4 \\ 2 & 2 \end{pmatrix}$, to find $\mathcal{A}^{-1}$,
$$\begin{pmatrix} 5 & 4 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \begin{array}{c} 5a + 4c = 1 \\ 2a + 2c = 0 \\ \text{and} \\ 5b + 4d = 0 \\ 2b + 2d = 1 \end{array} \quad \begin{array}{c} \Rightarrow \quad \begin{array}{c} 5a + 4c = 1 \\ -4a - 4c = 0 \end{array} \quad \Rightarrow \quad a = 1 \\ \\ \Rightarrow \quad \begin{array}{c} 5b + 4d = 0 \\ -4b - 4d = -2 \end{array} \quad \Rightarrow \quad b = -2 \end{array}$$
and substituting these into the second and fourth equations,
$$\begin{array}{lllll} a = 1 & \Rightarrow & 2(1) + 2c = 0 & \Rightarrow & c = -1\,, \\ b = -2 & \Rightarrow & 2(-2) + 2d = 1 & \Rightarrow & d = 5/2\,, \end{array} \quad \Rightarrow \quad \mathcal{A}^{-1} = \begin{pmatrix} 1 & -2 \\ -1 & 5/2 \end{pmatrix}.$$

To verify this result,

$$\mathcal{A}\,\mathcal{A}^{-1} = \begin{pmatrix} 5 & 4 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -1 & 5/2 \end{pmatrix} = \begin{pmatrix} 5-4 & -10+10 \\ 2-2 & -4+5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathcal{I}\,,$$

$$\mathcal{A}^{-1}\,\mathcal{A} = \begin{pmatrix} 1 & -2 \\ -1 & 5/2 \end{pmatrix} \begin{pmatrix} 5 & 4 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 5-4 & 4-4 \\ -5+5 & -4+5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathcal{I}\,.$$

$$\begin{pmatrix} 1 & -2 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \begin{array}{ccccc} \begin{array}{c} a-2c=1 \\ 4a-2c=0 \end{array} & \Rightarrow & \begin{array}{c} a-2c=1 \\ -4a+2c=0 \end{array} & \Rightarrow & a=-\dfrac{1}{3} \\ \text{and} & & & & \\ \begin{array}{c} b-2d=0 \\ 4b-2d=1 \end{array} & \Rightarrow & \begin{array}{c} -b+2d=0 \\ 4b-2d=1 \end{array} & \Rightarrow & b=\dfrac{1}{3} \end{array}$$

and substituting these into the first and third equations,

$$\begin{array}{l} a=-\dfrac{1}{3} \quad \Rightarrow \quad -\dfrac{1}{3}-2c=1 \quad \Rightarrow \quad c=-\dfrac{2}{3}\,, \\ b=\dfrac{1}{3} \quad \Rightarrow \quad \dfrac{1}{3}-2d=0 \quad \Rightarrow \quad d=\dfrac{1}{6}\,, \end{array} \quad \Rightarrow \quad \mathcal{D}^{-1} = \begin{pmatrix} -1/3 & 1/3 \\ -2/3 & 1/6 \end{pmatrix}.$$

To verify this result,

$$\mathcal{D}\,\mathcal{D}^{-1} = \begin{pmatrix} 1 & -2 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} -1/3 & 1/3 \\ -2/3 & 1/6 \end{pmatrix} = \begin{pmatrix} -1/3+4/3 & 1/3-1/3 \\ -4/3+4/3 & 4/3-1/3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathcal{I}\,,$$

$$\mathcal{D}^{-1}\,\mathcal{D} = \begin{pmatrix} -1/3 & 1/3 \\ -2/3 & 1/6 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 4 & -2 \end{pmatrix} = \begin{pmatrix} -1/3+4/3 & 2/3-2/3 \\ -2/3+2/3 & 4/3-1/3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathcal{I}\,.$$

Postscript: The fact that you want to retain from this problem is $\mathcal{A}\,\mathcal{A}^{-1} = \mathcal{A}^{-1}\,\mathcal{A} = \mathcal{I}$. We will use this fact frequently. The numerical examples of this problem are intended to be illustrative, but they are not important. We will never explicitly calculate the inverse of a matrix operator.

If an operator is singular, *i.e.*, if $\det\mathcal{A} = 0$, the operator does not have an inverse. The products $\mathcal{A}\,\mathcal{A}^{-1}$ and $\mathcal{A}^{-1}\,\mathcal{A}$ are, therefore, undefined. This condition is equivalent to prohibiting the division by zero in the real number system. We will employ inverse operators without explicitly determining that the operator is non-singular. In other words, we will consistently assume that an inverse of an operator exists. We make this assumption because the classes of operators common to quantum mechanics are typically non-singular.

---

6. Show that two dimensional matrix operators are associative.

Associativity is an often overlooked but frequently invoked property of matrix operators which says they can be grouped in any manner as long as the order does not change. Symbolically,

$$\mathcal{A}\,(\,\mathcal{B}\,\mathcal{C}\,) = (\,\mathcal{A}\,\mathcal{B}\,)\,\mathcal{C}\,.$$

By the way, $\mathcal{A}\,(\mathcal{B}\,\mathcal{C}\,)$ has the same meaning as $\mathcal{A}\,\mathcal{B}\,\mathcal{C}\,.$ Likely the simplest and clearest method of demonstrating associativity is using explicit forms of $\mathcal{A}\,,\ \mathcal{B}\,,$ and $\mathcal{C}\,,$ do the matrix multiplications and show that they are the same. An explicit form of $\mathcal{A}$ means

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

for instance. This problem is limited to two dimensions for reasons of simplicity. Your proof could be extended to any including infinite dimension.

We want to show

$$\mathcal{A}\,(\mathcal{B}\,\mathcal{C}\,) = (\mathcal{A}\,\mathcal{B}\,)\,\mathcal{C}\,. \tag{1}$$

The left side of the equation (1) is

$$\begin{aligned}\mathcal{A}\,(\mathcal{B}\,\mathcal{C}\,) &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\left[\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}\right] \\ &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11}c_{11}+b_{12}c_{21} & b_{11}c_{12}+b_{12}c_{22} \\ b_{21}c_{11}+b_{22}c_{21} & b_{21}c_{12}+b_{22}c_{22} \end{pmatrix}\end{aligned}$$

$$= \begin{pmatrix} a_{11}b_{11}c_{11}+a_{11}b_{12}c_{21}+a_{12}b_{21}c_{11}+a_{12}b_{22}c_{21} & a_{11}b_{11}c_{12}+a_{11}b_{12}c_{22}+a_{12}b_{21}c_{12}+a_{12}b_{22}c_{22} \\ a_{21}b_{11}c_{11}+a_{21}b_{12}c_{21}+a_{22}b_{21}c_{11}+a_{22}b_{22}c_{21} & a_{21}b_{11}c_{12}+a_{21}b_{12}c_{22}+a_{22}b_{21}c_{12}+a_{22}b_{22}c_{22} \end{pmatrix}$$

The right side of equation (1) is

$$\begin{aligned}(\mathcal{A}\,\mathcal{B}\,)\,\mathcal{C} &= \left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}\right]\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}\end{aligned}$$

$$= \begin{pmatrix} a_{11}b_{11}c_{11}+a_{12}b_{21}c_{11}+a_{11}b_{12}c_{21}+a_{12}b_{22}c_{21} & a_{11}b_{11}c_{12}+a_{12}b_{21}c_{12}+a_{11}b_{12}c_{22}+a_{12}b_{22}c_{22} \\ a_{21}b_{11}c_{11}+a_{22}b_{21}c_{11}+a_{21}b_{12}c_{21}+a_{22}b_{22}c_{21} & a_{21}b_{11}c_{12}+a_{22}b_{21}c_{12}+a_{21}b_{12}c_{22}+a_{22}b_{22}c_{22} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11}c_{11}+a_{11}b_{12}c_{21}+a_{12}b_{21}c_{11}+a_{12}b_{22}c_{21} & a_{11}b_{11}c_{12}+a_{11}b_{12}c_{22}+a_{12}b_{21}c_{12}+a_{12}b_{22}c_{22} \\ a_{21}b_{11}c_{11}+a_{21}b_{12}c_{21}+a_{22}b_{21}c_{11}+a_{22}b_{22}c_{21} & a_{21}b_{11}c_{12}+a_{21}b_{12}c_{22}+a_{22}b_{21}c_{12}+a_{22}b_{22}c_{22} \end{pmatrix}$$

when the second and third terms of each element are interchanged. This is the same as the first expansion. Since the two expansions equal the same thing, they must themselves be equal, or

$$\mathcal{A}\,(\mathcal{B}\,\mathcal{C}\,) = (\mathcal{A}\,\mathcal{B}\,)\,\mathcal{C}\,,\quad \text{and two dimensional matrix operators are associative.}$$

---

7. Show that $\left(\mathcal{A}\,\mathcal{B}\right)^{-1} = \mathcal{B}^{-1}\,\mathcal{A}^{-1}\,.$

This problem establishes the fact that the inverse of the product of two operators is the reverse order of the product of the inverses. It also introduces a technique known as **insertion of the identity** that is vital to many of the calculations of quantum mechanics. The identity operator is so named because anything upon which it operates remains identical. Use of the identity operator is analogous to multiplication by 1 in the real number system. The real utility of the identity operator (or any identity element including the number 1 in the real number system) is that it has an infinite number of different forms, $\mathcal{A}\mathcal{A}^{-1}$ and $\mathcal{B}\mathcal{B}^{-1}$ to name just two forms that are employed in this problem.

$\mathcal{I} = \mathcal{A}\mathcal{A}^{-1}$	statement of the identity
$\mathcal{I} = \mathcal{A}\mathcal{I}\mathcal{A}^{-1}$	insertion of the identity
$\mathcal{I} = \mathcal{A}\mathcal{B}\mathcal{B}^{-1}\mathcal{A}^{-1}$	use of the form of the identity $\mathcal{I} = \mathcal{B}\mathcal{B}^{-1}$
$\mathcal{I} = \left(\mathcal{A}\mathcal{B}\right)\mathcal{B}^{-1}\mathcal{A}^{-1}$	associative property
$\left(\mathcal{A}\mathcal{B}\right)^{-1}\mathcal{I} = \left(\mathcal{A}\mathcal{B}\right)^{-1}\left(\mathcal{A}\mathcal{B}\right)\mathcal{B}^{-1}\mathcal{A}^{-1}$	operate on both sides with $\left(\mathcal{A}\mathcal{B}\right)^{-1}$
$\left(\mathcal{A}\mathcal{B}\right)^{-1}\mathcal{I} = \mathcal{I}\mathcal{B}^{-1}\mathcal{A}^{-1}$	use of the form of the identity $\mathcal{I} = \left(\mathcal{A}\mathcal{B}\right)^{-1}\left(\mathcal{A}\mathcal{B}\right)$
$\left(\mathcal{A}\mathcal{B}\right)^{-1} = \mathcal{B}^{-1}\mathcal{A}^{-1}$	operation by the identity

---

Postscript: A short comment on lines 5 and 6 may be helpful. Remember that the product of two operators is another operator, so $\left(\mathcal{A}\mathcal{B}\right)$ is an operator, and its inverse is denoted $\left(\mathcal{A}\mathcal{B}\right)^{-1}$.

---

8. Show that the commutator of

$$\mathcal{A} = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix} \qquad \mathcal{B} = \begin{pmatrix} -1-i & 0 \\ 0 & -1+i \end{pmatrix} \quad \text{is zero.}$$

The order of the operators matters in matrix multiplication. Generally,

$$\mathcal{A}\mathcal{B} \neq \mathcal{B}\mathcal{A}, \qquad \text{or} \qquad \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} \neq 0.$$

The left side of the last equation is commonly denoted

$$\left[\mathcal{A},\, \mathcal{B}\right] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}.$$

The object $[\mathcal{A},\, \mathcal{B}]$ is called a **commutator**. Most of the physics of quantum mechanics can be generated using commutators. If $[\mathcal{A},\, \mathcal{B}] = 0$, the operators $\mathcal{A}$ and $\mathcal{B}$ are said to commute. Operators that represent observable quantities and commute become particularly focal.

Do the indicated multiplcations and you will find that the difference of the products is "zero." It is actually the zero operator. If all of the elements are zero, the operator is said to be zero.

$$\begin{aligned}
\left[\mathcal{A},\, \mathcal{B}\right] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} &= \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}\begin{pmatrix} -1-i & 0 \\ 0 & -1+i \end{pmatrix} - \begin{pmatrix} -1-i & 0 \\ 0 & -1+i \end{pmatrix}\begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix} \\
&= \begin{pmatrix} (1+i)(-1-i)+0 & 0+0 \\ 0+0 & 0+(1-i)(-1+i) \end{pmatrix} - \begin{pmatrix} (-1-i)(1+i)+0 & 0+0 \\ 0+0 & 0+(-1+i)(1-i) \end{pmatrix} \\
&= \begin{pmatrix} -1-i-i+1 & 0 \\ 0 & -1+i+i+1 \end{pmatrix} - \begin{pmatrix} -1-i-i+1 & 0 \\ 0 & -1+i+i+1 \end{pmatrix} \\
&= \begin{pmatrix} -2i & 0 \\ 0 & 2i \end{pmatrix} - \begin{pmatrix} -2i & 0 \\ 0 & 2i \end{pmatrix} = \begin{pmatrix} -2i+2i & 0 \\ 0 & 2i-2i \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.
\end{aligned}$$

Postscript: A commutator is itself an operator. The product of two operators is another operator, and the difference of two operators is an operator, so the difference of the products of two operators is an operator.

The operators in this problem are **diagonal operators**. A diagonal operator is a matrix that has non-zero elements only on the principal diagonal. The identity matrix is another example of a diagonal operator. All diagonal operators commute.

---

9. Is the operator $\mathcal{L}_y = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$ Hermitian?

A **Hermitian operator** is an operator that is identical to its adjoint, *i.e.*,

$$\mathcal{A} = \mathcal{A}^\dagger \,.$$

An adjoint is a transpose conjugate. Switch the rows and columns and conjugate all the elements.

The transpose of $\mathcal{L}_y$ is $\mathcal{L}_y^{\mathrm{T}} = \begin{pmatrix} 0 & i & 0 \\ -i & 0 & i \\ 0 & -i & 0 \end{pmatrix}$, and the conjugate of the transpose, or transpose conjugate, is the adjoint, or $\mathcal{L}_y^{\mathrm{T}\,*} = \mathcal{L}_y^\dagger = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} = \mathcal{L}_y \,,$ therefore, $\mathcal{L}_y$ is Hermitian.

Postscript: All observable quantities, such as energy, momentum, or position; are represented by Hermitian operators. *All observable quantities are represented by Hermitian operators.*

The operator $\mathcal{L}_y$ in this problem is one of the three component angular momentum operators.

An **anti-Hermitian operator** is an operator that is identical to the negative of its adjoint, *i.e.*,

$$\mathcal{A} = -\mathcal{A}^\dagger \,.$$

---

10. Is the operator $\mathcal{R} = \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ unitary?

A **unitary operator** is an operator whose product with its adjoint is the identity, *i.e.*,

$$\mathcal{U}\mathcal{U}^\dagger = \mathcal{U}^\dagger\mathcal{U} = \mathcal{I} \,.$$

Multiplication by a unitary operator preserves the norm, or in geometric terms, the length of a vector. A unitary operator is analogous to an identity operator with a rotation. Geometrically, operation by the identity operator changes neither the direction nor the length. Operation by a unitary operator does not change the length but does change the direction of a vector.

The adjoint of $\mathcal{R}$ is the tranpose conjugate, so

$$\mathcal{R}^{\mathrm{T}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad\Rightarrow\quad \mathcal{R}^{\mathrm{T}\,*} = \left(\frac{1}{\sqrt{2}}\right)^* \begin{pmatrix} 1^* & 1^* \\ -1^* & 1^* \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \mathcal{R}^{\dagger},$$

$$\Rightarrow \quad \mathcal{R}\,\mathcal{R}^{\dagger} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+1 & 1-1 \\ 1-1 & 1+1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathcal{I},$$

therefore $\mathcal{R}$ is unitary.

Postscript: If $\mathcal{U}\mathcal{U}^{\dagger} = \mathcal{I}$, then $\mathcal{U}^{\dagger}\mathcal{U} = \mathcal{I}$. Similarly, $\mathcal{U}^{\dagger}\mathcal{U} = \mathcal{I} \;\Rightarrow\; \mathcal{U}\mathcal{U}^{\dagger} = \mathcal{I}$. The proof is short and straightforward using techniques similar to problem 7. See problem 32.

The operator $\mathcal{R}$ in this problem rotates a vector clockwise through an angle $\pi/4$.

---

11. Show that if $\mathcal{A}$ and $\mathcal{B}$ are two-dimensional operators, $\left(\mathcal{A}\,\mathcal{B}\right)^{\dagger} = \mathcal{B}^{\dagger}\,\mathcal{A}^{\dagger}$.

This problem proves a useful result. Problem 12 is a more elegant reduction of this fact for unitary operators using operator algebra. Use

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad \text{and} \qquad \mathcal{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

for instance. Simply form the matrix products of the left and right side of $\left(\mathcal{A}\,\mathcal{B}\right)^{\dagger} = \mathcal{B}^{\dagger}\,\mathcal{A}^{\dagger}$, the adjoint of the product and the product of the adjoints in the reverse order, compare them, and they will be the same. Again, this proof can be easily extended to higher dimension.

$$\left(\mathcal{A}\,\mathcal{B}\right)^{\dagger} = \left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}\right]^{\dagger} = \left[\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}\right]^{\dagger}$$

$$= \begin{pmatrix} a_{11}^*b_{11}^* + a_{12}^*b_{21}^* & a_{21}^*b_{11}^* + a_{22}^*b_{21}^* \\ a_{11}^*b_{12}^* + a_{12}^*b_{22}^* & a_{21}^*b_{12}^* + a_{22}^*b_{22}^* \end{pmatrix}.$$

The product of the adjoints in the opposite order is

$$\mathcal{B}^{\dagger}\,\mathcal{A}^{\dagger} = \left[\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}\right]^{\dagger}\left[\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right]^{\dagger} = \begin{pmatrix} b_{11}^* & b_{21}^* \\ b_{12}^* & b_{22}^* \end{pmatrix}\begin{pmatrix} a_{11}^* & a_{21}^* \\ a_{12}^* & a_{22}^* \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}^*b_{11}^* + a_{12}^*b_{21}^* & a_{21}^*b_{11}^* + a_{22}^*b_{21}^* \\ a_{11}^*b_{12}^* + a_{12}^*b_{22}^* & a_{21}^*b_{12}^* + a_{22}^*b_{22}^* \end{pmatrix}.$$

This is, in fact, the same as the first expansion, therefore $\left(\mathcal{A}\,\mathcal{B}\right)^{\dagger} = \mathcal{B}^{\dagger}\,\mathcal{A}^{\dagger}$.

12. (a) Show that the adjoint of a non-singular unitary operator is equal to its inverse.

(b) Show that the product of unitary operators is unitary.

(c) Use operator algebra to show that if $\mathcal{A}$ and $\mathcal{B}$ are unitary, $\left(\mathcal{A}\,\mathcal{B}\right)^{\dagger} = \mathcal{B}^{\dagger}\,\mathcal{A}^{\dagger}$.

This problem is intended to provide the opportunity to use operator algebra. All three parts are intended to be completed without the use of an explicit matrix. You can operate on both sides of an equation with the same operator if you operate from the same side. Operation must be from the left on both sides or from the right on both sides because operators do not generally commute. Operator algebra allows insertion of the operator $\mathcal{I}$ in any circumstance that is convenient.

Part (a) is an application of the definitions of unitary and inverse relations, $\mathcal{A}\,\mathcal{A}^{\dagger} = \mathcal{I}$, and $\mathcal{A}\,\mathcal{A}^{-1} = \mathcal{I}$, for instance. To show that the product of any two unitary operators is unitary, start with the product of two arbitrary unitary operators, say $\mathcal{A}\,\mathcal{B} = \mathcal{C}$. Operate on both sides with appropriate operators to show that $\mathcal{C}\,\mathcal{C}^{\dagger} = \mathcal{I}$. Remember that you have $\left(\mathcal{A}\,\mathcal{B}\right)^{\dagger} = \mathcal{B}^{\dagger}\mathcal{A}^{\dagger}$ from problem 11. Part (c) is a more elegant determination of this fact for unitary operators. Start with $\mathcal{A}\,\mathcal{A}^{\dagger} = \mathcal{I}$ and insert the identity between the two operators. It is likely $\mathcal{B}\,\mathcal{B}^{\dagger}$ will be a convenient form for this problem. You will likely need the result of part (b) to complete part (c).

(a) If $\mathcal{A}$ is unitary, $\mathcal{A}\,\mathcal{A}^{\dagger} = \mathcal{I}$ and $\mathcal{A}\,\mathcal{A}^{-1} = \mathcal{I}$ for all non-singular operators, so

$\mathcal{A}\,\mathcal{A}^{\dagger} = \mathcal{A}\,\mathcal{A}^{-1}$	because both are equal to $\mathcal{I}$
$\mathcal{A}^{-1}\,\mathcal{A}\,\mathcal{A}^{\dagger} = \mathcal{A}^{-1}\,\mathcal{A}\,\mathcal{A}^{-1}$	operate from the left with $\mathcal{A}^{-1}$
$\mathcal{I}\,\mathcal{A}^{\dagger} = \mathcal{I}\,\mathcal{A}^{-1}$	because $\mathcal{A}^{-1}\,\mathcal{A} = \mathcal{I}$
$\mathcal{A}^{\dagger} = \mathcal{A}^{-1}$	which is the desired result.

(b) If $\mathcal{A}$ and $\mathcal{B}$ are unitary, then $\mathcal{A}\,\mathcal{A}^{\dagger} = \mathcal{B}\,\mathcal{B}^{\dagger} = \mathcal{I}$. Consider their product

$\mathcal{A}\;\mathcal{B} = \mathcal{C}$	the product of two operators is an operator
$\mathcal{A}\mathcal{B}\,\mathcal{B}^{\dagger} = \mathcal{C}\,\mathcal{B}^{\dagger}$	operate from the right with $\mathcal{B}^{\dagger}$
$\mathcal{A}\mathcal{I} = \mathcal{A} = \mathcal{C}\,\mathcal{B}^{\dagger}$	because $\mathcal{B}\,\mathcal{B}^{\dagger} = \mathcal{I}$
$\mathcal{A}\,\mathcal{A}^{\dagger} = \mathcal{C}\,\mathcal{B}^{\dagger}\,\mathcal{A}^{\dagger}$	operate from the right with $\mathcal{A}^{\dagger}$
$\mathcal{I} = \mathcal{C}\,\mathcal{B}^{\dagger}\mathcal{A}^{\dagger}$	because $\mathcal{A}\,\mathcal{A}^{\dagger} = \mathcal{I}$
$\mathcal{I} = \mathcal{C}\left(\mathcal{A}\mathcal{B}\right)^{\dagger}$	use of $\left(\mathcal{A}\,\mathcal{B}\right)^{\dagger} = \mathcal{B}^{\dagger}\mathcal{A}^{\dagger}$
$\mathcal{I} = \mathcal{C}\,\mathcal{C}^{\dagger}$	definition of $\mathcal{C}$

or $\mathcal{C}\,\mathcal{C}^{\dagger} = \mathcal{I}$, and therefore, the product of unitary operators is unitary.

(c) Given that $\mathcal{A}$ and $\mathcal{B}$ are unitary,

$$\begin{aligned}
\mathcal{A}\,\mathcal{A}^\dagger &= \mathcal{I} && \text{definition of unitary} \\
\mathcal{A}\,\mathcal{I}\,\mathcal{A}^\dagger &= \mathcal{I} && \text{insertion of the identity} \\
\mathcal{A}\,\mathcal{B}\,\mathcal{B}^\dagger\mathcal{A}^\dagger &= \mathcal{I} && \text{use of the form } \mathcal{I} = \mathcal{B}\,\mathcal{B}^\dagger \\
(\,\mathcal{A}\,\mathcal{B}\,)\,\mathcal{B}^\dagger\mathcal{A}^\dagger &= \mathcal{I} && \text{associative property} \\
(\,\mathcal{A}\,\mathcal{B}\,)^\dagger\,(\,\mathcal{A}\,\mathcal{B}\,)\,\mathcal{B}^\dagger\,\mathcal{A}^\dagger &= (\,\mathcal{A}\,\mathcal{B}\,)^\dagger\,\mathcal{I} && \text{operation from the left with } (\,\mathcal{A}\,\mathcal{B}\,)^\dagger \\
\mathcal{I}\,\mathcal{B}^\dagger\,\mathcal{A}^\dagger &= (\,\mathcal{A}\,\mathcal{B}\,)^\dagger && \text{because } (\,\mathcal{A}\,\mathcal{B}\,)^\dagger\,(\,\mathcal{A}\,\mathcal{B}\,) = \mathcal{I} \text{ from part (b)} \\
\mathcal{B}^\dagger\,\mathcal{A}^\dagger &= \left(\,\mathcal{A}\,\mathcal{B}\,\right)^\dagger && \text{or} \qquad \left(\,\mathcal{A}\,\mathcal{B}\,\right)^\dagger = \mathcal{B}^\dagger\,\mathcal{A}^\dagger\,.
\end{aligned}$$

---

13. Show that the determinant of any two-dimensional unitary matrix has unit modulus.

This problem establishes a property of unitary matrices just as problem 4 established three properties of determinants. This problem is fairly short and straightforward if you use the three results of problem 4 appropriately.

Show that $\mathcal{U}\,\mathcal{U}^\dagger = I$ implies $\det\left(\mathcal{U}\,\mathcal{U}^\dagger\,\right) = 1$. Then show that $\det\left(\mathcal{U}\,\mathcal{U}^\dagger\,\right) = \det\mathcal{U}\,\det\mathcal{U}^*$. Realize that $\det\mathcal{U}\,\det\mathcal{U}^*$ is the product of a complex number and its conjugate. Finally, remember that the modulus of a complex number is the square root of the product of the number and its complex conjugate.

For a unitary matrix,

$$\mathcal{U}\,\mathcal{U}^\dagger = \mathcal{I} \quad \Rightarrow \quad \det\left(\mathcal{U}\,\mathcal{U}^\dagger\,\right) = \det\mathcal{I} = 1\,,$$

so using previously developed results, we have

$$\begin{aligned}
\det\left(\mathcal{U}\,\mathcal{U}^\dagger\,\right) &= \det\mathcal{U}\,\det\mathcal{U}^\dagger && \text{from part (b) of problem 4} \\
&= \det\mathcal{U}\,\det\mathcal{U}^{T*} = \det\mathcal{U}\left(\det\mathcal{U}^T\,\right)^* && \text{from part (c) of problem 4} \\
&= \det\mathcal{U}\,\det\mathcal{U}^* = 1\,, && \text{from part (a) of problem 4}
\end{aligned}$$

The determinant is generally a complex scalar, so let $\det\mathcal{U} = a + ib \;\Rightarrow\; \det\mathcal{U}^* = a - ib$

$$\Rightarrow \quad \det\mathcal{U}\,\det\mathcal{U}^* = (a+ib)(a-ib) = a^2 + b^2 = 1 = |\det\mathcal{U}\,|^2\,,$$

$\Rightarrow \quad \det\mathcal{U} = \sqrt{a^2+b^2} = 1\,,$ therefore the determinant of a unitary matrix is of unit modulus.

---

14. Given the matrix operator $\begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$,

(a) show that $|\,2\!> \,= \begin{pmatrix} 1 \\ 1 \end{pmatrix}$,

(b) and that $\;|\,3\!> \,= \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

(c) Interpret these facts geometrically.

This problem is a numerical introduction to **eigenvalues** and **eigenvectors** and the **eigenvalue/eigenvector** equation. It also attempts to familiarize you with some popular notation.

The product of an operator and a vector is another vector, $\mathcal{A}\,|\,v> \;=\; |\,w>$. For every operator, there are products of the operator and special vectors such that the new vector is a product of a scalar and the original vector, *i.e.*,

$$\mathcal{A}\,|\,v_i\!> = \alpha_i\,|\,v_i\!> \;.$$

This eigenvalue/eigenvector equation actually describes a family of equations, thus subscripts are commonly used. There are as many matching scalars and vectors as the dimension of the space. The scalars are known as eigenvalues and the vectors are known as eigenvectors. It is popular to place the eigenvalue between the $|$ and $>$ that indicate the ket to identify the corresponding eigenvector. Thus, part (a) asks you to show that

$$\begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

which is simply a numerical statement of the eigenvalue/eigenvector equation. Consider "length" and "direction" of the vectors for part (c).

(a) $\begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+1 \\ -2+4 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, so for the operator $\mathcal{A} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$,

2 is an eigenvalue, and $|\,2\!> \,= \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is the corresponding eigenvector.

(b) $\begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1+2 \\ -2+8 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, so for the operator $\mathcal{A} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$,

3 is an eigenvalue, and $|\,3\!> \,= \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is the corresponding eigenvector.

In a 2 dimensional space, we expect exactly two eigenvalues and two eigenvectors.

(c) In part (a), the effect of the operation is a vector twice as long and in exactly the same direction as the original vector. In part (b), the effect of the operation is a vector three times as long as the original vector and in exactly the same direction. An operator acting on an eigenvector results in changing the length of the vector without rotating it. The length of the eigenvector is changed by a factor equal to its eigenvalue.

Postscript: Physicists use numerous other methods of denoting eigenvectors other than by placing the eigenvalue between the $|$ and $>$ that indicate the eigenvector or **eigenket**. For instance, $|\,v_1\!>$

or $|1>$ might identify the first eigenvector and $|v_2>$ or $|2>$ might denote the second eigenvector. Quantum numbers are often placed between the $|$ and $>$ that indicate the eigenket.

Also, be aware that the eigenvalue/eigenvector equation always denotes a family of equations. Subscripts are not always used. There are as many eigenvalues and eigenvectors as the dimension of the operator under consideration.

---

15. (a) Interpret the operation of $\begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ on the vector $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ geometrically.

(b) Show that the result of the operation of $\begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ on the vector $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ can be expressed as a linear combination of the eigenvectors given in problem 14.

An operator operating on an eigenvector results in an elongation or contraction of the vector. An operator operating on a vector that is not an eigenvector also includes some sort of rotation. There is generally an elongation or contraction unless the operator is unitary, but if the vector is not an eigenvector, there will always be something equivalent to a rotation.

Part (a) is intended that you recognize these facts. For instance, you should recognize that

$$\begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = a \begin{pmatrix} -1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 0 \\ 6 \end{pmatrix},$$

and interpret the result in a Cartesian coordinate system. Part (b) intends for you to solve

$$\begin{pmatrix} 0 \\ 6 \end{pmatrix} = a \begin{pmatrix} 1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

(a)
$$\begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1+1 \\ 2+4 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$$

which cannot be arranged into a scalar times the original vector. The new vector can be arranged

$$\begin{pmatrix} 0 \\ 6 \end{pmatrix} = \begin{pmatrix} -6 \\ 6 \end{pmatrix} + \begin{pmatrix} 6 \\ 0 \end{pmatrix} = 6 \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 6 \\ 0 \end{pmatrix}$$

which is a scalar times the original vector plus another vector. Given Cartesian coordinates, the first vector represents an elongation by a factor of six in the original direction, and the second vector represents 6 steps in the x direction which is a clockwise rotation around the origin.

(b)
$$\begin{pmatrix} 0 \\ 6 \end{pmatrix} = a \begin{pmatrix} 1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \Rightarrow \quad \begin{matrix} 0 = a+b \\ 6 = a+2b \end{matrix}.$$

The top equation says $a = -b$, and using this in the bottom equation yields $b = 6$ so $a = -6$, and

$$\begin{pmatrix} 0 \\ 6 \end{pmatrix} = -6 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 6 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

is a linear combination of $\begin{pmatrix} 0 \\ 6 \end{pmatrix}$ in terms of the eigenvectors of the operator.

Postscript: Operation on an eigenvector is analogous to an elongation or contraction without a rotation. Rotation is included as part of any operation on a vector that is not an eigenvector.

The linearly independent eigenvectors of this operator constitute a basis in $\mathbf{R}^2$. We will have frequent occasion to use the eigenvectors of an operator as a basis.

---

16. Find the eigenvalues and normalized eigenvectors of $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}$.

The eigenvalue/eigenvector problem is one of the preeminent calculations in quantum mechanics.

Here is a roadmap to solve an eigenvalue/eigenvector problem.

1) Set $\det\left(\mathcal{A} - \alpha\mathcal{I}\right) = 0$. This is known as the **characteristic equation**.
2) Solve the characteristic equation. The solutions are the eigenvalues.
3) Use the eigenvalue/eigenvector equation to solve for the eigenvectors.
4) Normalize the eigenvectors.

There are some conventions associated with the **eigenvalue/eigenvector** problem. We will work from the eigenvalue of least magnitude to the eigenvalue of greatest magnitude. The simultaneous equations that result from the eigenvalue/eigenvector equation are generally indeterminate. In this circumstance, we choose the first non-zero element of the eigenket to be positive and real, most often 1. This is a convention that is popular but not universal[3].

$$\det\Big(\mathcal{A} - \alpha\mathcal{I}\Big) = \det\left[\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} - \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}\right] = \det\begin{pmatrix} 1-\alpha & 1 & 2 \\ 0 & 2-\alpha & 0 \\ 0 & 1 & 3-\alpha \end{pmatrix}$$

$$= (1-\alpha)(2-\alpha)(3-\alpha) + 0 + 0 - 0 - 0 - 0 = (1-\alpha)(2-\alpha)(3-\alpha)$$

is the **characteristic polynomial**. The characteristic equation is formed by setting the characteristic polynomial equal to zero, or

$$(1-\alpha)(2-\alpha)(3-\alpha) = 0$$

which has the solutions $\alpha_i = 1,\ 2,$ and 3. These are the eigenvalues. For the eigenvalue of least magnitude, $\alpha_1 = 1$, the eigenvalue/eigenvector equation is

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 1\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{ccccccc} a & + & b & + & 2c & = & a \\ & & 2b & & & = & b \\ & & b & + & 3c & = & c \end{array}$$

---

[3] Shankar, *Principles of Quantum Mechanics* (Plenum Press, New York, 1994), 2nd ed., p. 34.

which is three equations in three unknowns. The middle equation implies $b = 0$. Using this in the bottom equation implies $c = 0$ also. Using both of these in the top equation yields $a = a$, which has an infinite number of solutions. This system is indeterminate. Following the convention of choosing the top element of the eigenket positive and real, we choose $a = 1$, so the eigenket is

$$|\,1\!> \,=\begin{pmatrix}1\\0\\0\end{pmatrix} \quad \text{for the eigenvalue} \quad \alpha_1 = 1.$$

This choice also has the advantage that the eigenvector is already normalized. For $\alpha_1 = 2$,

$$\begin{pmatrix}1 & 1 & 2\\0 & 2 & 0\\0 & 1 & 3\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix} = 2\begin{pmatrix}a\\b\\c\end{pmatrix} \quad \Rightarrow \quad \begin{array}{ccccccc} a & + & b & + & 2c & = & 2a\\ & & 2b & & & = & 2b\\ & & b & + & 3c & = & 2c\end{array}$$

The middle equation implies b is anything we want, and bottom equation indicates $b = -c$. Substituting $b = -c$ in the top equation yields $a = c$. We choose $a = 1$, which determines $c = 1$ and $b = -1$, so the eigenket is

$$|\,2\!> \,= A\begin{pmatrix}1\\-1\\1\end{pmatrix} \quad \text{with a normalization constant attached. Normalizing}$$

$$\big(1,\ -1,\ 1\big)A^*\,A\begin{pmatrix}1\\-1\\1\end{pmatrix} = |\,A\,|^2\big(1\!+\!1\!+\!1\big) = 3\,|\,A\,|^2 = 1 \quad\Rightarrow\quad A = \frac{1}{\sqrt{3}} \quad\Rightarrow\quad |\,2\!> \,= \frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\1\end{pmatrix}$$

for the eigenvalue $\alpha_2 = 2$. For the third eigenvalue

$$\begin{pmatrix}1 & 1 & 2\\0 & 2 & 0\\0 & 1 & 3\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix} = 3\begin{pmatrix}a\\b\\c\end{pmatrix} \quad \Rightarrow \quad \begin{array}{ccccccc} a & + & b & + & 2c & = & 3a\\ & & 2b & & & = & 3b\\ & & b & + & 3c & = & 3c\end{array}$$

The middle equation implies $b = 0$. Then the bottom equation implies $c =$ anything. We would like $a = 1$, which we can have if $c = 1$. The third eigenket is $|\,3\!> \,= A\begin{pmatrix}1\\0\\1\end{pmatrix}$. Normalizing,

$$\big(1,\ 0,\ 1\big)A^*\,A\begin{pmatrix}1\\0\\1\end{pmatrix} = |\,A\,|^2\big(1+0+1\big) = 2\,|\,A\,|^2 = 1 \quad\Rightarrow\quad A = \frac{1}{\sqrt{2}} \quad\Rightarrow\quad |\,3\!> \,= \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}.$$

---

17. What are the eigenvalues and eigenvectors for $\mathcal{L}_y = \begin{pmatrix}0 & -i & 0\\ i & 0 & -i\\ 0 & i & 0\end{pmatrix}$?

Use the procedures introduced in problem 16. The operator in this problem contains imaginary elements but is about the same degree of difficulty. There is another key difference that we will address in the postscript.

$$\det\left(\mathcal{L}_y - \alpha\mathcal{I}\right) = \det\begin{pmatrix} -\alpha & -i & 0 \\ i & -\alpha & -i \\ 0 & i & -\alpha \end{pmatrix} = (-\alpha)^3 - (-\alpha)(i)(-i) - (-i)(i)(-\alpha) = -\alpha^3 + \alpha + \alpha = 0$$

$$\Rightarrow \quad \alpha^3 - 2\alpha = 0 \quad \Rightarrow \quad \alpha(\alpha^2 - 2) = \alpha(\alpha - \sqrt{2}\,)(\alpha + \sqrt{2}\,) = 0 \quad \Rightarrow \quad \alpha = -\sqrt{2}\,,\ 0,\ \sqrt{2}$$

so $-\sqrt{2}\,,\ 0,$ and $\sqrt{2}$ are the eigenvalues. Calculating the eigenvectors starting with the smallest eigenvalue,

$$\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \left(-\sqrt{2}\right)\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{ccccc} & -bi & & = & -\sqrt{2}\,a \\ ai & - & ci & = & -\sqrt{2}\,b \\ & bi & & = & -\sqrt{2}\,c \end{array}$$

Adding the top and bottom equation, we get $-\sqrt{2}\,a - \sqrt{2}\,c = 0 \quad \Rightarrow \quad a = -c$. Using this in the middle equation, we get $ai + ai = -\sqrt{2}\,b \quad \Rightarrow \quad -\dfrac{2ai}{\sqrt{2}} = b \quad \Rightarrow \quad b = -\sqrt{2}\,ai$. If we choose $a = 1$, then $b = -\sqrt{2}\,i$, and $c = -1$, so the eigenket is

$$|-\sqrt{2}> = A\begin{pmatrix} 1 \\ -\sqrt{2}i \\ -1 \end{pmatrix} \quad \Rightarrow <-\sqrt{2}\,|-\sqrt{2}> = \left(1,\ \sqrt{2}\,i,\ -1\right) A^* A \begin{pmatrix} 1 \\ -\sqrt{2}\,i \\ -1 \end{pmatrix} = \left|\,A\,\right|^2 \left(1+2+1\right)$$

$$= 4\left|\,A\,\right|^2 = 1 \quad \Rightarrow \quad |-\sqrt{2}> = \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2}\,i \\ -1 \end{pmatrix} \quad \text{for the eigenvalue} \quad -\sqrt{2}\,.$$

To find the eigenvector corresponding to the next smallest eigenvalue 0,

$$\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{ccccc} & -bi & & = & 0 \\ ai & - & ci & = & 0 \\ & bi & & = & 0 \end{array}$$

The top and bottom equation say $b = 0$, and the middle equation says $a = c$, so by convention, $a = 1 \quad \Rightarrow \quad c = 1$, and the eigenket is

$$|\,0> = A\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \Rightarrow \quad <0\,|\,0> = \left(1,\ 0,\ 1\right) A^* A \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \left|\,A\,\right|^2 \left(1 + 0 + 1\right) = 2\left|\,A\,\right|^2 = 1$$

$$\Rightarrow \quad |\,0> = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{for the eigenvalue } 0\,.$$

For the eigenvector corresponding to the largest eigenvalue,

$$\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \sqrt{2}\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{ccccc} & -bi & & = & \sqrt{2}\,a \\ ai & - & ci & = & \sqrt{2}\,b \\ & bi & & = & \sqrt{2}\,c \end{array}$$

Adding the top and bottom equation, we get $\sqrt{2}\,a + \sqrt{2}\,c = 0 \quad \Rightarrow \quad a = -c\,.$ Using this in the middle equation, we get

$$ai + ai = \sqrt{2}\,b \quad \Rightarrow \quad \frac{2ai}{\sqrt{2}} = b \quad \Rightarrow \quad b = \sqrt{2}\,ai\,.$$

If we choose $a = 1\,,$ then $b = \sqrt{2}\,i,$ and $c = -1\,,$ the eigenket is

$$|\,\sqrt{2}> = A\begin{pmatrix} 1 \\ \sqrt{2}\,i \\ -1 \end{pmatrix} \quad \Rightarrow \quad <\sqrt{2}\,|\,\sqrt{2}> = \big(\,1,\ -\sqrt{2}\,i,\ -1\big)\,A^*A\begin{pmatrix} 1 \\ \sqrt{2}\,i \\ -1 \end{pmatrix} = \big|\,A\,\big|^2\big(1+2+1\big)$$

$$= 4\,\big|\,A\,\big|^2 = 1 \quad \Rightarrow \quad |\,\sqrt{2}> = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2}i \\ -1 \end{pmatrix} \quad \text{for the eigenvalue } \sqrt{2}\,.$$

Postscript: Notice that $\mathcal{L}_y$ in this problem is Hermitian, and that operator $\mathcal{A}$ of problem 16 is not Hermitian. Hermitian operators have two particularly desirable properties.

(1) The eigenvalues of Hermitian operators are real numbers, and

(2) the eigenvectors of Hermitian operators are orthogonal,

thus can be made orthonormal by the process of normalization. Property (1) is a necessity because any physical observable, for instance mass, charge, energy, momentum, or position, is quantified using a real number. Measurements in the real world require real numbers as outcomes. Property (2) is valuable because orthonormality is a necessity for a quantum mechanical basis. Use the procedures of problem 15 from part 1 to show that the eigenvectors of $\mathcal{L}_y$ are orthogonal, (or orthonormal if you include the normalization constants), if you want. Hermitian operators are central to quantum mechanical calculation because Hermitian operators represent all physically observable quantities.

---

18. What are the eigenvalues and eigenvectors for $\mathcal{D} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$?

The operator $\mathcal{D}$ is a diagonal operator. All non-zero elements are on the principal diagonal and all elements that are not on the principal diagonal are zero. Diagonal operators have some valuable properties. This problem intends to expose you to these properties. Solve the eigenvalue/eigenvector equation in the traditional way. You will find that the eigenvalues are the elements on the principal diagonal and the eigenvectors are the unit vectors that correspond to the position of the eigenvalue. And, since they are unit vectors, the eigenvectors are orthonormal unit vectors that do not require normalization.

$$\det\Big(\mathcal{D} - \alpha\mathcal{I}\Big) = \det\left[\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} - \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}\right]$$

$$= \begin{pmatrix} 2-\alpha & 0 & 0 \\ 0 & 3-\alpha & 0 \\ 0 & 0 & 4-\alpha \end{pmatrix} = (2-\alpha)(3-\alpha)(4-\alpha) + 0 + 0 - 0 - 0 - 0 = (2-\alpha)(3-\alpha)(4-\alpha)$$

$$\Rightarrow \quad (2-\alpha)(3-\alpha)(4-\alpha) = 0$$

is the characteristic equation, so the eigenvalues are $\alpha_i = 2,\ 3,$ and 4. The eigenvectors are

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 2 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{rcl} 2a &=& 2a \\ 3b &=& 2b \\ 4c &=& 2c \end{array}$$

The middle and bottom equations tell us $b = c = 0$. We choose $a = 1$, so the eigenket is

$$|\,2> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{for the eigenvalue} \quad \alpha_1 = 2\,.$$

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 3 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{rcl} 2a &=& 3a \\ 3b &=& 3b \\ 4c &=& 3c \end{array}$$

The top and bottom equations tell us $a = c = 0$. We choose $b = 1$, so the eigenket is

$$|\,3> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{for the eigenvalue} \quad \alpha_2 = 3\,.$$

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 4 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{rcl} 2a &=& 4a \\ 3b &=& 4b \\ 4c &=& 4c \end{array}$$

The top and middle equations tell us $a = b = 0$. We choose $c = 1$, so the eigenket is

$$|\,4> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{for the eigenvalue} \quad \alpha_3 = 4\,.$$

Postscript: The eigenvalues and eigenvectors of a diagonal operator are found by inspection. Also, orthonormality of the eigenvectors is inherent. For instance,

$$\begin{pmatrix} -3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix} \quad \Rightarrow \quad |-3> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\,1> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |\,5> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

without calculation. We would clearly prefer to work with diagonal operators if we had the opportunity. In fact, all Hermitian operators can be transformed into diagonal operators using a **unitary transformation**.

---

19. Diagonalize $\mathcal{L}_y = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$.

Unitary operators were introduced because they are central to the process of **diagonalization**. All Hermitian operators can be transformed into diagonal operators. Having the eigenvalues as the elements on the diagonal and the unit vectors as eigenvectors can be a significant convenience. The process of diagonalization can be accomplished using a **unitary transformation**. For an appropriate choice of $\mathcal{U}$, the transformation

$$\mathcal{U}^\dagger \mathcal{A} \mathcal{U} = \mathcal{A}'$$

yields an operator $\mathcal{A}'$ that is diagonal. The appropriate unitary operator is constructed from the eigenvectors of the operator to be diagonalized. There is, however, more than one way to arrange the eigenvectors to attain a unitary operator that will diagonalize the original operator. The convention that we will generally use is to keep the eigenvectors in columns, and place them in a matrix from left to right in the order of the vector corresponding to the lowest eigenvalue to the vector corresponding to the highest eigenvalue. For instance, given that the eigenvectors of $\mathcal{L}_y$ are

$$|-\sqrt{2}> = \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2}\,i \\ -1 \end{pmatrix}, \qquad |0> = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \text{and} \qquad |\sqrt{2}> = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2}\,i \\ -1 \end{pmatrix},$$

the unitary operator of interest is

$$\mathcal{U} = \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -i\sqrt{2}/2 & 0 & i\sqrt{2}/2 \\ -1/2 & 1/\sqrt{2} & -1/2 \end{pmatrix}.$$

$$\mathcal{U} = \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -i\sqrt{2}/2 & 0 & i\sqrt{2}/2 \\ -1/2 & 1/\sqrt{2} & -1/2 \end{pmatrix} \quad \Rightarrow \quad \mathcal{U}^\dagger = \begin{pmatrix} 1/2 & i\sqrt{2}/2 & -1/2 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -i\sqrt{2}/2 & -1/2 \end{pmatrix}$$

$$\Rightarrow \quad \mathcal{U}^\dagger \mathcal{A} \mathcal{U} = \begin{pmatrix} 1/2 & i\sqrt{2}/2 & -1/2 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -i\sqrt{2}/2 & -1/2 \end{pmatrix} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -i\sqrt{2}/2 & 0 & i\sqrt{2}/2 \\ -1/2 & 1/\sqrt{2} & -1/2 \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 & i\sqrt{2}/2 & -1/2 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -i\sqrt{2}/2 & -1/2 \end{pmatrix} \begin{pmatrix} -\sqrt{2}/2 & 0 & \sqrt{2}/2 \\ i & 0 & i \\ \sqrt{2}/2 & 0 & -\sqrt{2}/2 \end{pmatrix}$$

$$= \begin{pmatrix} -\sqrt{2}/4 - \sqrt{2}/2 - \sqrt{2}/4 & 0 & \sqrt{2}/4 - \sqrt{2}/2 + \sqrt{2}/4 \\ -1/2 + 1/2 & 0 & 1/2 - 1/2 \\ -\sqrt{2}/4 + \sqrt{2}/2 - \sqrt{2}/4 & 0 & \sqrt{2}/4 + \sqrt{2}/2 + \sqrt{2}/4 \end{pmatrix} = \begin{pmatrix} -\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

where the eigenvalues are on the principal diagonal in order of the smallest at the upper left to the largest at the lower right. The eigenvectors of this transformed operator are now the unit vectors, *i.e.*,

$$|-\sqrt{2}> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |0> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad |\sqrt{2}> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

A unitary transformation is a **change of basis**. It is another expression of the same operator in terms of a different set of linearly independent vectors. In a sense, our unitary transformation "rotates" the eigenvectors to align with the unit vectors. The transformed operator contains the same information as the original operator.

Of particular note is that *every Hermitian operator may be diagonalized by a unitary change of basis*, which is a result we will use, but will rely on other texts, such as Shankar[4], to support.

Our convention for forming a unitary operator is to keep the eigenvectors in columns, and place them in a matrix from left to right in the order of the eigenvectors that correspond to ascending eigenvalues. This convention aids organization and provides eigenvalues on the main diagonal in ascending order. There are other ways to form a unitary matrix that will diagonalize an operator. We will demonstrate an option in chapter 2.

---

20. Show that the eigenvalues of a matrix are invariant under a unitary transformation.

The postulates of quantum mechanics indicate that the eigenvalues of any operator are the only possible result of any measurement. Eigenvalues, therefore, are of preeminent importance to quantum mechanical measurement. This problem is intended to amplify and further legitimize the unitary transformation of problem 19. "Rotating" the eigenvectors to align with unit vectors does not affect the eigenvalues.

If you can show that $\Omega' = \mathcal{U}^\dagger \Omega \mathcal{U}$ and Ω have the same characteristic equation, then you are essentially done since the respective eigenvalues are the roots of identical characteristic equations. Start by letting $\Omega' = \mathcal{U}^\dagger \Omega \mathcal{U}$, a statement that says only that an operator subject to a unitary transformation is another operator. Work to justify $\Omega' - \omega \mathcal{I} = \mathcal{U}^\dagger (\Omega - \omega \mathcal{I}) \mathcal{U}$. We will insert an identity of the form $\mathcal{I} = \mathcal{U}^\dagger \mathcal{U}$ and use some of the operator algebra developed previously to attain this intermediate result. Remember that a determinant is a scalar, and that scalars commute with anything. Part (b) of problem 4 is pertinent. Identity operators also commute with anything. Conclude that $\det\big(\Omega' - \omega \mathcal{I}\big) = \det\big(\Omega - \omega \mathcal{I}\big)$ and reiterate that respective eigenvalues are the roots of identical characteristic equations.

Consider $\Omega' = \mathcal{U}^\dagger \Omega \mathcal{U}$ where $\mathcal{U}$ is unitary. The characteristic equation is formed by setting the determinant $\det(\Omega' - \omega I) = \det(\mathcal{U}^\dagger \Omega \mathcal{U} - \omega I)$ equal to zero. Consider first

$$\begin{aligned} \Omega' - \omega \mathcal{I} &= \mathcal{U}^\dagger \Omega \mathcal{U} - \omega \mathcal{I} = \mathcal{U}^\dagger \Omega \mathcal{U} - \omega \mathcal{I}\mathcal{I} \\ &= \mathcal{U}^\dagger \Omega \mathcal{U} - \omega \mathcal{I} \mathcal{U}^\dagger \mathcal{U} = \mathcal{U}^\dagger \Omega \mathcal{U} - \mathcal{U}^\dagger \omega \mathcal{I} \mathcal{U} = \mathcal{U}^\dagger (\Omega - \omega \mathcal{I}) \mathcal{U}. \end{aligned}$$

---

[4] Shankar, *Principles of Quantum Mechanics* (Plenum Press, New York, 1994), 2nd ed., p. 40.

Since these matrices are equal, their determinants are equal. So

$$\begin{aligned}\det\left(\Omega' - \mathcal{I}\,\omega\right) &= \det\left[\mathcal{U}^\dagger\left(\Omega - \mathcal{I}\,\omega\right)\mathcal{U}\right] = \det\left(\mathcal{U}^\dagger\right)\det\left(\Omega - \mathcal{I}\,\omega\right)\det\left(\mathcal{U}\right)\\ &= \det\left(\mathcal{U}^\dagger\right)\det\left(\mathcal{U}\right)\det\left(\Omega - \mathcal{I}\,\omega\right) = \det\left(\mathcal{U}^\dagger\mathcal{U}\right)\det\left(\Omega - \mathcal{I}\,\omega\right)\\ &= \det\left(\mathcal{I}\right)\det\left(\Omega - \mathcal{I}\,\omega\right) = \left(1\right)\det\left(\Omega - \mathcal{I}\,\omega\right) = \det\left(\Omega - \mathcal{I}\,\omega\right)\\ \text{therefore}\quad &\det\left(\mathcal{U}^\dagger\Omega\,\mathcal{U} - \mathcal{I}\,\omega\right) = \det\left(\Omega - \mathcal{I}\,\omega\right).\end{aligned}$$

Since the determinants of $\Omega - \mathcal{I}\,\omega$ and $\Omega' - \mathcal{I}\,\omega = \mathcal{U}^\dagger\Omega\,\mathcal{U} - \mathcal{I}\,\omega$ are equal, their characteristic equations are identical, and consequently, the eigenvalues of the operators Ω and $\Omega' = \mathcal{U}^\dagger\Omega\,\mathcal{U}$ are identical since the respective eigenvalues are the roots of identical characteristic equations. Therefore, the eigenvalues of an operator are not changed by a unitary change of basis.

---

21. Consider the two operators $\mathcal{A} = \begin{pmatrix} 3 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 3 \end{pmatrix}$ and $\mathcal{B} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & -3 & 0 \\ -1 & 0 & 0 \end{pmatrix}$.

(a) Are they Hermitian?

(b) Do they commute?

(c) Find the eigenvalues and eigenvectors of $\mathcal{A}$.

(d) Find $\mathcal{U}^\dagger\mathcal{A}\,\mathcal{U}$ where $\mathcal{U}$ is constructed from the eigenvectors of $\mathcal{A}$. Is it diagonal and what are its eigenvectors?

(e) Find $\mathcal{U}^\dagger\mathcal{B}\,\mathcal{U}$ using the same $\mathcal{U}$ constructed for part (d). Is it diagonal and what are its eigenvectors?

This is a toy problem meant to introduce a concept. This problem illustrates feasible mathematical mechanics to **simultaneously diagonalize** two matrix operators. Physics will likely never ask you to do such a calculation. Quantum mechanics, however, demands that you grasp the underlying concept. Simultaneous diagonalization of more than one operator is the mechanism that leads to a **complete set of commuting observables** required to describe many realistic physical systems.

All physical observables are represented by Hermitian matrices. Thus, part (a) asks you to check that $\mathcal{A} = \mathcal{A}^\dagger$, for instance. As stated in the last problem, every Hermitian operator can be diagonalized by a unitary change of basis, therefore, both of these Hermitian matrices can be diagonalized. When can they be simultaneously diagonalized? The answer is when they commute! *Two Hermitian operators that commute will both be diagonalized by the same unitary transformation.* This means that they share a common **eigenbasis**. Given the appropriate rotation, the eigenvalues do not change but both operators will have the same eigenvectors! The common eigenbasis for Hermitian matrices will be unit vectors. Part (b) asks you to check that the commutator, $[\,\mathcal{A},\ \mathcal{B}\,] = \mathcal{A}\mathcal{B} - \mathcal{B}\,\mathcal{A} = 0$. You will find that these two operators commute.

You need to solve the eigenvalue/eigenvector problem for one of the two operators. The eigenvectors of either operator are satisfactory to form an appropriate unitary matrix. We arbitrarily choose to solve the eigenvalue/eigenvector problem for $\mathcal{A}$ in part (c) of this problem. The eigenvalues of $\mathcal{A}$ are 2, 3, and 4. Use the procedures of problem 20 to form your unitary matrix. You will find that both $\mathcal{U}^\dagger\mathcal{A}\,\mathcal{U}$ and $\mathcal{U}^\dagger\mathcal{B}\,\mathcal{U}$ are diagonal operators in parts (d) and (e).

(a) Both operators are Hermitian, because $\mathcal{A} = \mathcal{A}^\dagger$ and $\mathcal{B} = \mathcal{B}^\dagger$.

(b) The commutator is

$$\big[\mathcal{A},\, \mathcal{B}\big] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = \begin{pmatrix} 3 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 3 \end{pmatrix}\begin{pmatrix} 0 & 0 & -1 \\ 0 & -3 & 0 \\ -1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & -1 \\ 0 & -3 & 0 \\ -1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 3 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 0+0+1 & 0+0+0 & -3+0+0 \\ 0+0+0 & 0-9+0 & 0+0+0 \\ 0+0-3 & 0+0+0 & 1+0+0 \end{pmatrix} - \begin{pmatrix} 0+0+1 & 0+0+0 & 0+0-3 \\ 0+0+0 & 0-9+0 & 0+0+0 \\ -3+0+0 & 0+0+0 & 1+0+0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & -3 \\ 0 & -9 & 0 \\ -3 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & -3 \\ 0 & -9 & 0 \\ -3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0 \cdot \mathcal{I}\,, \quad \text{or just } 0\,,$$

therefore $\mathcal{A}$ and $\mathcal{B}$ commute. This means that there is a basis of common eigenvectors that simultaneously diagonalize both $\mathcal{A}$ and $\mathcal{B}$.

(c) The unitary transformation to this basis can be constructed from the eigenvectors of either operator, so we arbitrarily pick $\mathcal{A}$ and find its eigenvalues and eigenvectors.

$$\det\big(\mathcal{A} - \mathcal{I}\,\alpha\big) = \begin{pmatrix} 3-\alpha & 0 & -1 \\ 0 & 3-\alpha & 0 \\ -1 & 0 & 3-\alpha \end{pmatrix} = (3-\alpha)^3 + 0 + 0 - 0 - 0 - (3-\alpha) = 0$$

$$\Rightarrow \quad 27 - 27\alpha + 9\alpha^2 - \alpha^3 - 3 + \alpha = 0 \quad \Rightarrow \quad \alpha^3 - 9\alpha^2 + 26\alpha - 24 = 0$$

$$\Rightarrow \quad (\alpha-2)(\alpha-3)(\alpha-4) = 0 \quad \Rightarrow \quad \alpha = 2,\ 3, \text{ and } 4 \text{ are the eigenvalues of } \mathcal{A}\,.$$

Next, find the eigenvectors.

$$\alpha = 2 \quad \Rightarrow \quad \begin{pmatrix} 3 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 3 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 2\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{matrix} 3a - c & = & 2a \\ 3b & = & 2b \\ -a + 3c & = & 2c \end{matrix} \quad \Rightarrow \quad \begin{matrix} a & = & c \\ b & = & 0 \\ a & = & c \end{matrix}$$

Choose $a = 1 \;\Rightarrow\; c = 1$, and $b = 0$, so $\;|\,2\!> = A\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, and $<\!2\,|\,2\!> = 1$ means that

$$\big(1,\ 0,\ 1\big)A^*A\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = |\,A\,|^2(1+0+1) = 2\,|\,A\,|^2 = 1 \quad \Rightarrow \quad A = \frac{1}{\sqrt{2}} \quad \Rightarrow \quad |\,2\!> = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

$$\alpha = 3 \quad \Rightarrow \quad \begin{pmatrix} 3 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 3 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 3\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{matrix} 3a - c & = & 3a \\ 3b & = & 3b \\ -a + 3c & = & 3c \end{matrix} \quad \Rightarrow \quad \begin{matrix} c & = & 0 \\ b & = & 1 \\ a & = & 0 \end{matrix}$$

$$|\,3\!> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{which is a unit vector so is already normalized.}$$

$$\alpha = 4 \quad \Rightarrow \quad \begin{pmatrix} 3 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 4 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{rcl} 3a - c & = & 4a \\ 3b & = & 4b \\ -a + 3c & = & 4c \end{array} \quad \Rightarrow \quad \begin{array}{rcl} -c & = & a \\ b & = & 0 \\ -a & = & c \end{array}$$

Choose $a = 1 \Rightarrow c = -1$, and $b = 0$, so $|\,4\!> = A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, and $<\!4\,|\,4\!> = 1$ means that

$$(1,\ 0, -1) A^* A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \big|\, A \,\big|^2 (1+0+1) = 2 \big|\, A \,\big|^2 = 1 \ \Rightarrow \ A = \frac{1}{\sqrt{2}} \ \Rightarrow \ |\,4\!> = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

(d) Form $\mathcal{U}$ from the eigenvectors of $\mathcal{A}$ by placing them in a matrix from left to right in order of ascending eigenvalue.

$$\mathcal{U} = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \quad \Rightarrow \quad \mathcal{U}^\dagger = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix}$$

$$\begin{aligned} \Rightarrow \quad \mathcal{U}^\dagger \mathcal{A} \mathcal{U} &= \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3/\sqrt{2} - 1/\sqrt{2} & 0 & 3/\sqrt{2} + 1/\sqrt{2} \\ 0 & 3 & 0 \\ -1/\sqrt{2} + 3/\sqrt{2} & 0 & -1/\sqrt{2} - 3/\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 2/\sqrt{2} & 0 & 4/\sqrt{2} \\ 0 & 3 & 0 \\ 2/\sqrt{2} & 0 & -4/\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 2/2 + 2/2 & 0 & 4/2 - 4/2 \\ 0 & 3 & 0 \\ 2/2 - 2/2 & 0 & 4/2 + 4/2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}. \end{aligned}$$

This operator is diagonal. The eigenvalues on the principal diagonal are as calculated in part (c), but the eigenvectors are now

$$|\,2\!> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\,3\!> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |\,4\!> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

(e) Applying the same transformation to $\mathcal{B}$,

$$\begin{aligned} \mathcal{U}^\dagger \mathcal{B} \mathcal{U} &= \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & -3 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & -3 & 0 \\ -1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \end{aligned}$$

$$= \begin{pmatrix} -1/2 - 1/2 & 0 & 1/2 - 1/2 \\ 0 & -3 & 0 \\ -1/2 + 1/2 & 0 & 1/2 + 1/2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This is a diagonal matrix. The eigenvalues and corresponding eigenvectors are

$$|-1> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |-3> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |1> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Postscript: Notice that after the unitary transformations are completed the common eigenvectors of both operators are the unit vectors.

Suppose that we did part (e) before part (d). Given that we solved the eigenvalue/eigenvector problem for $\mathcal{A}$ in part (c) so have the eigenvalues, the diagonal form of $\mathcal{A}$ can be constructed by inspection. Do you understand why?

The diagonal operator for $\mathcal{B}$ has eigenvalues on the principal diagonal but they are not in ascending order. That is because the unitary matrix was formed from the eigenvectors of $\mathcal{A}$. If the unitary matrix was formed from the eigenvectors of $\mathcal{B}$, the eigenvalues of $\mathcal{B}$ would be in ascending order and the eigenvalues of the transformed $\mathcal{A}$ would appear in the order 3, 2, 4.

The key thing to take from this problem is that any two Hermitian operators that commute have a common set of eigenvectors so can be simultaneously diagonalized.

---

22. Given that

$$\mathcal{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{and that the state function is the superposition} \quad |\psi> = \beta \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

what are the possible results of a measurement of $\mathcal{D}$?

This problem introduces substance that is the reason for the next two problems. It informally introduces two of the postulates of quantum mechanics. The primary reason for this lengthy mathematics preliminary is to make the postulates of quantum mechanics and the calculations that follow from them accessible.

The eigenvalues and corresponding eigenvectors of $\mathcal{D}$ are

$$|1> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |2> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |3> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The first postulate of quantum mechanics indicates that any possible state vector is a linear combination of the eigenvectors, *i.e.*,

$$|\psi> = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

where α, β, and γ are scalars. This type of linear combination is known as a **superposition**. A state vector is a linear combination or superposition of the eigenvectors. The third postulate of quantum mechanics indicates that the only possible results of a measurement are the eigenvalues of the operator that represents the observable. In general, the possible results of a measurement of $\mathcal{D}$ are 1, 2, and 3. The third postulate also indicates, however, that the state vector must contain the eigenvector in order to measure the corresponding eigenvalue. The given state vector contains only two of the three eigenvectors, so only two of the three eigenvalues are possible results of a measurement.

Possible outcomes of a measurement are 2, the eigenvalue corresponding to the second eigenvector, and 3, the eigenvalue corresponding to the third eigenvector. A measurement with a result of 1 is impossible, because the state vector does not contain the eigenvector that corresponds to the eigenvalue 1.

Postscript: The condition that the state vector does not contain the first eigenvector is equivalent to $\alpha = 0$. We will examine the meaning of the coefficients α, β, and γ in chapter 2.

We have examined only operators that have the property that there is one eigenvector for each eigenvalue to this point. Consider the operator $\mathcal{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. This has eigenvalues of 1, 1, and 2. An eigenvalue is repeated. The eigenvectors are

$$|\,1> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\,1> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |\,2> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

$\mathcal{A}$ is Hermitian so could represent a measurable quantity. Only 1 or 2 are possible outcomes of a measurement of $\mathcal{A}$ because those are its only eigenvalues. But two eigenvectors are associated with the eigenvalue 1. If the outcome of a measurement is 1, to which eigenvector does this correspond? It is impossible to tell. This is an example of **degeneracy**. An operator that has an eigenvalue corresponding to two or more eigenvectors is known as degenerate.

---

23. Is $\mathcal{B} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ degenerate?

Solve the characteristic equation. If $\mathcal{B}$ has any repeated eigenvalues, then that eigenvalue will correspond to more than one eigenvector and $\mathcal{B}$ is degenerate.

$$\det\left[\mathcal{B} - \alpha\,\mathcal{I}\right] = \det\begin{pmatrix} 1-\alpha & 0 & 1 \\ 0 & 2-\alpha & 0 \\ 1 & 0 & 1-\alpha \end{pmatrix} = (1-\alpha)^2(2-\alpha) - (2-\alpha)$$

$$= (1-2\alpha+\alpha^2)(2-\alpha) - (2-\alpha) = (1-2\alpha+\alpha^2-1)(2-\alpha) = (-2\alpha+\alpha^2)(2-\alpha) \;\;\Rightarrow\;\; -\alpha\,(2-\alpha)^2 = 0$$

(because $(-2\alpha+\alpha^2)=-\alpha\,(2-\alpha)\,)$, is the characteristic equation. The eigenvalues are 0, 2, and 2. The eigenvalue 2 corresponds to more than one eigenstate, so the operator $\mathcal{B}$ is degenerate.

Postscript: Notice that we have referred to an eigenvector as an **eigenstate**. The two terms are used interchangeably. Since the superposition of eigenvectors is known as the "state vector," the term eigenstate frequently preferred within the realm of physics.

---

24. Simultaneously diagonalize $\quad \mathcal{A}=\begin{pmatrix}1&0&1\\0&2&0\\1&0&1\end{pmatrix}$ and $\mathcal{B}=\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}$.

The procedures to this problem parallel problem 22 except that operator $\mathcal{A}$ is degenerate. Both operators are Hermitian. The next question is do they commute? Check to see if these operators commute and they will... because this is a non-problem if they do not commute. Operators that do not commute do not have a common eigenbasis so cannot be diagonalized simultaneously. Solve the eigenvalue/eigenvector problem for operator $\mathcal{B}$. The choice was arbitrary in problem 22. It is not arbitrary in this problem because $\mathcal{A}$ is degenerate. Solve the eigenvalue/eigenvector problem for $\mathcal{B}$ because it is the non-degenerate operator. Procedures for forming the unitary matrix and doing the unitary transformations are the same as problem 22.

$$\left[\mathcal{A},\mathcal{B}\right]=\begin{pmatrix}1&0&1\\0&2&0\\1&0&1\end{pmatrix}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}-\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}1&0&1\\0&2&0\\1&0&1\end{pmatrix}$$
$$=\begin{pmatrix}0&2&0\\2&0&2\\0&2&0\end{pmatrix}-\begin{pmatrix}0&2&0\\2&0&2\\0&2&0\end{pmatrix}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}.$$

These operators commute so they can be simultaneously diagonalized. $\mathcal{A}$ is degenerate from problem 23, so solve the eigenvalue/eigenvector problem for $\mathcal{B}$ and hope that it is not degenerate.

$$\det\begin{pmatrix}-\alpha&1&0\\1&-\alpha&1\\0&1&-\alpha\end{pmatrix}=(-\alpha)^3-(-\alpha)-(-\alpha)=-\alpha^3+2\alpha\quad\Rightarrow\quad-\alpha^3+2\alpha=0$$

$$\Rightarrow\quad\alpha^3-2\alpha=0\quad\Rightarrow\quad\alpha\,(\alpha^2-2)=0\quad\Rightarrow\quad\alpha\,(\alpha+\sqrt{2}\,)(\alpha-\sqrt{2}\,)=0$$

so the eigenvalues are $\alpha=-\sqrt{2},\ 0,\ \sqrt{2}$. $\mathcal{B}$ is not degenerate. Therefore, $\mathcal{A}$ and $\mathcal{B}$ can be diagonalized simultaneously by forming the unitary matrix from the eigenvectors of $\mathcal{B}$. We need to find the eigenvectors.

$$\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix}=\left(-\sqrt{2}\right)\begin{pmatrix}a\\b\\c\end{pmatrix}\quad\Rightarrow\quad\begin{aligned}b&=-\sqrt{2}\,a\\a+c&=-\sqrt{2}\,b\\b&=-\sqrt{2}\,c\end{aligned}$$

The top and bottom equations tell us $a=c$. Using this in the middle equation,

$$a+a=-\sqrt{2}\,b\quad\Rightarrow\quad b=-\frac{2}{\sqrt{2}}\,a\quad\Rightarrow\quad b=-\sqrt{2}\,a\,,$$

so choosing $a = 1$ determines b and c, or $|-\sqrt{2}> = A \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$. Normalizing,

$$\left(1,\ -\sqrt{2},\ 1\right) A^* A \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} = 4\,|\,A\,|^2 = 1 \quad \Rightarrow \quad A = \frac{1}{2} \ \Rightarrow \ |-\sqrt{2}> = \frac{1}{2} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}.$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{matrix} b & = & 0 \\ a + c & = & 0 \\ b & = & 0 \end{matrix}$$

$\Rightarrow \ b = 0$ and $a = -c$, so choosing $a = 1$, $|\,0> = A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Normalizing,

$$\left(1,\ 0,\ -1\right) A^* A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = |\,A\,|^2 \left(1+0+1\right) = 2\,|\,A\,|^2 = 1 \ \Rightarrow \ A = \frac{1}{\sqrt{2}} \ \Rightarrow \ |\,0> = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \sqrt{2} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{matrix} b & = & \sqrt{2}\,a \\ a + c & = & \sqrt{2}\,b\,. \\ b & = & \sqrt{2}\,c \end{matrix}$$

The top and bottom equations tell us $a = c$. Using this in the middle equation,

$$a + a = \sqrt{2}\,b \quad \Rightarrow \quad b = \frac{2}{\sqrt{2}}\,a \quad \Rightarrow \quad b = \sqrt{2}\,a\,,$$

so choosing $a = 1$ determines b and c, or $|\sqrt{2}> = A \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$. Normalizing,

$$\left(1,\ \sqrt{2},\ 1\right) A^* A \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} = |\,A\,|^2 \left(1+2+1\right) = 4\,|\,A\,|^2 = 1 \ \Rightarrow \ A = \frac{1}{2} \ \Rightarrow \ |\sqrt{2}> = \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}.$$

Now that we have the eigenvectors, we can form the appropriate unitary operator,

$$\mathcal{U} = \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} & 1/2 \end{pmatrix} \quad \Rightarrow \quad \mathcal{U}^\dagger = \begin{pmatrix} 1/2 & -1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & 1/\sqrt{2} & 1/2 \end{pmatrix}$$

$$\Rightarrow \ \mathcal{U}^\dagger \mathcal{A} \mathcal{U} = \begin{pmatrix} 1/2 & -1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & 1/\sqrt{2} & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} & 1/2 \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 & -1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & 1/\sqrt{2} & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 + 1/2 & 1/\sqrt{2} - 1/\sqrt{2} & 1/2 + 1/2 \\ -2/\sqrt{2} & 0 & 2/\sqrt{2} \\ 1/2 + 1/2 & 1/\sqrt{2} - 1/\sqrt{2} & 1/2 + 1/2 \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 & -1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & 1/\sqrt{2} & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ -2/\sqrt{2} & 0 & 2/\sqrt{2} \\ 1 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 + 2/2 + 1/2 & 0 & 1/2 - 2/2 + 1/2 \\ 1/\sqrt{2} - 1/\sqrt{2} & 0 & 1/\sqrt{2} - 1/\sqrt{2} \\ 1/2 - 2/2 + 1/2 & 0 & 1/2 + 2/2 + 1/2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

We do not need to do the diagonalization for $\mathcal{B}$ because of our method of constructing the unitary matrix. Place the eigenvalues on the diagonal in ascending order, and the eigenvectors are the unit vectors by inspection, meaning

$$\mathcal{U}^\dagger \mathcal{B} \mathcal{U} = \begin{pmatrix} -\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}.$$

Postscript: The eigenvectors for both $\mathcal{A}$ and $\mathcal{B}$ are the unit vectors where the eigenvalues correspond to the unit eigenvector appropriate to that position. Look at what has developed. If we measure that $\mathcal{A}$ has the eigenvalue 2 and $\mathcal{B}$ has the eigenvalue $-\sqrt{2}$, we can ascertain the eigenvector is $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. If we measure that $\mathcal{A}$ has the eigenvalue 2 and $\mathcal{B}$ has the eigenvalue $\sqrt{2}$, the eigenvector is $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. We can now tell which eigenvector corresponds to the eigenvalue of 2 for $\mathcal{A}$ by measuring both $\mathcal{A}$ and $\mathcal{B}$. We have removed the degeneracy, however, we must specify two eigenvalues to identify one eigenstate. That is one of the consequence of degeneracy.

The concept is much more important than the mathematical mechanics. This is a toy problem intended for illustration and exercise. Physical systems are commonly degenerate. The process used to remove the degeneracy in problem 24 is the same process used to remove degeneracy in realistic physical systems. We have developed what is known as a **complete set of commuting observables** for the operators of the last problem.

A complete set of commuting observables is used to remove a degeneracy. The first step in developing a complete set of commuting observables is to find a non-degenerate operator with which the degenerate operator commutes. The eigenvectors of the non-degenerate operator can be used to form a unitary matrix for unitary transformations that will simultaneously diagonalize both operators.

In the instance that $\mathcal{A}$ and $\mathcal{B}$ are both degenerate, it is necessary to find an operator $\mathcal{C}$ with which both $\mathcal{A}$ and $\mathcal{B}$ commute to develop a complete set of commuting observables. In this circumstance, three eigenvalues would be required to specify one eigenstate.

It is not a mathematical necessity to use the eigenvectors of the non-degenerate operator to form the unitary matrix, though it is a practical necessity. The indeterminacy in the eigenvector/eigenvalue problem of the degenerate operator make it difficult to attain a unitary matrix that leads to transformations that diagonalize operators other than itself.

---

25. Show that for any Hermitian operator Ω,

$$\det \Omega = \text{the product of the eigenvalues of } \Omega = \prod_{i=1}^{n} \omega_i \,.$$

Remember that any Hermitian operator can be diagonalized. That the determinant of a matrix is the same as the product of the eigenvalues is a fact that is fairly accessible if the matrix is diagonal.

You need to show that the determinant of an operator is invariant under a unitary transformation. You already know that the eigenvalues are not changed by a unitary change of basis. A unitary transformation diagonalizes Ω. So if you can show that the statement is true in the diagonal basis, then you are essentially done. The fact that the determinant of a matrix is the product of the eigenvalues is a simple consistency check for eigenvalue/eigenvector problems.

If Ω is Hermitian, there exists a unitary operator $\mathcal{U}$ such that $\mathcal{U}^\dagger \Omega \mathcal{U}$ is diagonal. That the eigenvalues of $\mathcal{U}^\dagger \Omega \mathcal{U}$ and Ω are the same has been demonstrated previously. Further, the determinants of $\mathcal{U}^\dagger \Omega \mathcal{U}$ and Ω are identical because

$$\det \mathcal{U}^\dagger \Omega \mathcal{U} = \det \mathcal{U}^\dagger \det \Omega \det \mathcal{U} = \det \mathcal{U}^\dagger \det \mathcal{U} \det \Omega = \det \left(\mathcal{U}^\dagger \mathcal{U} \right) \det \Omega = \left(1 \right) \det \Omega = \det \Omega \,.$$

Let $\Lambda = \mathcal{U}^\dagger \Omega \mathcal{U}$. Λ is diagonal because $\mathcal{U}^\dagger \Omega \mathcal{U}$ is diagonal. To find the eigenvalues,

$$\det \left(\mathcal{U}^\dagger \Omega \mathcal{U} - \mathcal{I} \omega \right) = \det \left(\Lambda - \omega I \right) = 0$$

$$\Rightarrow \quad \det \begin{pmatrix} \Lambda_{11} - \omega & 0 & 0 & \cdots & 0 \\ 0 & \Lambda_{22} - \omega & 0 & \cdots & 0 \\ 0 & 0 & \Lambda_{33} - \omega & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Lambda_{nn} - \omega \end{pmatrix} = 0 \,.$$

This has the characteristic equation $\left(\Lambda_{11} - \omega \right) \left(\Lambda_{22} - \omega \right) \left(\Lambda_{33} - \omega \right) \cdots \left(\Lambda_{nn} - \omega \right) = 0$, which has n roots $\omega_i = \Lambda_{ii}$, $i = 1, 2, \ldots, n$. The elements on the principal diagonal of a diagonal matrix are the eigenvalues. Therefore, the form of $\det \Lambda$ or $\det U^\dagger \, \Omega \, U$ in the diagonal basis is

$$\det \Omega = \det \left(\mathcal{U}^\dagger \Omega \mathcal{U} \right) = \det \begin{pmatrix} \omega_1 & 0 & 0 & \cdots & 0 \\ 0 & \omega_2 & 0 & \cdots & 0 \\ 0 & 0 & \omega_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \omega_n \end{pmatrix} = (\omega_1)(\omega_2)(\omega_3) \cdots (\omega_n) = \prod_{i=1}^{n} \omega_i \,. \quad \text{Q.E.D.}$$

Supplementary Problems

26. Show that

$$\mathcal{A} = \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{is linear.}$$

Like problem 1, show that $\mathcal{A}$ commutes with scalars and that it distributes over vectors.

27. Use explicit calculations to show that

(a) $\det\big(\mathcal{A}\,\mathcal{A}\big) = \det\mathcal{A}\det\mathcal{A}$, and

(b) $\det\big(\mathcal{A}\,\mathcal{B}\big) = \det\mathcal{A}\det\mathcal{B}$,

$$\text{for} \quad \mathcal{A} = \begin{pmatrix} 2 & i & 1 \\ -i & 1 & -i \\ 1 & i & 2 \end{pmatrix} \qquad \mathcal{B} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 3 \end{pmatrix}.$$

This is a numerical exercise intended to reinforce the fact that the determinant of the products is the same as the product of the determinants. Simply form the appropriate determinants and products for the right and left sides of the equations, and they will be the same.

28. Verify that the operator $\mathcal{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ is unitary.

This problem is practice in forming adjoints and should help solidify the meaning of a unitary matrix. $\mathcal{R}$ represents a rotation of $\pi/2$ around the x-axis in three dimensions. Calculate $\mathcal{R}^\dagger$ and $\mathcal{R}\mathcal{R}^\dagger$. You will find that $\mathcal{R}\mathcal{R}^\dagger = \mathcal{I}$, so can conclude that this matrix is unitary.

29. Show that $\big(\mathcal{A} + \mathcal{B}\big)^\dagger = \mathcal{A}^\dagger + \mathcal{B}^\dagger$ in two dimensions.

That the adjoint of the sum is the same as the sum of the adjoints is not obvious but the fact is used frequently. The proof in two dimensions is simplest. The result extends to arbitrary or infinite dimension. Use

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad \text{and} \qquad \mathcal{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Calculate the adjoint of the sum and the sum of the adjoints. They will be the same.

30. (a) Show that any two diagonal matrices commute.

(b) Show that a diagonal matrix does not commute with an arbitrary matrix.

The fact that diagonal operators commute is both a useful tool and another reason to diagonalize operators when the opportunity exists. Part (b) warns you that there is an end to this good fortune because you cannot generally assume commutivity. For part (a), explicitly multiply two diagonal matrices of arbitrary dimension. Your result should be another diagonal matrix. The elements of the resulting diagonal matrix will be the product of two numbers. Is each of these products the same if you reverse the order of the factors? Why? For part (b), explicitly multiply a matrix with all non-zero elements by a diagonal matrix. You should show that the off-diagonal elements are different when you change the order of matrix multiplication.

---

31. Show that $\mathcal{U}\mathcal{U}^\dagger = \mathcal{I} \quad \Rightarrow \quad \mathcal{U}^\dagger\mathcal{U} = \mathcal{I}\,.$

This problem is an application of the insertion of the identity. Remember that both $\mathcal{U}\mathcal{U}^\dagger = \mathcal{I}$ and $\mathcal{U}^{-1}\mathcal{U} = \mathcal{I}$. Assume that $\mathcal{U}$ is non-singular.

---

32. Given that Ω and Λ are Hermitian, show that

(a) $\Omega\,\Lambda$ is not Hermitian,

(b) $\Omega\Lambda + \Lambda\Omega$ is Hermitian,

(c) $\big[\,\Omega,\,\Lambda\,\big]$ is anti-Hermitian, and

(d) $i\big[\,\Omega,\,\Lambda\,\big]$ is Hermitian.

This problem is designed to deepen your understanding of the properties of Hermitian operators. It should also provide a vehicle for use of commutators and some previous results. All parts use the result of problem 11. As an example of how you want to proceed, here is the solution to part (d). Hermitian for part (d) means $\big(\,i\,\big[\,\Omega,\,\Lambda\,\big]\,\big)^\dagger = i\,\big[\,\Omega,\,\Lambda\,\big]$. To show that this is true,

$$\begin{aligned}
\big(\,i\,\big[\,\Omega,\,\Lambda\,\big]\,\big)^\dagger &= (i)^*\,\big(\Omega\,\Lambda - \Lambda\,\Omega\big)^\dagger \quad = -i\,\Big(\big(\Omega\,\Lambda\big)^\dagger - \big(\Lambda\,\Omega\big)^\dagger\Big) \\
&= -i\,\big(\Lambda^\dagger\,\Omega^\dagger - \Omega^\dagger\,\Lambda^\dagger\big) \;= -i\,\big(\Lambda\,\Omega - \Omega\,\Lambda\big) \\
&= -i\,\big(-\Omega\,\Lambda + \Lambda\,\Omega\big) \quad = \;\; i\,\big(\Omega\,\Lambda - \Lambda\,\Omega\big) \\
&= \;\; i\,\big[\,\Omega,\,\Lambda\,\big], \quad \text{so it is Hermitian.}
\end{aligned}$$

---

33. Consider the matrix $\quad \Omega = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 2 & 0 \end{pmatrix}.$

(a) Is it Hermitian?

(b) Find its eigenvalues and normalized eigenvectors.

(c) Verify that $\mathcal{U}^\dagger \Omega \mathcal{U}$ is diagonal, where $\mathcal{U}$ is the matrix of eigenvectors of Ω.

The eigenvalue/eigenvector problem is one of the preeminent problems of quantum mechanics, so you need to understand it thoroughly. Part (a) is simply to recognize Hermiticity. Show that $\Omega = \Omega^\dagger$. Solve the eigenvalue/eigenvector problem. You should find

$$|-2> = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \qquad |0> = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad |4> = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Use the convention of keeping the column vectors in columns and placing them in a matrix from left to right in the ascending order of the eigenvalues to form the unitary matrix. The unitary matrix that diagonalizes Ω is

$$\mathcal{U} = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \end{pmatrix}$$

using our convention. $\mathcal{U}^\dagger \Omega \mathcal{U}$ is diagonal, with eigenvalues from left to right in ascending order.

---

34. Consider the matrix

$$\Omega = \begin{pmatrix} 0 & 4i & 0 \\ -4i & 0 & -3i \\ 0 & 3i & 0 \end{pmatrix}.$$

(a) Is it Hermitian?

(b) Find its eigenvalues and normalized eigenvectors.

(c) Verify that $\mathcal{U}^\dagger \Omega \mathcal{U}$ is diagonal, where $\mathcal{U}$ is the unitary matrix constructed from the eigenvectors of Ω.

The intent and procedures for this problem are essentially identical to the previous one, however, this operator has elements that are imaginary. The operator Ω is Hermitian. The eigenvalues are integral but the eigenvectors are slightly more challenging than the last problem. Using the convention that the leading non-zero component should be 1, you should find

$$|-5> = \frac{4}{5\sqrt{2}}\begin{pmatrix} 1 \\ i5/4 \\ 3/4 \end{pmatrix}, \quad |0> = \frac{3}{5}\begin{pmatrix} 1 \\ 0 \\ -4/3 \end{pmatrix}, \quad \text{and} \quad |5> = \frac{4}{5\sqrt{2}}\begin{pmatrix} 1 \\ -i5/4 \\ 3/4 \end{pmatrix}.$$

The eigenvectors above lead to the unitary matrix

$$\mathcal{U} = \begin{pmatrix} 4/5\sqrt{2} & 3/5 & 4/5\sqrt{2} \\ i/\sqrt{2} & 0 & -i/\sqrt{2} \\ 3/5\sqrt{2} & -4/5 & 3/5\sqrt{2} \end{pmatrix}$$

if formed using our convention. The transformation $\mathcal{U}^\dagger \Omega \mathcal{U}$ does diagonalize this matrix.

---

35. Consider the matrix

$$\Omega = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 4 & -i \\ 0 & i & 0 \end{pmatrix}.$$

(a) Is it Hermitian?

(b) Find its eigenvalues and normalized eigenvectors.

(c) Verify that $\mathcal{U}^\dagger \Omega \mathcal{U}$ is diagonal, where $\mathcal{U}$ is the unitary matrix constructed from the eigenvectors of Ω.

This problem has both real and imaginary elements. Again, this operator is Hermitian. You should find

$$|-1> = \sqrt{\frac{2}{3}} \begin{pmatrix} 1 \\ -1/2 \\ i/2 \end{pmatrix}, \quad |0> = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 0 \\ -2i \end{pmatrix}, \quad \text{and} \quad |5> = \frac{2}{\sqrt{30}} \begin{pmatrix} 1 \\ 5/2 \\ i/2 \end{pmatrix},$$

for eigenvalues and eigenvectors. These lead to the unitary matrix

$$\mathcal{U} = \begin{pmatrix} 2/\sqrt{6} & 1/\sqrt{5} & 2/\sqrt{30} \\ -1/\sqrt{6} & 0 & 5/\sqrt{30} \\ i/\sqrt{6} & -2i/\sqrt{5} & i/\sqrt{30} \end{pmatrix}$$

if formed using our convention. We have done some arithmetic so that denominators are common in each column. If you have the foresight to do this, you save having to find common denominators in the following steps. Again, the transformation $\mathcal{U}^\dagger \Omega \mathcal{U}$ must diagonalize this matrix.

---

36. Consider the operator $\Omega = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$.

(a) Show that Ω is unitary.

(b) Show that Ω's eigenvalues are $e^{i\theta}$ and $e^{-i\theta}$.

(c) Find the corresponding eigenvectors, and show that they are orthogonal.

(d) Verify that $\mathcal{U}^\dagger \Omega \mathcal{U}$ is diagonal, where $\mathcal{U}$ is the unitary matrix constructed from the eigenvectors of Ω.

You should proceed forward undaunted even though this operator has components that are trigonometric functions. You will probably want to use the quadratic formula to solve the eigenvalue problem. You may also want to recall Euler's equation $e^{\pm i\theta} = \cos\theta \pm i\sin\theta$.

---

37. Show that $\mathcal{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ $\mathcal{B} = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$ cannot be simultaneously diagonalized.

If they do not commute, then they cannot be simultaneously diagonalized because there is no common eigenbasis!

---

38. Consider the operators $\mathcal{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \mathcal{B} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 2 \end{pmatrix}$.

(a) Show that they are Hermitian.

(b) Show that they commute.

(c) Diagonalize $\mathcal{A}$ and $\mathcal{B}$ using a unitary matrix formed from the eigenvectors of $\mathcal{B}$.

The operators are Hermitian. Part (b) means to check if $\left[\mathcal{A}, \mathcal{B}\right] = 0$, and it does. Part (c) means solve the eigenvalue/eigenvector problem for $\mathcal{B}$ and form $\mathcal{U}$ from the eigenvectors of that operator so as to diagonalize both operators using the transformations $\mathcal{U}^\dagger \mathcal{A} \mathcal{U}$ and $\mathcal{U}^\dagger \mathcal{B} \mathcal{U}$. You want to form the unitary operator from the eigenvectors of $\mathcal{B}$ because $\mathcal{A}$ is degenerate.

The most important idea here is that when one operator is degenerate, simultaneous diagonalization is essentially a necessary procedure. It is the procedure used in problems where degeneracy occurs to attain a complete set of commuting operators. In fact, where two or more **quantum numbers** are required to specify an eigenstate, you can be sure that simultaneous diagonalization or an equivalent process has been accomplished.

If a degenerate operator is encountered, you can lift the degeneracy using simultaneous diagonalization by finding another non-degenerate operator with which it commutes, and using the eigenvectors of the non-degenerate operator to form a unitary transformation that diagonalizes both operators. If you have two operators which are both degenerate, you have to find a third operator that commutes with both of these that is not degenerate. That is why it takes at least three quantum numbers to specify an eigenstate for the hydrogen atom, two of the quantum numbers specify eigenstates of degenerate operators. The concept is more important than the mathematical mechanics, but you are not likely to comprehend the concept without mastering the math.

---

39. Consider the two operators $\mathcal{A} = \begin{pmatrix} 4 & -i & 0 \\ i & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \mathcal{B} = \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(a) Show that they are Hermitian.

(b) Show that they commute.

(c) What are $\mathcal{A}$ and $\mathcal{B}$ in the eigenbasis in which both are diagonal?

Use the same procedures as the previous problem. These operators have both real and imaginary elements which makes this problem more challenging than the last one. One of the operators is degenerate. If the characteristic equation yields duplicate eigenvalues for the operator you choose, then you want to solve the eigenvalue/eigenvector equation for the other operator and use the eigenvectors of the non-degenerate operators to form your unitary matrix.

---

40. Consider the operators $\mathcal{A} = \begin{pmatrix} 3 & i & 0 \\ -i & 3 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ $\mathcal{B} = \begin{pmatrix} -1 & 2i & 0 \\ -2i & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. What are $\mathcal{A}$ and $\mathcal{B}$ in the eigenbasis in which both are diagonal?

This should look a lot like the previous two problems. These operators are Hermitian and they commute. You may have perceived that these properties are important. If you have not, perceive so now! Physical quantities are represented by Hermitian operators. Operators that commute have a common eigenbasis so can be simultaneously diagonalized. Pick one of the two operators and solve the eigenvalue/eigenvector problem. If it is not degenerate, you can form a unitary matrix from its eigenvectors and do a unitary transformation for both operators.

An up scale restaurant with a messenger dressing in concrete hues to match his granite disposition and a mansion that Gatsby couldn't imagine complete with provocative servants. Now a posh men's store with attendants whose names were "Pierre" and "Armin." He found himself looking at a rack of $2200 suits. When "Armin" asks if he can assist my selection, maybe I'll suggest that my wardrobe requires a replacement rubberized overcoat... Before Armin could intervene, a gentleman wearing a white jacket with a red carnation in the lapel approached. He offered the greeting "How 'bout dem Bums?" to parry the gentleman's fixed attention. The gentleman responded by handing him a ticket to a Mets, Dodgers game and added the even but unexpected comment "Fourier found a need to dress warmly and should have been more cautious of stairs ... "

The Mathematics of Quantum Mechanics, Part 3

In the last two parts, calculations using small dimensional, discrete systems are featured. We will continue to illustrate calculations in small dimensional systems to provide a picture of the mathematical mechanics. These calculations are useful when working in a small dimensional subspace. They are also useful in generalizing to larger dimensional and even infinite dimensional systems. The generalization is essential because infinite dimensional systems model physical realism. Part 3 emphasizes the generalization to infinite dimensions.

Projection operators and the **completeness relation** depend upon an understanding of the **outer product**. You should have the a grasp on the rudiments of the notation quantum mechanics routinely uses to describe infinite dimensional systems in **position space**, **momentum space**, and **energy space**. The **Dirac delta function** and its derivative are used to express the second postulate of quantum mechanics. You need to understand its use within integrals and some of its properties. The **theta function** is the integral of the Dirac delta function. An infinite dimensional vector is equivalent to a continuous function. Calculations concerning infinite dimensional vectors expressed as functions are calculus problems. Any continuous function can be expressed as a **Fourier series**. Fourier series are precursors to **Fourier integrals** and **Fourier transforms** A Fourier transform is used to change the domain or the basis of a continuous function. For instance, a quantum mechanical Fourier transform is used change a state function from the position basis to the momentum basis or vice versa.

---

1. Given that $\;|\,v> = \begin{pmatrix} 1 \\ 2 \end{pmatrix}\;$ and $\;<w\,| = \begin{pmatrix} 3, & 4 \end{pmatrix},\;$ find $\;|\,v><w\,|$.

The object $|\,v><w\,|$ is an **outer product**. If $\;|\,v> = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \end{pmatrix}\;$ and $\;<w\,| = \begin{pmatrix} w_1, & w_2, & w_3, & \cdots \end{pmatrix},$

$$\text{then} \quad |\,v><w\,| = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \end{pmatrix} \begin{pmatrix} w_1, & w_2, & w_3, & \cdots \end{pmatrix} = \begin{pmatrix} v_1w_1 & v_1w_2 & v_1w_3 & \cdots \\ v_2w_1 & v_2w_2 & v_2w_3 & \cdots \\ v_3w_1 & v_3w_2 & v_3w_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

is the outer product. The outer product of two vectors is an operator.

$$|v><w| = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 3, & 4 \end{pmatrix} = \begin{pmatrix} 1 \cdot 3 & 1 \cdot 4 \\ 2 \cdot 3 & 2 \cdot 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix}$$

---

2. (a) Find the projection operator for the ket $|2>$ in the basis of unit vectors where $|1>$ is the first unit vector.

(b) Apply this projection operator to an arbitrary ket.

(c) Apply this projection operator to an arbitrary bra.

The projection operator is defined $\mathcal{P}_i = |i><i|$ for the ket $|i>$, where $|i>$ and $<i|$ are corresponding unit vectors. Here, $|2> = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}$. Use $|v> = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \end{pmatrix}$ as the arbitrary ket.

(a)
$$|2><2| = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \begin{pmatrix} 0, & 1, & 0, & \cdots \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

(b)
$$|2><2|v> = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ v_2 \\ 0 \\ \vdots \end{pmatrix}.$$

(c)
$$<v|2><2| = \begin{pmatrix} v_1, & v_2, & v_3, & \cdots \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} 0, & v_2, & 0, & \cdots \end{pmatrix}.$$

Postscript: The projection operator $|2><2|$ selects the second component from any ket or bra. It "projects" the second component from the original vector. The result of operation with the projection operator $|2><2|$ is a vector of the original type with the original second component and all other components zero. Similarly, $\mathcal{P}_4$ will project the fourth component from any vector, and $\mathcal{P}_i$ will project the i^{th} component from any vector.

---

3. What is the meaning of the consecutive projection operators $\mathcal{P}_i \mathcal{P}_j$?

Assume that the operator $\mathcal{P}_i \mathcal{P}_j$ acts on a ket to the right, for instance, $\mathcal{P}_i \mathcal{P}_j |v>$. Then $\mathcal{P}_j$ selects the j^{th} component, and all other components are zero. So $\mathcal{P}_i$ operates on a vector with one non-zero component. If the non-zero component is in the i^{th} place in the vector, the operation with $\mathcal{P}_i$ will return the same ket with one non-zero component. If the non-zero component is in other than the i^{th} place in the vector, the operation with $\mathcal{P}_i$ will return the zero vector. See if you can express these ideas symbolically using Dirac notation and the Kronecker delta.

$$\mathcal{P}_i \, \mathcal{P}_j = |\, i><i \,|\, j><j \,| = |\, i> \ \delta_{ij} \ <j \,| = |\, i><j \,| \ \delta_{ij} = |\, i><i \,| = \mathcal{P}_i \, .$$

Postscript: Notice that $< j \,|\, \delta_{ij} = < i \,|$. It says that the only condition under which this product is non-zero is when $i = j$. A state vector that is zero indicates the absence of a system. Quantum mechanics is not interested in the absence of systems so cases that are zero are usually ignored.

You could finish the problem $|\, i> \ \delta_{ij} \ <j \,| = |\, j><j \,| = \mathcal{P}_j$ if that suits your purpose.

---

4. Express the completeness relation in three dimensions.

In any given space, there are as many projection operators as the dimension of the space. Add them all to attain an identity operator, that is

$$\sum_i \mathcal{P}_i = \sum_i |\, i><i \,| = \mathcal{I} \, .$$

This is known as the **completeness relation**. It is a **resolution of the identity**, or another way to express an identity operator. It may be the most useful form of the identity for quantum mechanical calculations. This problem should illustrate the concept. Of course, the completeness relation can be expressed in arbitrary or infinite dimensions.

$$\sum_{i=1}^{3} \mathcal{P}_i = \sum_{i=1}^{3} |\, i><i \,| = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \big(1, \ 0, \ 0\big) + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \big(0, \ 1, \ 0\big) + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \big(0, \ 0, \ 1\big)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathcal{I} \, .$$

---

5. Show how the unitary transformation $\mathcal{U}^\dagger \mathcal{A} \mathcal{U}$ determines the transformation of vectors.

This problem is intended to amplify the concept of coordinate independence. It should also amplify some of the notation. In particular, one of the characteristics of Dirac notation is that it is independent of basis. Said another way, Dirac notation is coordinate independent.

An inner product is a scalar. An operator acting on a vector, such as $\mathcal{A} \,|\, v>$ is a vector, $\mathcal{A} \,|\, v> = |\, v'>$ for instance. Since $< w \,|\, \mathcal{A} \,|\, v> = < w \,|\, v'>$, it follow that $< w \,|\, \mathcal{A} \,|\, v>$ is a scalar because $<w\,|\,v'>$ is a scalar. We will encounter numerous brakets where a bra and a ket sandwich an operator. This type of braket is a scalar, just as an inner product is a scalar.

When conducting a unitary transformation on an operator, any associated bras and kets also require transformation. Insert the identity in the braket $<w\,|\,\mathcal{A}\,|\,v>$ to the immediate left and right of $\mathcal{A}$, use the fact that the identity can be expressed $\mathcal{U}\mathcal{U}^\dagger = \mathcal{I}$, and examine pieces.

$$\begin{aligned} <w\,|\,\mathcal{A}\,|\,v> &= <w\,|\,\mathcal{I}\,\mathcal{A}\,\mathcal{I}\,|\,v> \\ &= <w\,|\; \mathcal{U}\mathcal{U}^\dagger\;\mathcal{A}\;\mathcal{U}\mathcal{U}^\dagger\;|\,v> \\ &= \underbrace{\Big(<w\,|\,\mathcal{U}\Big)}_{\text{transform bras}}\;\overbrace{\Big(\mathcal{U}^\dagger\,\mathcal{A}\mathcal{U}\Big)}^{\text{transform operators}}\;\underbrace{\Big(\mathcal{U}^\dagger\,|\,v>\Big)}_{\text{transform kets}}\,. \end{aligned}$$

The object in the center set of parenthesis is our unitary transformation. For general $<w\,|$ and $|\,v>$, the last equation says we must transform a bra as $<w\,|\mathcal{U}$ and a ket as $\mathcal{U}^\dagger\,|\,v>$ to be consistent with the form of the unitary transformation we have chosen.

Postscript: Notice that if we had chosen $\mathcal{U}\,\mathcal{A}\mathcal{U}^\dagger$ as our unitary transformation, $<w\,|\,\mathcal{U}^\dagger$ would be the correct method to transform a bra and $\mathcal{U}\,|\,v>$ would be the manner to transform a ket.

---

6. Given that $|\,f> = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $|\,g> = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$, show that $<f\,|\,g>$ is invariant under a unitary transformation where $\mathcal{U} = \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} & 1/2 \end{pmatrix}$.

This numerical example is intended to illustrate the procedures and a portion of the meaning of the coordinate independence addressed in the last problem. First find $<f\,|\,g>$. Then find $<f\,|\,\mathcal{U} = <f'\,|$ and $\mathcal{U}^\dagger\,|\,g> = |\,g'>$. You will find that $<f\,|\,g> = <f'\,|\,g'>$. The unitary operator given is from problem 24 in part 2. Of course, the inner product must be the same for any unitary transformation.

$$<f\,|\,g> = \big(1,\;2,\;3\big)\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 4 + 10 + 18 = 32\,.$$

$$\mathcal{U} = \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} & 1/2 \end{pmatrix} \;\Rightarrow\; \mathcal{U}^\dagger = \begin{pmatrix} 1/2 & -1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & 1/\sqrt{2} & 1/2 \end{pmatrix}$$

$$
<f\,|\,\mathcal{U} = \left(1,\ 2,\ 3\right)\begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} & 1/2 \end{pmatrix}
$$

$$
= \left(\frac{1}{2} - \frac{2}{\sqrt{2}} + \frac{3}{2},\ \frac{1}{\sqrt{2}} - \frac{3}{\sqrt{2}},\ \frac{1}{2} + \frac{2}{\sqrt{2}} + \frac{3}{2}\right) = \left(2-\sqrt{2},\ -\sqrt{2},\ 2+\sqrt{2}\right) = \,<f'\,|
$$

$$
\mathcal{U}^\dagger\,|\,g> = \begin{pmatrix} 1/2 & -1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & 1/\sqrt{2} & 1/2 \end{pmatrix}\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 2-5/\sqrt{2}+3 \\ 4/\sqrt{2}-6/\sqrt{2} \\ 2+5/\sqrt{2}+3 \end{pmatrix} = \begin{pmatrix} 5-5/\sqrt{2} \\ -\sqrt{2} \\ 5+5/\sqrt{2} \end{pmatrix} = \,|\,g'> .
$$

$$
\begin{aligned}
<f'\,|\,g'> &= \left(2-\sqrt{2},\ -\sqrt{2},\ 2+\sqrt{2}\right)\begin{pmatrix} 5-5/\sqrt{2} \\ -\sqrt{2} \\ 5+5/\sqrt{2} \end{pmatrix} \\
&= 10 - \frac{10}{\sqrt{2}} - 5\sqrt{2} + 5 + 2 + 10 + \frac{10}{\sqrt{2}} + 5\sqrt{2} + 5 = 32\,.
\end{aligned}
$$

Postscript: The focal concept is basis independence. A vector expressed in Cartesian coordinates has different components than the same vector expressed in spherical coordinates. Similarly, the vectors $<f\,|$ and $<f'\,|$, and $|\,g>$ and $|\,g'>$, are the same vectors represented in different bases. The basis of $<f\,|$ and $|\,g>$ is Cartesian. The basis for $<f'\,|$ and $|\,g'>$ is the basis of eigenvectors of the operator $\mathcal{B}$ of problem 24 in part 2.

---

7. Provide an informal argument that a vector of infinite dimension can contain the same information as a continuous function.

The concept in this problem is essential to understanding the physics that follows. A proof is beyond the scope of this book. We explain the primary difficulty in the postscript. You may never have the formal math background to fully comprehend the proof, but you can and must comprehend the concept.

You should carefully read the following comments now and try to recreate the critical steps after you have absorbed the content.

Consider a curve segment with boundaries 0 and a. A function $f(x)$ could be used to describe this curve segment. If 20 points are thought to adequately describe this curve segment, then

$$
f(x) = f(x_1)\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + f(x_2)\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + f(x_3)\begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + f(x_{19})\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix} + f(x_{20})\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}
$$

$$= \begin{pmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \vdots \\ f(x_{20}) \end{pmatrix} = \sum_{i=1}^{20} f(x_i)\,|\,i>,$$

Figure 1-3-1.

where $|\,i>$ are unit vectors having 20 components with 1 as the i^{th} component, and the x_i are the positions of the i^{th} unit vectors on the x-axis. Though it is likely a better representation than a description using just 5 points or 10 points, 20 points cannot adequately describe a continuous function. If an arbitrarily large number, say n, is thought to adequately describe the given curve segment,

$$f(x) = f(x_1)\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + f(x_2)\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + f(x_3)\begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + f(x_{n-1})\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix} + f(x_n)\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \vdots \\ f(x_n) \end{pmatrix} = \sum_{i=1}^{n} f(x_i)\,|\,i>,$$

where $|\,i>$ are unit vectors having n components and the 1 is the i^{th} component, and the x_i are the positions of the i^{th} unit vectors on the x-axis. This is an adequate representation of a curve segment for most macroscopic purposes for adequately large n. Of course, for mathematical or quantum mechanical purposes, any finite number of $f(x_i)$ cannot describe a continuous function. Given two successive points on the x-axis, say x_j and x_k, the point midway between them is excluded so the curve segment cannot be continuous. But in the limit of $n \to \infty$, or equivalently, as the interval between successive points on the x-axis $\Delta x \to 0$,

$$f(x) = f(x_1)\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} + f(x_2)\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} + f(x_3)\begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix} + \cdots = \begin{pmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \vdots \end{pmatrix} = \sum_{i=1}^{\infty} f(x_i)\,|\,i>, \qquad (1)$$

where $|\,i>$ are unit vectors having infinite components where the 1 is the i^{th} component, and the x_i are the positions of the i^{th} unit vectors on the x-axis.

Postscript: This procedure is related to numerous applications. A beginning student of algebra will choose a finite number of points to manually graph a function. Make Δx adequately small and the eye cannot distinguish the space between the points on a curve segment. This fact is used by computer monitors, printers, digital cameras, and scanners to name just a few applications. The difference is that neither a student doing paper and pencil graphing nor a computer graphing application place the functional values in a column vector. An infinity of functional values can be placed into a column vector and that is the key point in the argument that a continuous function and a vector of infinite dimension are equivalent.

The difficult question is when do the function and the vector of infinite dimension converge? This question is addressed by the field of functional analysis. We do not address the question of convergence. However, the concept of the equivalence of continuous functions and vectors of infinite dimension introduced here informally is essential to many of the developments that are encountered in the study of quantum mechanics.

The unit vectors in this problem correspond to infinitesimally small increments in position. We are, therefore, using the **position basis** in **position space**. If the unit vectors represent infinitesimally small increments in momentum, *i.e.*, if the independent variable is momentum, the same paradigm applies except that it is using the **momentum basis** in **momentum space**. If this paradigm is used when energy is the independent variable, the unit vectors represent infinitesimally small increments in energy and the description is using the **energy basis** in **energy space**.

---

8. Write an infinite dimensional vector that is equivalent to a continuous function $f(p)$ in momentum space.

This is the same as the last problem with the exception that momentum is the independent variable. Parallel equation (1) from problem 7 using the momentum basis.

$$f(p) = f(p_1)\begin{pmatrix}1\\0\\0\\\vdots\end{pmatrix} + f(p_2)\begin{pmatrix}0\\1\\0\\\vdots\end{pmatrix} + f(p_3)\begin{pmatrix}0\\0\\1\\\vdots\end{pmatrix} + \cdots = \begin{pmatrix}f(p_1)\\f(p_2)\\f(p_3)\\\vdots\end{pmatrix} = \sum_{i=1}^{\infty} f(p_i)\,|\,i> \ .$$

---

9. Provide an informal explanation of the relations

$$<\mathbf{x}\,|\,g> = g(x)\,, \quad <\mathbf{p}\,|\,g> = g(p)\,, \quad \text{and} \quad <\mathbf{E}\,|\,g> = g(E)\,.$$

The relations introduced in this problem are fundamental. Your ability to use these relations will likely be significantly enhanced if they have meaning beyond rote memorization.

Like problem 7, study the following comments and then try to recreate the critical steps. You will encounter two new objects. Also, realize that the following argument depends directly on the fact that an infinite dimensional vector can contain the same information as a continuous function.

To this point, components of bras and kets have been scalars. Consider $<\mathbf{x}\,|$ which is not a typical bra, but a bra of unit ket vectors in position space, *i.e.*,

$$<\mathbf{x}\,| = \left(\begin{pmatrix}1\\0\\0\\\vdots\end{pmatrix},\ \begin{pmatrix}0\\1\\0\\\vdots\end{pmatrix},\ \begin{pmatrix}0\\0\\1\\\vdots\end{pmatrix},\ \cdots\right).$$

In words, $<\mathbf{x}\,|$ is a bra that represents the position basis. Also consider $|\,g>$, an arbitrary vector,

$$|\,g> = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ \vdots \end{pmatrix}.$$

Notice that no independent variable is specified for the ket $|\,g>$ or the components of $|\,g>$. This is because $|\,g>$ is completely abstract. This is another new object. It is an abstract vector with abstract components, and we know that the components exist, but we do not know what they are. We know only that the vector exists and it has a first component, a second component, and so on, ad infinitum. The abstract vector $|\,g>$ is independent of basis. Independence from basis is one of the primary features of Dirac notation. Were we to form an inner product with $<\mathbf{x}\,|$, we attain

$$<\mathbf{x}\,|\,g> = g_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} + g_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} + g_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix} + \cdots$$

where we have fixed the location of the first component of $|\,g>$ as that associated with $\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}$, so this is the first x coordinate for $|\,g>$, or equivalently, the first coordinate for $|\,g>$ is at x_1. The location of the second component of $|\,g>$ is that associated with $\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}$, so this is the second x coordinate for $|\,g>$, or the second coordinate for $|\,g>$ is at x_2, and so on, ad infinitum. We have fixed each component of $|\,g>$ at a specific location in the position basis. In other words, it is the same as

$$<\mathbf{x}\,|\,g> = g\,(x_1) \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} + g\,(x_2) \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} + g\,(x_3) \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix} + \cdots = \begin{pmatrix} g\,(x_1) \\ g\,(x_2) \\ g\,(x_3) \\ \vdots \end{pmatrix} = g\,(x)\,.$$

In short $<\mathbf{x}\,|\,g> = g\,(x)$, the inner product of an abstract ket with the bra that describes the position basis is equivalent to a function of position.

There exists a bra of unit ket vectors that describes the momentum basis, and another bra of unit kets that describes the energy basis. The arguments in momentum space and energy space are only cosmetically different than that for position space. The inner product of an abstract ket with the bra of unit ket vectors that describes the momentum basis is equivalent to a function of momentum, and the inner product of an abstract ket with the bra of unit ket vectors that describes the energy basis is equivalent to a function of energy. This means that $<\mathbf{p}\,|\,g> = g\,(p)$ and $<\mathbf{E}\,|\,g> = g\,(E)$.

Postscript: These three relations are known as **representations** of the abstract vector $|\, g >$. The last two relations are refined $< \mathbf{p}\,|\, g > \,=\, \widehat{g}\,(p)$ and $< \mathbf{E}\,|\, g > \,=\, \widetilde{g}\,(E)$, using a hat for a quantity in momentum space and a tilde for a quantity in energy space to clearly denote the momentum or energy basis.

The first postulate of quantum mechanics indicates that a state vector $|\,\psi >$ describes a system that we know to exist. Like $|\, g >$ in this problem, the abstract ket $|\,\psi >$ is a function without an argument. The functional argument is present in the basis chosen. The concept is that $|\,\psi >$ exists and is expressible in <u>any</u> appropriate basis. A bra of unit kets is used to attain a functional form $\psi\,(x)$, or $\widehat{\psi}\,(p)$, or $\widetilde{\psi}\,(E)$, like $< \mathbf{x}\,|\,\psi > \,=\, \psi\,(x)$ for instance. You cannot calculate abstract components of g_i or ψ_i without reference to a basis just the same as you cannot express the components of a two-dimensional vector without reference to Cartesian or polar or another coordinate system. You may want to calculate $g\,(x_i)$ or $\widehat{\psi}\,(p_i)$ after the ket is represented. Representations are often used to transition from Dirac notation and the language of linear algebra to the language of calculus.

The symbol between the slashes and the angles for the bras are boldfaced in this problem to distinguish the new object. Most authors require the reader make the distinction between the bra of unit kets and a bra whose components are scalars. We will use the conventional notation $<x\,|\,g> = g\,(x)$, $<p\,|\,g> = g\,(p)$, and $<E\,|\,g> = g\,(E)$, following the introduction presented in this and the next problems.

---

10. Provide an informal explanation of the representation relations in the dual spaces, *i.e.*,

$$<g\,|\,\mathbf{x}> = g^*\,(x)\,, \quad <g\,|\,\mathbf{p}> = \widehat{g}^*\,(p)\,, \quad \text{and} \quad <g\,|\,\mathbf{E}> = \widetilde{g}^*\,(E)\,.$$

This problem extends the concept provided in problem 9 to the dual space. It also exposes pertinent notation and its meaning. Again, you are likely to have more utility with these relations if you have an idea of their meaning.

This problem parallels problem 9. The ket $|\,\mathbf{x}>$ represents the entire basis in the dual space. It is a ket of unit bras. You need only to develop the argument in position space and state that the development of the momentum space and energy space cases are identical except that the different bases result in different functional arguments. Remember that the components of a ket that is the dual to a given bra has components that are conjugates of the given bra.

$$|\,\mathbf{x}> \;=\; \begin{pmatrix} (1,\ 0,\ 0,\ \cdots) \\ (0,\ 1,\ 0,\ \cdots) \\ (0,\ 0,\ 1,\ \cdots) \\ \vdots \end{pmatrix} \quad \Rightarrow \; <g\,|\,\mathbf{x}> = \big(g_1,\ g_2,\ g_3, \cdots\big) \begin{pmatrix} (1,\ 0,\ 0,\ \cdots) \\ (0,\ 1,\ 0,\ \cdots) \\ (0,\ 0,\ 1,\ \cdots) \\ \vdots \end{pmatrix}$$

$$= g_1\big(1,\ 0,\ 0,\ \cdots\big) + g_2\big(0,\ 1,\ 0,\ \cdots\big) + g_3\big(0,\ 0,\ 1,\ \cdots\big) + \cdots$$

$$= \big(g\,(x_1),\ 0,\ 0,\ \cdots\big) + \big(0,\ g\,(x_2),\ 0,\ \cdots\big) + \big(0,\ 0,\ g\,(x_3),\ \cdots\big) + \cdots$$

$$= \big(g\,(x_1),\ g\,(x_2),\ g\,(x_3),\ \cdots\big) = \begin{pmatrix} g^*\,(x_1) \\ g^*\,(x_2) \\ g^*\,(x_3) \\ \vdots \end{pmatrix} = g^*\,(x)\,.$$

The inner product of $<g\,|$ with the ket of unit bras that represents the momentum basis fixes the components of $<g\,|$ in momentum space so $<g\,|\,\mathbf{p}> = \widehat{g}^*\,(p)$. The inner product of $<g\,|$ with the ket of unit bras that represents the energy basis fixes the components of $<g\,|$ in energy space so $<g\,|\,\mathbf{E}> = \widetilde{g}^*\,(E)$.

---

11. Evaluate

(a) $2\int_{-\pi}^{\pi} \cos(2\theta)\,\delta(\theta)\,d\theta,$

(b) $\int_{0}^{\infty} A\,\cos(x)\,e^{x^2}\,\delta(x+1)\,dx,$

(c) $\int \cos\left(\frac{n\pi x}{l}\right)\,\sin\left(\frac{n\pi x}{l}\right)\,\delta(x-l)\,dx, \quad n = 1, 2, 3, \ldots$

(d) $\int_{0}^{1} dr \int_{-\infty}^{\infty} d\theta\,\frac{1}{r^2}\,\sin(\theta)\,\delta(r-a)\,\delta\left(\theta - \frac{\pi}{4}\right), \quad 0 < a < 1\,.$

This problem introduces the procedures to evaluate integrals that contain **Dirac delta functions**, often called just delta functions or δ functions. The Dirac delta function is defined

$$\int_{-\infty}^{\infty} \delta(x)\,dx = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x \neq 0 \end{cases}$$

To get a picture for a delta function, consider a normalized function of the real variable x that vanishes everywhere except inside a domain of length ϵ. The integral over this domain is unity because it is given to be normalized. If we diminish the domain, the function must become "taller" to remain normalized. In the limit of $\epsilon \to 0$, the functional value goes to ∞. The delta function is an idealization of an infinitely high, infinitely thin spike that is normalized. We refer you to Boas[1], Arfken[2], or your favorite text on mathematical physics for greater depth.

Following Dirac, the most important property of $\delta(x)$ is

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

Informally, you can picture the value of $\delta(x)$ being zero everywhere except at the origin, so zero is the only location at which the product of the function $f(x)$ and the delta function is non-zero. The value of the function at the origin is $f(x)\big|_{x=0} = f(0)$, which is a constant, so can be moved outside of the integral. The result is the constant $f(0)$ times the defining integral which has a value of 1, or $f(0)\cdot 1 = f(0)$. The Dirac delta function in an integrand "picks out" the value of the function evaluated at the origin. If we "shift" the origin to $x = a$, the last equation becomes

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

[1] Boas, *Mathematical Methods in the Physical Sciences* (John Wiley & Sons, New York, 1983), 2nd ed., pp. 665-670.

[2] Arfken, *Mathematical Methods for Physicists* (Academic Press, New York, 1970), 2nd ed., pp. 413-415.

If the value that makes the argument of a delta function zero is not between the limits of integration, the integral of that delta function is zero on that interval, thus the product is zero everywhere, and therefore the original integral is zero. For example, if $f(x) = A\,e^{x^2}$, then

$$\int_{-\infty}^{\infty} A\,e^{x^2}\,\delta(x)\,dx = \left(A\,e^{x^2}\Big|_{x=0}\right)\int_{-\infty}^{\infty}\delta(x)\,dx = A\,e^{0}\,(1) = A\,.$$

$$\int_{-\infty}^{\infty} A\,e^{x^2}\,\delta(x-a)\,dx = \left(A\,e^{x^2}\Big|_{x=a}\right)\int_{-\infty}^{\infty}\delta(x-a)\,dx = A\,e^{a^2}\,(1) = A\,e^{a^2}\,.$$

$$\int_{-\infty}^{0} A\,e^{x^2}\,\delta(x-1)\,dx = 0 \quad \text{because 1 is not between the limits of integration.}$$

One of the four integrals is zero because the value that makes the argument of the delta function zero is not between the limits of integration. The integral for part (c) has no limits. It is common to suppress the limits $-\infty$ and ∞, particularly when these limits are encountered repetitively. Part (d) has two variables so you must address each coordinate independently.

(a) The argument of the delta function is zero at $\theta = 0$, so

$$2\int_{-\pi}^{\pi}\cos(2\theta)\,\delta(\theta)\,d\theta = 2\,\cos(2\theta)\Big|_{\theta=0} = 2\,\cos(0) = 2\cdot 1 = 2\,.$$

(b) The argument of the delta function is zero at $x + 1 = 0 \quad\Rightarrow\quad x = -1$, which is not within the limits of integration, therefore

$$\int_{0}^{\infty} A\,\cos(x)\,e^{x^2}\,\delta(x+1)\,dx = 0\,.$$

(c) The argument of the delta function is zero at $x - l = 0 \quad\Rightarrow\quad x = l$, so

$$\int_{-\infty}^{\infty}\cos\left(\frac{n\pi x}{l}\right)\sin\left(\frac{n\pi x}{l}\right)\delta(x-l)\,dx = \cos\left(\frac{n\pi x}{l}\right)\sin\left(\frac{n\pi x}{l}\right)\Big|_{x=l}$$

$$= \cos(n\pi)\,\sin(n\pi) = \cos(n\pi)\,(0) = 0 \quad \text{for all } n\,.$$

(d) The argument of the angular delta function is zero at $x = \pi/4$ and the argument of the radial delta function is zero at $r = a$, so addressed in the order of integration,

$$\int_0^1 dr\int_{-\infty}^{\infty} d\theta\,\frac{1}{r^2}\,\sin(\theta)\,\delta(r-a)\,\delta\left(\theta-\frac{\pi}{4}\right) = \int_0^1 dr\,\frac{1}{r^2}\left(\Big[\sin(\theta)\Big]_{\theta=\pi/4}\right)\delta(r-a)$$

$$= \int_0^1 dr\,\frac{1}{r^2}\,\frac{\sqrt{2}}{2}\,\delta(r-a) = \frac{\sqrt{2}}{2}\int_0^1 dr\,\frac{1}{r^2}\,\delta(r-a) = \frac{\sqrt{2}}{2}\,\frac{1}{r^2}\Big|_{r=a} = \frac{\sqrt{2}}{2a^2} \quad \text{assuming } 0 < a < 1\,.$$

Postscript: The Dirac delta function is a generalization of the Kronecker delta. It is usually considered a generalized function[3] or distribution, though Dirac, the originator, considered it an

[3] Lighthill, *An Introduction to Fourier Analysis and Generalised Functions* (Cambridge University Press, Cambridge, England, 1958).

improper function. Some authors will use it only as a factor in an integrand. Even Dirac says "...it will be something which is to be used ultimately in an integrand." Nevertheless, delta functions commonly appear outside an integral in modern literature. The Dirac delta function is an idealization, much the same way as a mathematical point will model the location of a classical particle. Dirac indicates it is "...merely a convenient notation, enabling us to express in a concise form certain relations which we could, if necessary, rewrite in a form not involving improper functions, but only in a cumbersome way which would tend to obscure the argument."[4] If the system is discrete, use of the Kronecker delta is appropriate to model an "all or none" quantity. If the system is continuous, use of the Dirac delta function is appropriate to model an "all or none" situation.

The Dirac delta function is an even function. A function is even if $f(-x) = f(x)$, like a cosine, and a function is odd if $f(-x) = -f(x)$, like a sine. The graph of an even function is symmetric about the vertical axis. If the graph of an even function is folded on the vertical axis, the two half curves overlap. Notice that figure 1–3–2 is symmetric about the vertical axis. The graph of an odd function is symmetric about the origin. If the graph of an odd function is rotated 180 degrees around the origin, the two half curves overlap. Most functions do not have any central symmetry so are neither even nor odd. For instance, $f(x) = x + 1$ is neither even nor odd. The fact that the Dirac delta function is even is often expressed

$$\delta(-x) = \delta(x) \quad \text{and} \quad \delta(x - x') = \delta(x' - x).$$

This notation requires explanation because a delta function is "something which is to used ultimately in an integrand." A function and integral with the limits $-\infty$ and ∞ are implicit, *i.e.*,

$$\delta(x - x') = \delta(x' - x) \quad \Rightarrow \quad \int_{-\infty}^{\infty} f(x)\,\delta(x - x')\,dx = \int_{-\infty}^{\infty} f(x)\,\delta(x' - x)\,dx.$$

The function, integral, and limits often provide only clutter. These can be attached to complete the meaning of relationships involving "naked" delta functions, as seen in the next problem.

---

12. Show that $\delta(ax) = \dfrac{\delta(x)}{|a|}, \quad a \neq 0.$

The second postulate of quantum mechanics is often stated using a delta function and its derivative. Some ability to use delta functions is necessary. Problems 12 to 17 are oriented toward building dexterity and exposing some of properties of the delta function, its integral, and its derivative.

Approach this problem by changing variables. Let

$$y = ax \quad \Rightarrow \quad x = \frac{y}{a} \quad \Rightarrow \quad dx = \frac{dy}{a}. \quad \text{Change the variables in} \quad \int_{-\infty}^{\infty} f(x)\,\delta(ax)\,dx$$

for an arbitrary $f(x)$. There are two cases, $a > 0$ and $a < 0$. Pay attention to the limits of integration to get the absolute value of a to emerge. You want to show that

$$\int_{-\infty}^{\infty} f(x)\,\delta(ax)\,dx = \int_{-\infty}^{\infty} f(x)\,\frac{\delta(x)}{|a|}\,dx,$$

[4] Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, England, 1958), 4th ed., pp. 58-59.

and that is the explicit meaning of the statement $\delta(ax) = \dfrac{\delta(x)}{|a|}$.

$$y = ax \quad \Rightarrow \quad x = \frac{y}{a} \quad \Rightarrow \quad dx = \frac{dy}{a}. \quad \text{Consider} \quad \int_{-\infty}^{\infty} f(x)\,\delta(ax)\,dx$$

for an arbitrary $f(x)$. There are two cases. Case I. $a > 0 \;\Rightarrow\; \infty \to \infty$ and $-\infty \to -\infty$, so

$$\int_{-\infty}^{\infty} f(x)\,\delta(ax)\,dx = \int_{-\infty}^{\infty} f\left(\frac{y}{a}\right)\delta(y)\,\frac{dy}{a} = \frac{1}{a}\int_{-\infty}^{\infty} f\left(\frac{y}{a}\right)\delta(y)\,dy = \frac{f(0)}{a} = \frac{f(0)}{|a|}$$

since $a = |a|$. Case II. $a < 0 \;\Rightarrow\; \infty \to -\infty$ and $-\infty \to \infty$, so

$$\int_{-\infty}^{\infty} f(x)\,\delta(ax)\,dx = \int_{\infty}^{-\infty} f\left(\frac{y}{a}\right)\delta(y)\,\frac{dy}{a} = -\frac{1}{a}\int_{-\infty}^{\infty} f\left(\frac{y}{a}\right)\delta(y)\,dy = -\frac{f(0)}{a} = \frac{f(0)}{|a|}$$

since $a < 0$. Combining cases I and II for a general constant a,

$$\int_{-\infty}^{\infty} f(x)\,\delta(ax)\,dx = \frac{f(0)}{|a|} = \int_{-\infty}^{\infty} f(x)\,\frac{\delta(x)}{|a|}\,dx, \quad \text{therefore,} \quad \delta(ax) = \frac{\delta(x)}{|a|}. \quad \text{Q.E.D.}$$

Postscript: The motivation to change variables is to attain a delta function with a single variable in the argument, like $\delta(x)$ or $\delta(r)$. The meaning as part of an integrand is then clear.

---

13. Show that $\quad \delta\Big(f(x)\Big) = \displaystyle\sum_i^n \frac{\delta(x_i - x)}{|\,f'(x_i)\,|} \quad$ where the x_i are the zeros of $f(x)$.

The integral $\int \delta\big(f(x)\big)\,g(x)\,dx$ will have a non-zero value where the argument of the delta function is zero, *i.e.*, where $f(x) = 0$. To find the zeros, expand $f(x)$ in a Taylor series. Near each zero,

$$f(x) \simeq (x - x_i)\,f'(x_i), \qquad \text{where} \qquad f(x_i) = 0$$

and quadratic and higher order terms are negligible, so only terms with first derivatives remain. Substitute each $(x - x_i)\,f'(x_i)$, into the argument of the delta function. You then have a series of integrals containing delta functions. Each delta function has an argument that is a product of a difference and a first derivative evaluated at a point. The first derivative evaluated at a point is a scalar so you can use the result of problem 12. You need to use the fact that the delta function is even to attain the order of the argument that is specified in the problem. Express the result as a summation. The meaning of the premise is

$$\int_{-\infty}^{\infty} \delta\Big(f(x)\Big)\,g(x)\,dx = \int_{-\infty}^{\infty} \sum_i^n \frac{\delta(x_i - x)}{|\,f'(x_i)\,|}\,g(x)\,dx.$$

Consider $f(x)$ which is a continuous, analytic function on the interval $-\infty < x < \infty$. Then $f(x)$ may be expanded in a Taylor series

$$f(x) = f(x_i) + (x - x_i)\, f'(x_i) + \frac{(x - x_i)^2}{2!}\, f''(x_i) + \frac{(x - x_i)^3}{3!}\, f'''(x_i) + \cdots$$

Each x_i is a zero of $f(x) \;\Rightarrow\; f(x) \simeq (x - x_i)\, f'(x_i)$ at points near x_i, where $f(x_i) = 0$, and quadratic and higher order terms are negligible near zeros. Under these conditions the integral is

$$\begin{aligned}\int_{-\infty}^{\infty} \delta\Big(f(x)\Big)\, g(x)\, dx \simeq& \int_{-\infty}^{\infty} \delta\Big((x - x_1)\, f'(x_1)\Big)\, g(x)\, dx \\ &+ \int_{-\infty}^{\infty} \delta\Big((x - x_2)\, f'(x_2)\Big)\, g(x)\, dx + \int_{-\infty}^{\infty} \delta\Big((x - x_3)\, f'(x_3)\Big)\, g(x)\, dx + \cdots \\ =& \int_{-\infty}^{\infty} \frac{\delta(x - x_1)}{|\, f'(x_1)\,|}\, g(x)\, dx + \int_{-\infty}^{\infty} \frac{\delta(x - x_2)}{|\, f'(x_2)\,|}\, g(x)\, dx + \int_{-\infty}^{\infty} \frac{\delta(x - x_3)}{|\, f'(x_3)\,|}\, g(x)\, dx + \cdots\end{aligned}$$

using the result of the problem 12. Expressing the series in terms of a summation and using the fact that the Dirac delta function is even,

$$\int_{-\infty}^{\infty} \delta\Big(f(x)\Big)\, g(x)\, dx = \sum_i^n \int_{-\infty}^{\infty} \frac{\delta(x - x_i)}{\big|\, f'(x_i)\,\big|}\, g(x)\, dx = \sum_i^n \int_{-\infty}^{\infty} \frac{\delta(x_i - x)}{\big|\, f'(x_i)\,\big|}\, g(x)\, dx\,.$$

The last expression is the same as the integral of the sum, essentially reversing the direction of $\int \big(h(x) + j(x)\big)\, dx = \int h(x)\, dx + \int j(x)\, dx$, so can be written

$$\int_{-\infty}^{\infty} \delta\Big(f(x)\Big)\, g(x)\, dx = \int_{-\infty}^{\infty} \sum_i \frac{\delta(x_i - x)}{\big|\, f'(x_i)\,\big|}\, g(x)\, dx\,, \;\Rightarrow\; \delta\Big(f(x)\Big) = \sum_i \frac{\delta(x_i - x)}{\big|\, f'(x_i)\,\big|}\,. \quad \text{Q.E.D}$$

---

14. Show that $\displaystyle\int \delta(x)\, dx = \Theta(x)$.

This problem introduces the **theta function**. The delta function is the derivative of the theta function, as this problem demonstrates in integral form. The theta function is useful in establishing some of the properties of the delta function and in problems involving discontinuities.

The theta function, also known as the step function or the Heaviside step function, is defined

$$\Theta(x) = \begin{cases} 1\,, & \text{if } x > 0\,, \\ 0\,, & \text{if } x < 0\,. \end{cases}$$

If we "shift" the origin to $x = x'$,

$$\Theta(x - x') = \begin{cases} 1\,, & \text{if } x > x'\,, \\ 0\,, & \text{if } x < x'\,. \end{cases}$$

The last equation is illustrated in the figure. Notice that it is not defined at $x = 0$, and it is not continuous at $x = 0$. Like the delta function, the theta function is another "improper" or "pathological" function. It does not possess the properties of continuity and differentiability at the location where the argument is equal to zero.

As usual, assume the limits of integration are $-\infty$ and ∞ when a proper integral is anticipated and limits are not stated. Limits are the key to the meaning of the premise. Break the integral into two integrals at an arbitrarily small $\epsilon < 0$. This means the limits on the first integral are $-\infty$ to $-\epsilon$ and the limits on the second integral are $-\epsilon$ to ∞. Examine the behavior of both integrals in the limit $\epsilon \to 0$.

$$\int_{-\infty}^{\infty} \delta(x)\, dx = \int_{-\infty}^{-\epsilon} \delta(x)\, dx + \int_{-\epsilon}^{\infty} \delta(x)\, dx,$$

where ϵ is an arbitrarily small number. There are two cases, either $x > 0$ or $x < 0$. If $x > 0$,

$$\int_{-\infty}^{-\epsilon} \delta(x)\, dx + \int_{-\epsilon}^{\infty} \delta(x)\, dx = 0 + 1 = 1\,,$$

where the first integral is zero because values of $x > 0$ are not within the limits of the integral, and the second integral is 1 per the definition of the Dirac delta function. If $x < 0$, then

$$\int_{-\infty}^{-\epsilon} \delta(x)\, dx + \int_{-\epsilon}^{\infty} \delta(x)\, dx = 0 + 0 = 0\,,$$

where the first integral is zero because the argument of the delta function cannot assume the value of $x = 0$ within limits that are both less than zero. The second integral is zero because the values $x < 0$ are not within the limits of the integral. If the circumstance $-\epsilon < x < 0$ were to occur, the second integral would be 1, but picking another ϵ such that $x < -\epsilon$ restores the situation in which both integrals vanish. Therefore, in the limit that $\epsilon \to 0$,

$$\int \delta(x)\, dx = \begin{cases} 1\,, & \text{if } x > 0\,, \\ 0\,, & \text{if } x < 0\,. \end{cases} \;=\; \Theta(x)\,.$$

Postscript: The differential form of this relationship is $\dfrac{d\,\Theta(x)}{dx} = \delta(x)$. The integral form is often more useful for calculation because of sensitivity to limits.

---

properties of delta parts, I know of no book that actually has an problem 17 may be unique in modern literature. By the way, since I confirm this procedure through any of the research that I have done, I hope that both

15. Use integration by parts to show that $\displaystyle\int_a^c f(x)\,\delta(x-b)\, dx = f(b)\,, \quad a < b < c\,.$

As Dirac indicates, the fact that $\int f(x)\,\delta(x)\,dx = f(0)\,,$ is a property. Establishing properties and understanding the methods by which they are established is more useful than rote memorization. Many of the properties of delta functions are established using integration by parts. For a definite integral, integration by parts means

$$\int_a^c u\,dv = uv\Big|_a^c - \int_a^c v\,du\,.$$

You need the result of the last problem to apply this formula. You also need to use the technique of breaking the resulting $\int_a^c v\,du$ into two integrals as demonstrated in the last problem. Use ϵ slightly less than b. You need the definition of the theta function to evaluate every term following the integration by parts. Finally, taking the limit $\epsilon \to 0$ establishes the premise.

$$\text{Given} \quad \int_a^c f(x)\,\delta(x-b)\,dx, \quad \text{and using} \quad \begin{array}{rclcrcl} u &=& f(x) & \Rightarrow & du &=& f'(x)\,dx \\ dv &=& \delta(x-b)\,dx & \Rightarrow & v &=& \Theta(x-b) \end{array}$$

$$\Rightarrow \int_a^c f(x)\,\delta(x-b)\,dx = f(x)\,\Theta(x-b)\Big|_a^c - \int_a^c \Theta(x-b)\,f'(x)\,dx$$

$$= f(c)\,\Theta(c-b) - f(a)\,\Theta(a-b) - \int_a^{b-\epsilon} \Theta(x-b)\,f'(x)\,dx - \int_{b-\epsilon}^c \Theta(x-b)\,f'(x)\,dx \qquad (1)$$

where ϵ is arbitrarily close to zero. The condition $a < b < c$ determines the first two Θ functions in equation (1) as $\Theta(c-b) = 1\,,$ since $c > b$ and $\Theta(a-b) = 0$ since $a < b$. Further,

$$\int_a^{b-\epsilon} \Theta(x-b)\,f'(x)\,dx = 0$$

since $\Theta(x-b) = 0$ because $x < b$ for all possible values of x within the range of integration. Similarly,

$$\int_{b-\epsilon}^c \Theta(x-b)\,f'(x)\,dx = \int_{b-\epsilon}^c f'(x)\,dx$$

since $\Theta(x-b) = 1$ for all possible values of x within that range of integration because $x > b$. Using all these in equation (1),

$$\begin{aligned} \int_a^c f(x)\,\delta(x-b)\,dx &= f(c)\cdot 1 - f(a)\cdot 0 - 0 - \int_{b-\epsilon}^c f'(x)\,dx \\ &= f(c) - f(x)\Big|_{b-\epsilon}^c \\ &= f(c) - f(c) + f(b-\epsilon) = f(b-\epsilon)\,. \end{aligned}$$

In the limit of $\epsilon \to 0$, $f(b-\epsilon) \to f(b)\,,$ and the integral is

$$\int_a^c f(x)\,\delta(x-b)\,dx = f(b)\,.$$

Postscript: The argument for infinite limits is the same except $a = -\infty$ and $c = \infty$.

---

16. Evaluate

(a) $\displaystyle\int V_0 \left(\frac{a}{x} - \frac{x}{a} \right)^2 \delta'(x-a)\, dx$, and

(b) $\displaystyle\int V_0 \left(\frac{a}{x} - \frac{x}{a} \right)^2 \delta'(x-b)\, dx$.

This problem introduces derivatives of delta functions. Like delta functions, derivatives of delta functions are always intended to be used as a portion of an integrand though they will frequently appear without an integral or companion function for reasons of notational economy. The derivative of the Dirac delta function is the negative of the derivative of the companion function evaluated at the point that makes the argument of δ' zero. Symbolically,

$$\int \delta'(x-a)\, f(x)\, dx = -f'(a).$$

For $f(x) = x^2$, $f'(x) = 2x$, so $\displaystyle\int \delta'(x-a)\, x^2\, dx = -2a$, for instance. A second example is $f(x) = A\, e^{-x^2} \;\Rightarrow\; f'(x) = A\, e^{-x^2}(-2x) = -2A\, xe^{x^2}$, so $\displaystyle\int \delta'(x-b) A\, e^{x^2}\, dx = 2A\, be^{b^2}$. We could attach the limits understood to be $-\infty$ and ∞ but they add only clutter here so have been suppressed. The potential given in this problem is realistic and is sometimes used to model an assymetric potential well in radial coordinates (*i.e.*, $x \to r$ and $r > 0$).

(a) The derivative of the function is

$$V_0 \frac{d}{dx} \left(ax^{-1} - \frac{x}{a} \right)^2 = V_0\, 2 \left(\frac{a}{x} - \frac{x}{a} \right) \left(-ax^{-2} - \frac{1}{a} \right) = -2\, V_0 \left(\frac{a}{x} - \frac{x}{a} \right) \left(\frac{a}{x^2} + \frac{1}{a} \right)$$

$$\Rightarrow \quad \int V_0 \left(\frac{a}{x} - \frac{x}{a} \right)^2 \delta'(x-a)\, dx = -\left[-2\, V_0 \left(\frac{a}{x} - \frac{x}{a} \right) \left(\frac{a}{x^2} + \frac{1}{a} \right) \right]_{x=a}$$

$$= 2\, V_0 \left(\frac{a}{a} - \frac{a}{a} \right) \left(\frac{a}{a^2} + \frac{1}{a} \right) = 2\, V_0\, (0) \left(\frac{2}{a} \right) = 0.$$

(b) Having calculated the derivative for part (a),

$$\int V_0 \left(\frac{a}{x} - \frac{x}{a} \right)^2 \delta'(x-b)\, dx = -\left[-2\, V_0 \left(\frac{a}{x} - \frac{x}{a} \right) \left(\frac{a}{x^2} + \frac{1}{a} \right) \right]_{x=b}$$

$$= 2\, V_0 \left(\frac{a}{b} - \frac{b}{a} \right) \left(\frac{a}{b^2} + \frac{1}{a} \right) = 2\, V_0 \left(\frac{a^2}{b^3} + \frac{1}{b} - \frac{1}{b} - \frac{b}{a^2} \right)$$

$$= 2\, V_0 \left(\frac{a^2}{b^3} - \frac{b}{a^2} \right).$$

---

17. Show that $\displaystyle\int \delta'(x-a)\, f(x)\, dx = -f'(a)$ using integration by parts.

This problem is meant to further develop your appreciation of the Dirac delta function and its derivative. It also further illustrates the process of integration by parts that is consistently referenced but rarely demonstrated in discussions of the Dirac delta function.

Consider $\displaystyle\int_a^c \delta'(x-b)\, f(x)\, dx$ for $a < b < c$. Integration by parts is reviewed in problem 15. You are left with two terms that have delta functions that are not part of an integrand following the integration by parts. Both terms vanish because the arguments of the delta functions cannot be zero. The rest is straightforward.

$$\text{Given} \quad \int_a^c f(x)\, \delta'(x-b)\, dx\,, \quad \text{let} \quad \begin{array}{rclcrcl} u & = & f(x) & \Rightarrow & du & = & f'(x)\, dx \\ dv & = & \delta'(x-b)\, dx & \Rightarrow & v & = & \delta(x-b) \end{array}$$

$$\Rightarrow \quad \int_a^c \delta'(x-b)\, f(x)\, dx = f(x)\, \delta(x-b)\Big|_a^c - \int_a^c \delta(x-b)\, f'(x)\, dx\,.$$

The term $f(x)\, \delta(x-b)\Big|_a^c = f(c)\, \delta(c-b) - f(a)\, \delta(a-b) = 0$ because $c-b \neq 0$, and $a-b \neq 0$, so the argument of neither delta function can be zero, therefore, the integrals with respect to any variable of integration is necessarily zero. More formally,

$$\begin{aligned} \int dc \left(\int_a^c \delta'(x-b)\, f(x)\, dx \right) &= \int f(c)\, \delta(c-b)\, dc \\ &\quad + \int dc \left(-f(a)\, \delta(a-b) - \int_a^c \delta(x-b)\, f'(x)\, dx \right) \\ &= 0 + \int dc \left(-f(a)\, \delta(a-b) - \int_a^c \delta(x-b)\, f'(x)\, dx \right) \end{aligned}$$

$$\Rightarrow \quad \int_a^c \delta'(x-b)\, f(x)\, dx = -f(a)\, \delta(a-b) - \int_a^c \delta(x-b)\, f'(x)\, dx\,,$$

and then repeat the process with respect to a. In either event,

$$\int_a^c \delta'(x-b)\, f(x)\, dx = -\int_a^c \delta(x-b)\, f'(x)\, dx = -f'(b) \quad \text{since} \quad a < b < c\,.$$

to you, (it looks on the page.

---

18. Establish relations applicable to continuous systems in position space, momentum space, and energy space, for

(a) the orthonormality condition and

(b) the completeness relation.

Infinities are inherent in descriptions of systems that are continuous. The infinite dimensional vector that contains the same information as a continuous function is but one example. Though not mathematically rigorous, physicists frequently reason using finite dimensional analogs treating the infinite as a generalization[5]. This seems a practical if not a mathematical necessity.

There is enough new notation in this problem that we recommend that you simply read the discussion carefully. If you prefer to venture an effort, realize that in position space, for instance,

$$i \ \text{ or } \ x_i \longrightarrow x\,, \quad \delta_{ij} \longrightarrow \delta\,(x - x')\,, \quad \text{and} \quad \sum_i \longrightarrow \int dx\,,$$

when generalizing from the discrete to the continuous. Suppress the limits of $-\infty$ and ∞ to keep the notation clean.

(a) The orthonormality condition is denoted $< i \,|\, j > = \delta_{ij}$ in general. If specifically in position space, it can be written $< x_i \,|\, x_j > = \delta_{ij}$ where x_i and x_j denote unit vectors that corresponded to the i^{th} and j^{th} positions on the x-axis. If $i = j$, then $< x_i \,|\, x_j > = 1$. Let the distance between unit vectors approaches zero, modelling a continuous system. The discrete x_i and x_j become indistinguishable as the distance between them approaches zero, so are superceded by x and x' for continuous systems. Symbology that describes coincidence or non-coincidence is then,

$$< x_i \,|\, x_j > = \delta_{ij} \quad \longrightarrow \quad < x \,|\, x' > = \delta\,(x - x') \ \text{ for a continuous system in position space.}$$

If the argument is made in momentum space, $< p_i \,|\, p_j > = \delta_{ij} \quad \longrightarrow \quad < p \,|\, p' > = \delta\,(p - p')\,.$

If the argument is made in energy space, $< E_i \,|\, E_j > = \delta_{ij} \quad \longrightarrow \quad < E \,|\, E' > = \delta\,(E - E')\,.$

(b) The completeness relation is $\displaystyle\sum_{i=1}^{n} |\, i >< i \,| = \mathcal{I}$ or $\displaystyle\sum_{i=1}^{n} |\, x_i >< x_i \,| = \mathcal{I}$ for position space. In the limit as $n \to \infty$, $|\, x_i > \to |\, x >$ and $\displaystyle\sum_i \to \int dx$ in position space, so

$$\sum_{i=1}^{n} |\, x_i >< x_i \,| = \mathcal{I} \quad \longrightarrow \quad \int |\, x >< x \,|\, dx = \mathcal{I} \ \text{ for position space,}$$

$$\sum_{i=1}^{n} |\, p_i >< p_i \,| = \mathcal{I} \quad \longrightarrow \quad \int |\, p >< p \,|\, dp = \mathcal{I} \ \text{ for momentum space, and}$$

$$\sum_{i=1}^{n} |\, E_i >< E_i \,| = \mathcal{I} \quad \longrightarrow \quad \int |\, E >< E \,|\, dE = \mathcal{I} \ \text{ for energy space}\,.$$

Postscript: The relations established in this problem should be viewed as definitions because of the question of convergence. The extension $\displaystyle\sum_i \to \int dx$ is subject to the question of

[5] Cohen-Tannoudji, Diu, & Laloe, *Quantum Mechanics* (John Wiley & Sons, New York, 1977), 4th ed., pp. 94.

convergence but is more fundamentally a redefinition. The development of a Riemann integral often concerns the area under a curve where the area is approximated by $\sum_i^n f(x_i)\,\Delta x$. Then as $n \to \infty$, $\Delta x \to 0$. The Δx in this development is the distance between the infinite dimensional unit vectors, so Δx is implicit in the index i for the extension $\sum_i \to \int dx$.

---

19. Use the technique of insertion of the identity to establish expressions applicable to continuous systems in position space, momentum space, and energy space, for an inner product.

This problem establishes a fundamental tenet while applying representations. It should start to convince you of the economy and convenience inherent in Dirac notation. Each of the three solutions require only about one line.

Start with an inner product of two abstract vectors, say $< f\,|\,g >$. Insert the identity to get $< f\,|\,\mathcal{I}\,|\,g >$. Use the completeness relation for a continuous system from problem 18 appropriate to each space, then find the representations of problems 9 and 10. You should find $<f\,|\,g> = \int f^*(x)\,g(x)\,dx$ for position space, for instance.

$$<f\,|\,g> = <f\,|\,\mathcal{I}\,|\,g> = <f\,|\left(\int |\,x><x\,|\,dx\right)|\,g> = \int <f\,|\,x><x\,|\,g>\;dx = \int f^*(x)\,g(x)\,dx\,.$$

$$<f\,|\,g> = <f\,|\,\mathcal{I}\,|\,g> = <f\,|\left(\int |\,p><p\,|\,dp\right)|\,g> = \int <f\,|\,p><p\,|\,g>\;dp = \int \widehat{f}^*(p)\,\widehat{g}(p)\,dp\,.$$

$$\begin{aligned} <f\,|\,g> &= <f\,|\,\mathcal{I}\,|\,g> = <f\,|\left(\int |\,E><E\,|\,dE\right)|\,g> \\ &= \int <f\,|\,E><E\,|\,g>\;dE = \int \widetilde{f}^*(E)\,\widetilde{g}(E)\,dE\,. \end{aligned}$$

Postscript: See problem 41 for a comment concerning the validity of moving the bra and ket inside the integrand.

---

20. (a) Use the result of problem 19 to establish an expression of the normalization condition in position space.

(b) Normalize $f(x) = \mathrm{A}\,e^{-\mathrm{B}x^2}$.

This is a practical problem. First, you know the representation of $<f\,|\,g>$ in position space, so the what is the representation of $<f\,|\,f>$? Transition to a representation such as position space is often accompanied by a transition to calculus based arguments. You may find form 3.323.2 from Gradshteyn and Ryzhik useful. It is

$$\int_{-\infty}^{\infty} \exp\left(-p^2x^2 \pm qx\right) dx = \exp\left(\frac{q^2}{4p^2}\right)\frac{\sqrt{\pi}}{p}\,, \quad p > 0\,.$$

(a) The normalization condition is $<f\,|\,f> = 1 \quad \Rightarrow \quad \int f^*(x)\, f(x)\, dx = 1$ in position space.

(b) $$<f\,|\,f> = 1 \;\Rightarrow\; \int_{-\infty}^{\infty} \left(\mathrm{A}\, e^{-\mathrm{B}x^2}\right)^* \left(\mathrm{A}\, e^{-\mathrm{B}x^2}\right) dx = 1 \;\Rightarrow\; \left|\,\mathrm{A}\,\right|^2 \int_{-\infty}^{\infty} e^{-2\mathrm{B}x^2}\, dx = 1\,.$$

Using form 3.323.2 from Gradshteyn and Ryzhik, our integral has $p^2 = 2\mathrm{B}\,,\;\; q = 0\,,$ therefore

$$\left|\,\mathrm{A}\,\right|^2 e^0 \sqrt{\frac{\pi}{2\mathrm{B}}} = 1 \quad\Rightarrow\quad \left|\,\mathrm{A}\,\right|^2 \sqrt{\frac{\pi}{2\mathrm{B}}} = 1 \quad\Rightarrow\quad \mathrm{A} = \left(\frac{2\mathrm{B}}{\pi}\right)^{1/4} \quad\Rightarrow\quad f(x) = \left(\frac{2\mathrm{B}}{\pi}\right)^{1/4} e^{-\mathrm{B}x^2}\,.$$

---

21. (a) Show that

$\psi_n(x) = \cos\left(\frac{n\pi x}{l}\right)$ and $\psi_m(x) = \sin\left(\frac{m\pi x}{l}\right)$ are orthogonal on the interval $-\,l < x < l\,.$

(b) Orthonormalize $\psi_n(x)$ and $\psi_m(x)\,.$

There are many reasons for this problem at this point. Part (a) is an application of an inner product in position space. Cosines, sines, and orthogonality are intrinsic to the discussion of Fourier series that follows. Finally, we will see these wave functions again when discussing the infinite square well, meaning that they are realistic enough to be useful. A substantial portion of the reason that they are useful is the property of orthonormality addressed in part (b).

To show orthogonality for part (a), you must find that the inner product of each wavefunction with itself is non-zero, *i.e.*, $<\psi_n\,|\,\psi_n> \neq 0$ and $<\psi_m\,|\,\psi_m> \neq 0$. Then you must show that

$$<\psi_{n_i}\,|\,\psi_{n_j}> \;=\; <\psi_{m_i}\,|\,\psi_{m_j}> \;=\; <\psi_n\,|\,\psi_m> \;=\; 0.$$

Here are some indefinite integrals taken from the 30th edition of CRC Standard Mathematical Tables and Formulae that you should find useful.

$$\int \cos^2 ax\, dx = \frac{x}{2} + \frac{1}{4a}\sin 2ax\,, \qquad \int \sin^2 ax\, dx = \frac{x}{2} - \frac{1}{4a}\sin 2ax\,,$$

$$\int (\sin ax)(\cos ax)\, dx = \frac{1}{2a}\sin^2 ax\,,$$

and some others from Gradshteyn and Ryzhik,

$$\int (\cos ax)(\cos bx)\, dx = \frac{\sin(a-b)\,x}{2\,(a-b)} + \frac{\sin(a+b)\,x}{2\,(a+b)}\,, \quad a^2 \neq b^2\,,$$

$$\int (\sin ax)(\sin bx)\, dx = \frac{\sin(a-b)\,x}{2\,(a-b)} - \frac{\sin(a+b)\,x}{2\,(a+b)}\,, \quad a^2 \neq b^2\,,$$

$$\int (\sin ax)(\cos bx)\, dx = -\frac{\cos(a-b)\,x}{2\,(a-b)} - \frac{\cos(a+b)\,x}{2\,(a+b)}\,, \quad a^2 \neq b^2\,.$$

The appropriate limits on these integrals are $-l$ and l. Indices are used only to enumerate eigenstates, thus, all n and m are positive integers. Also remember that cosine is an even function and sine is an odd function.

$$\begin{aligned}
<\psi_n\,|\,\psi_n> &= \int_{-l}^{l}\left(\cos\left(\frac{n\pi x}{l}\right)\right)^*\left(\cos\left(\frac{n\pi x}{l}\right)\right)dx = \int_{-l}^{l}\cos^2\left(\frac{n\pi x}{l}\right)dx \\
&= \left[\frac{x}{2}+\frac{l}{4n\pi}\sin\,\frac{2n\pi x}{l}\right]_{-l}^{l} = \frac{l}{2}-\frac{-l}{2}+\frac{l}{4n\pi}\sin\,\frac{2n\pi l}{l}-\frac{l}{4n\pi}\sin\,\frac{-2n\pi l}{l} \\
&= l+\frac{l}{4n\pi}\Big(\sin\,2n\pi+\sin\,2n\pi\Big) = l+\frac{l}{2n\pi}\cancel{\sin}\,2n\pi = l \neq 0\,,
\end{aligned}$$

where the sine is struck because the sine of any integral multiple of π is zero. Similarly

$$\begin{aligned}
<\psi_m\,|\,\psi_m> &= \int_{-l}^{l}\left(\sin\left(\frac{m\pi x}{l}\right)\right)^*\left(\sin\left(\frac{m\pi x}{l}\right)\right)dx = \int_{-l}^{l}\sin^2\left(\frac{m\pi x}{l}\right)dx \\
&= \left[\frac{x}{2}-\frac{l}{4m\pi}\sin\,\frac{2m\pi x}{l}\right]_{-l}^{l} = \frac{l}{2}-\frac{-l}{2}-\frac{l}{4m\pi}\sin\,\frac{2m\pi l}{l}+\frac{l}{4m\pi}\sin\,\frac{-2m\pi l}{l} \\
&= l-\frac{l}{4m\pi}\Big(\sin\,2m\pi+\sin\,2m\pi\Big) = l-\frac{l}{2m\pi}\cancel{\sin}\,2m\pi = l \neq 0\,.
\end{aligned}$$

Conjugation is cosmetic when functions are real so is hereafter suppressed. Then

$$\begin{aligned}
<\psi_{n_i}\,|\,\psi_{n_j}> &= \int_{-l}^{l}\left(\cos\,\frac{n_i\pi x}{l}\right)\left(\cos\,\frac{n_j\pi x}{l}\right)dx \\
&= \left[\frac{\sin\left(\frac{n_i\pi x}{l}-\frac{n_j\pi x}{l}\right)}{2\left(\frac{n_i\pi}{l}-\frac{n_j\pi}{l}\right)}+\frac{\sin\left(\frac{n_i\pi x}{l}+\frac{n_j\pi x}{l}\right)}{2\left(\frac{n_i\pi}{l}+\frac{n_j\pi}{l}\right)}\right]_{-l}^{l}
\end{aligned}$$

$$= \frac{l}{2\pi}\frac{\sin\left(n_i-n_j\right)\pi}{\left(n_i-n_j\right)}-\frac{l}{2\pi}\frac{\sin\left(-n_i+n_j\right)\pi}{\left(n_i-n_j\right)}+\frac{l}{2\pi}\frac{\sin\left(n_i+n_j\right)\pi}{\left(n_i+n_j\right)}-\frac{l}{2\pi}\frac{\sin\left(-n_i-n_j\right)\pi}{\left(n_i+n_j\right)}\,.$$

Remember that $n_i \neq n_j$, because the case $n_i = n_j = n$ is covered by $<\psi_n\,|\,\psi_n>$. The arguments of all the sine functions are sums and differences of integers times π, so are integral multiples of π. The numerators of all terms are therefore zero, so $<\psi_{n_i}\,|\,\psi_{n_j}> = 0$. Next

$$\begin{aligned}
<\psi_{m_i}\,|\,\psi_{m_j}> &= \int_{-l}^{l}\left(\sin\,\frac{m_i\pi x}{l}\right)\left(\sin\,\frac{m_j\pi x}{l}\right)dx \\
&= \left[\frac{\sin\left(\frac{m_i\pi x}{l}-\frac{m_j\pi x}{l}\right)}{2\left(\frac{m_i\pi}{l}-\frac{m_j\pi}{l}\right)}-\frac{\sin\left(\frac{m_i\pi x}{l}+\frac{m_j\pi x}{l}\right)}{2\left(\frac{m_i\pi}{l}+\frac{m_j\pi}{l}\right)}\right]_{-l}^{l}
\end{aligned}$$

$$= \frac{l}{2\pi}\frac{\sin\left(m_i-m_j\right)\pi}{\left(m_i-m_j\right)}-\frac{l}{2\pi}\frac{\sin\left(-m_i+m_j\right)\pi}{\left(m_i-m_j\right)}-\frac{l}{2\pi}\frac{\sin\left(m_i+m_j\right)\pi}{\left(m_i+m_j\right)}+\frac{l}{2\pi}\frac{\sin\left(-m_i-m_j\right)\pi}{\left(m_i+m_j\right)}\,.$$

Again, $m_i \neq m_j$, because the case $m_i = m_j = m$ is addressed by $<\psi_m \,|\, \psi_m>$. Again, the arguments of all the sine functions are sums and differences of integers times π, so are integral multiples of π. The numerators of all terms are again zero, so $<\psi_{m_i} \,|\, \psi_{m_j}> = 0$. Then

$$
\begin{aligned}
<\psi_m \,|\, \psi_n> &= \int_{-l}^{l} \left(\sin \frac{m\pi x}{l}\right)\left(\cos \frac{n\pi x}{l}\right) dx \\
&= \left[-\frac{\cos\left(\frac{m\pi x}{l} - \frac{n\pi x}{l}\right)}{2\left(\frac{m\pi}{l} - \frac{n\pi}{l}\right)} - \frac{\cos\left(\frac{m\pi x}{l} + \frac{n\pi x}{l}\right)}{2\left(\frac{m\pi}{l} + \frac{n\pi}{l}\right)} \right]_{-l}^{l}
\end{aligned}
$$

$$
= -\frac{l}{2\pi}\frac{\cos(m-n)\pi}{(m-n)} + \frac{l}{2\pi}\frac{\cos(-m+n)\pi}{(m-n)} - \frac{l}{2\pi}\frac{\cos(m+n)\pi}{(m+n)} + \frac{l}{2\pi}\frac{\cos(-m-n)\pi}{(m+n)}.
$$

The cosine is an even function, so $\cos(-m+n)\pi = \cos(m-n)\pi$ and $\cos(-m-n)\pi = \cos(m+n)\pi$. Substituting in the last line,

$$
<\psi_m \,|\, \psi_n> = -\frac{l}{2\pi}\frac{\cos(m-n)\pi}{(m-n)} + \frac{l}{2\pi}\frac{\cos(m-n)\pi}{(m-n)} - \frac{l}{2\pi}\frac{\cos(m+n)\pi}{(m+n)} + \frac{l}{2\pi}\frac{\cos(m+n)\pi}{(m+n)},
$$

and the first and second terms are identical except they have opposite signs so sum to zero, similarly, the third and fourth terms sum to zero, so we conclude $<\psi_m \,|\, \psi_n> = 0, \quad m \neq n$. The qualification $m \neq n$ applies to the last calculation because $m = n$ is not covered by any previous case. Since both indices are equal for this calculation, let $m = n = k$

$$
<\psi_k \,|\, \psi_k> = \int_{-l}^{l} \left(\sin \frac{k\pi x}{l}\right)\left(\cos \frac{k\pi x}{l}\right) dx = \frac{l}{2k\pi} \sin^2 \frac{k\pi x}{l}\Bigg|_{-l}^{l}
$$

$$
= \frac{l}{2k\pi}\left(\sin^2 k\pi - \sin^2(-k\pi)\right) = \frac{l}{2k\pi}\left(\sin^2 k\pi - \sin^2 k\pi\right) = 0
$$

since $\sin^2$ is an even function. Also, the sine of an integral multiple of π is zero so the squares of integral multiples of π are zero. This exhausts all possibilities for indices that are positive integers. Therefore,

$$
\psi_n(x) = \cos\left(\frac{n\pi x}{l}\right) \quad \text{and} \quad \psi_m(x) = \sin\left(\frac{m\pi x}{l}\right) \quad \text{are orthogonal on the interval} \quad -l < x < l.
$$

(b) In the first two calculations we found $<\psi_n \,|\, \psi_n> = <\psi_m \,|\, \psi_m> = l$, so

$$
<\psi_n \,|\, \mathrm{A}^* \,\mathrm{A} \,|\, \psi_n> = \left|\mathrm{A}\right|^2 <\psi_n \,|\, \psi_n> = \left|\mathrm{A}\right|^2 l = 1 \quad \Rightarrow \quad \mathrm{A} = \frac{1}{\sqrt{l}}
$$

$$
\Rightarrow \quad \psi_n(x) = \frac{1}{\sqrt{l}}\cos\left(\frac{n\pi x}{l}\right) \quad \text{and} \quad \psi_m(x) = \frac{1}{\sqrt{l}}\sin\left(\frac{n\pi x}{l}\right)
$$

are orthonormal on the interval $-l < x < l$.

---

22. (a) Show that if indices of $n < 0$ where to apply, $a_{-n} = a_n$, and that $a_0 = \dfrac{1}{l}\displaystyle\int_{-l}^{l} f(x)\,dx$ for a basic Fourier series.

(b) Show that if indices of $n < 0$ where to apply, $b_{-n} = -b_n$, and that $b_0 = 0$ for a basic Fourier series.

(c) Express the c_n that are the coefficients of the exponential format of the Fourier series in terms of the basic Fourier coefficients a_n and b_n, and show that $c_0 = a_0/2$.

(d) Demonstrate the equivalence of the basic Fourier series and the exponential format of the Fourier series.

---

the same period? A **Fourier series** is a representation of a periodic function as a linear combination of all cosine and sine functions that have the same period. The periodic function $f(x) = f(x + 2l)$, has a period of $2l$, or repeats itself every $2l$. In a basic Fourier series,

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{l}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{l}\right),$$

where the coefficients are

$$a_n = \frac{1}{l}\int_{-l}^{l} f(x)\cos\left(\frac{n\pi x}{l}\right) dx, \qquad \text{and} \qquad b_n = \frac{1}{l}\int_{-l}^{l} f(x)\sin\left(\frac{n\pi x}{l}\right) dx,$$

given that the period is $2l$. The Fourier series in a exponential format for a function that is periodic over the length $2l$ is

$$f(x) = \sum_{-\infty}^{\infty} c_n\, e^{in\pi x/l}, \qquad \text{where} \qquad c_n = \frac{1}{2l}\int_{-l}^{l} f(x)\, e^{-in\pi x/l}\, dx.$$

Use the defining integrals for a_n, b_n, and c_n, for parts (a), (b), and (c). Cosines are even functions and sines are odd functions. Substituting these results into the summation that defines the exponential form yields the basic Fourier series required for part (d). It is wise to divide the summation into three parts that coincide with the three cases, namely: $n < 0$ implies a summation from $-\infty$ to -1, $n = 0$ is one term of the summation, and $n > 0$ implies a summation from 1 to ∞. The tricky part is switching the indices on the summation from $-\infty$ to -1. It may require some reflection, but realize that

$$\sum_{-\infty}^{-1} \frac{1}{2}(a_n - ib_n)\, e^{in\pi x/l} = \sum_{1}^{\infty} \frac{1}{2}(a_n + ib_n)\, e^{-in\pi x/l}.$$

(a)
$$a_{-n} = \frac{1}{l}\int_{-l}^{l} f(x)\cos\left(\frac{-n\pi x}{l}\right) dx = \frac{1}{l}\int_{-l}^{l} f(x)\cos\left(\frac{n\pi x}{l}\right) dx = a_n,$$

$$\text{and} \quad a_0 = \frac{1}{l}\int_{-l}^{l} f(x)\cos(0)\, dx = \frac{1}{l}\int_{-l}^{l} f(x)\, dx.$$

(b)
$$b_{-n} = \frac{1}{l}\int_{-l}^{l} f(x)\sin\left(\frac{-n\pi x}{l}\right) dx = -\frac{1}{l}\int_{-l}^{l} f(x)\sin\left(\frac{n\pi x}{l}\right) dx = -b_n,$$

$$\text{and} \quad b_0 = \frac{1}{l} \int_{-l}^{l} f(x) \sin(0)\, dx = 0\,.$$

(c)
$$\begin{aligned}
c_n &= \frac{1}{2l} \int_{-l}^{l} f(x)\, e^{-in\pi x/l}\, dx = \frac{1}{2l} \int_{-l}^{l} f(x) \left[\cos\left(\frac{n\pi x}{l}\right) - i \sin\left(\frac{n\pi x}{l}\right) \right] dx \\
&= \frac{1}{2l} \int_{-l}^{l} f(x) \cos\left(\frac{n\pi x}{l}\right) dx - \frac{i}{2l} \int_{-l}^{l} f(x) \sin\left(\frac{n\pi x}{l}\right) dx = \frac{1}{2}\left(a_n - ib_n\right).
\end{aligned}$$

When $n = 0$, the argument of the exponential is zero, so

$$c_0 = \frac{1}{2l} \int_{-l}^{l} f(x)\, e^0\, dx = \frac{1}{2l} \int_{-l}^{l} f(x)\, dx = \frac{1}{2}\, a_0\,.$$

(d) For convenience, let $\theta = \pi x / l\,.$ The summation is

$$\begin{aligned}
f(x) &= \sum_{-\infty}^{\infty} c_n\, e^{in\theta} = \sum_{-\infty}^{-1} \frac{1}{2}\left(a_n - ib_n\right) e^{in\theta} + \frac{1}{2}\, a_0 + \sum_{1}^{\infty} \frac{1}{2}\left(a_n - ib_n\right) e^{in\theta} \\
&= \frac{1}{2}\, a_0 + \sum_{1}^{\infty} \left(\frac{a_n}{2} + \frac{ib_n}{2} \right) e^{-in\theta} + \sum_{1}^{\infty} \left(\frac{a_n}{2} - \frac{ib_n}{2} \right) e^{in\theta} && (1) \\
&= \frac{1}{2}\, a_0 + \sum_{1}^{\infty} \left(\frac{a_n}{2}\, e^{-in\theta} + \frac{ib_n}{2}\, e^{-in\theta} + \frac{a_n}{2}\, e^{in\theta} - \frac{ib_n}{2}\, e^{in\theta} \right) \\
&= \frac{1}{2}\, a_0 + \sum_{1}^{\infty} \left(\frac{a_n}{2} \cos n\theta - \frac{a_n}{2}\, i \cancel{\sin n\theta} + \frac{ib_n}{2} \cancel{\cos n\theta} + \frac{b_n}{2} \sin n\theta \right. \\
&\qquad\qquad \left. + \frac{a_n}{2} \cos n\theta + \frac{a_n}{2}\, i \cancel{\sin n\theta} - \frac{ib_n}{2} \cancel{\cos n\theta} + \frac{b_n}{2} \sin n\theta \right) && (2) \\
&= \frac{1}{2}\, a_0 + \sum_{1}^{\infty} \left(a_n \cos n\theta + b_n \sin n\theta \right)
\end{aligned}$$

which is identical to the basic Fourier series upon the substitution of $\theta = \pi x / l\,.$ We have used the results of parts (a) and (b) in equation (1). The sign of the argument of the first exponential in equation (1) changes because we have changed the sign of the index of the summation. The signs of the fourth and eighth terms in the summation of equation (2) are correct because both include factors of $i \cdot i = -1$.

Postscript: The cosine and sine functions are orthogonal over a finite interval, say $-l < x < l\,,$ and can be made orthonormal on that interval as demonstrated in problem 21. Orthonormal vectors or orthonormal functions constitute a basis. Any function can be expressed in terms of a linear combination of the vectors, or in this case, the functions that constitute the basis.

The infinite square well, also known as a particle in a box, uses the orthogonal cosine and sine functions as a basis.

The exponential format of the Fourier series is often called **complex Fourier series**. We are interested in the complex Fourier series because it is the form that is most easily generalized to the **Fourier integral**.

23. Sketch the graphs of the following functions and their the Fourier transforms.

(a) $f(x) = \begin{cases} 1\,, & -a < x < a\,, \\ 0\,, & \text{elsewhere,} \end{cases}$ and (b) $f_1(x) = \begin{cases} 1/2\,, & -b < x < b\,, \\ 0\,, & \text{elsewhere,} \end{cases}$ where $b = 2a$.

Fourier integrals are generalizations of Fourier series. A **Fourier integral** is a representation of a non-periodic function that may be regarded as the limit of the Fourier series as the period approaches infinity. Two **Fourier transforms** compose the Fourier integral. Fourier transforms are the objects that we want to develop further for quantum mechanical calculation. We refer the reader to Byron and Fuller[6], or your favorite text on mathematical physics for greater depth. In the limit of $l \to \infty$, the complex Fourier series becomes

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

The function $f(x)$ is the Fourier transform of $g(k)$, and $g(k)$ is the Fourier transform of $f(x)$. Substitution of one transform into the other expressing the combination as one relation, *i.e.*,

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x')\, e^{-ikx'}\, dx' \right) e^{ikx}\, dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} f(x')\, e^{ik(x-x')}\, dx' \quad \text{is the Fourier integral.} \end{aligned}$$

The functions given for parts (a) and (b) are constants for specified intervals and zero elsewhere which means that you can find $g(k)$ and $g_1(k)$ by integrating between $-a$ and a, and $-b$ and b instead of using infinite limits. Express your answer for part (b) in terms of the constant a and look for relations between the graphs of parts (a) and (b). This problem uses constant functions which are among the easiest functions to integrate both as an appropriate place to start and to highlight relations between parts (a) and (b) for which we provide comment in the postscript.

(a)

$$\begin{aligned} g(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} f(x)\, e^{-ikx}\, dx \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{(-ik)}\, e^{-ikx} \Big|_{-a}^{a} = \frac{1}{\sqrt{2\pi}} \frac{e^{-ika} - e^{ika}}{-ik} \\ &= \frac{2}{\sqrt{2\pi}\, k} \frac{e^{-ika} - e^{ika}}{-2i} = \frac{2}{\sqrt{2\pi}\, k} \frac{e^{ika} - e^{-ika}}{2i} = \frac{2}{\sqrt{2\pi}} \frac{\sin(ka)}{k}\,. \end{aligned}$$

The graphs of $f(x)$ and $g(k)$, are

[6] Byron and Fuller, *Mathematics of Classical and Quantum Physics* (Dover Publications, Inc., New York, 1970), pp. 239–253, 566–570.

Graph of $f(x)$ Graph of $g(k)$

(b) The Fourier transform of the given function where $b = 2a$, double the width and half the height, so that the area under the curve is equal for both $f(x)$ and $f_1(x)$, is

$$g_1(k) = \frac{1}{\sqrt{2\pi}}\frac{\sin(kb)}{k} = \frac{1}{\sqrt{2\pi}}\frac{\sin(2ka)}{k}.$$

Graph of $f_1(x)$ Graph of $g_1(k)$

Postscript: The graph of $f_1(x)$ is broader than the graph of $f(x)$, but the graph of $g_1(k)$ is sharper than the graph of $g(k)$. If one of the Fourier transforms is distributed and broad, the other will be localized and sharp. This phenomena is a precursor of the Heisenberg uncertainty relations.

While $f_i(x)$ has a limited domain, the domain of $g_i(k)$ is infinite. This is another general feature of Fourier transforms.

---

24. Derive the Fourier transforms from the complex Fourier series.

The basic Fourier series is useful for introductions and many applications, but the complex Fourier series is the form from which Fourier transforms are most easily attained. This is a challenging problem but provides an opportunity to deepen your understanding of Fourier transforms.

Start with the notation

$$F(y) = \sum_{-\infty}^{\infty} c_n\, e^{in\pi y/l}, \qquad \text{where} \qquad c_n = \frac{1}{2l}\int_{-l}^{l} F(y)\, e^{-in\pi y/l}\, dy.$$

Substitute the integral form of c_n into the summation. Symmetrize the factor of $1/2l$. This means to rearrange the summation so a factor of $1/\sqrt{2l}$ is in both the $F(y)$ and the c_n. You now have $F(y)$ and c_n that are different than originally defined. Change variables by letting $x = \sqrt{\frac{\pi}{l}}\, y$. Also, let $k_n = n\sqrt{\frac{\pi}{l}}$. You can now find that $\Delta k_n = \sqrt{\frac{\pi}{l}}$ and $\frac{n\pi y}{l} = k_n x$. Redefine $F\left(\sqrt{\frac{l}{\pi}}\, x\right) = f(x)$. Introduce a new symbol for the coefficients that indicates the index k_n replaces n, so $c_n \to g_{k_n}$. Finally, let $l \to \infty$, so that the Fourier transforms

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(k)\, e^{ikx}\, dk \quad \text{and} \quad g(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx \quad \text{are exposed.}$$

$$F(y) = \sum_{n=-\infty}^{\infty} \left(\frac{1}{2l} \int_{-l}^{l} F(y)\, e^{-in\pi y/l}\, dy \right) e^{in\pi y/l} = \sum_{n=-\infty}^{\infty} \left(\frac{1}{\sqrt{2l}} \int_{-l}^{l} F(y)\, e^{-in\pi y/l}\, dy \right) \frac{1}{\sqrt{2l}}\, e^{in\pi y/l}.$$

We can now identify a symmetrized function and coefficient as

$$F(y) = \sum_{n=-\infty}^{\infty} c_n \frac{1}{\sqrt{2l}}\, e^{in\pi y/l} \quad \text{and} \quad c_n = \frac{1}{\sqrt{2l}} \int_{-l}^{l} F(y)\, e^{-in\pi y/l}\, dy\,. \tag{1}$$

We are going to change variables to $\sqrt{\frac{\pi}{l}}\, y = x \quad \Rightarrow \quad y = \sqrt{\frac{l}{\pi}}\, x \quad \Rightarrow \quad dy = \sqrt{\frac{l}{\pi}}\, dx$. We are also going to use $k_n = n\sqrt{\frac{\pi}{l}} \quad \Rightarrow \quad \Delta k_n = k_{n+1} - k_n = (n+1)\sqrt{\frac{\pi}{l}} - n\sqrt{\frac{\pi}{l}} = \sqrt{\frac{\pi}{l}}$. These substitutions allow us to rewrite a portion of the argument of the exponentials, $\frac{n\pi y}{l} = n\sqrt{\frac{\pi}{l}}\sqrt{\frac{\pi}{l}}\, y = k_n x$. The difference in the new index is $\Delta k_n = \sqrt{\frac{\pi}{l}} \quad \Rightarrow \quad \frac{\Delta k_n}{\sqrt{2\pi}} = \frac{1}{\sqrt{2l}}$. The limits of the integral become $y = l \quad \Rightarrow \quad x = \sqrt{\frac{\pi}{l}}\, l = \sqrt{\pi l}$, and $y = -l \quad \Rightarrow \quad x = -\sqrt{\pi l}$. We finally introduce a new symbol for the coefficients that indicates the index is now k_n vice n so $c_n \to g_{k_n}$. Using these substitutions and developments in equation (1),

$$F(y) = F\left(\sqrt{\frac{l}{\pi}}\, x\right) = \sum_{k_n=-\infty}^{\infty} g_{k_n} \frac{1}{\sqrt{2\pi}}\, e^{ik_n x}\, \Delta k_n = \frac{1}{\sqrt{2\pi}} \sum_{k_n=-\infty}^{\infty} g_{k_n}\, e^{ik_n x}\, \Delta k_n$$

$$\text{and} \qquad g_{k_n} = \frac{1}{\sqrt{2l}} \int_{-\sqrt{\pi l}}^{\sqrt{\pi l}} F\left(\sqrt{\frac{l}{\pi}}\, x\right) e^{-ik_n x} \sqrt{\frac{l}{\pi}}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{\pi l}}^{\sqrt{\pi l}} F\left(\sqrt{\frac{l}{\pi}}\, x\right) e^{-ik_n x}\, dx$$

Redefine $F\left(\sqrt{\frac{l}{\pi}}\, x\right) = f(x)$. Now, let the period $l \to \infty \quad \Rightarrow \quad \Delta k_n = \sqrt{\frac{\pi}{l}} \to 0$. The difference between successive $k_n \to 0$ so that k_n assumes all real values and thus becomes a continuous variable. The summation over discrete k_n becomes an integral over the continuous variable k, the coefficient g_{k_n} becomes a continuous function $g(k)$, the difference Δk_n becomes the differential dk, and

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

Postscript: The factor of $1/2\pi$ must be in the Fourier integral but it does not need to be symmetrized in the Fourier transforms. Some authors and tables place the entire factor $1/2\pi$ with one integral so that there is no coefficient in front of the other integral. Also, some authors and tables will change the signs of the exponentials from what we have presented. Asymmetric treatment of the factor of $1/2\pi$ and/or use of a different sign convention result in slightly different functional forms for resulting Fourier transforms.

25. Find the momentum space wavefunction corresponding to a position space wavefunction that is Gaussian.

Gaussian wavefunctions are foundational as you will find in chapter 3 and elsewhere. But first, what does it mean to change a wavefunction in position space to a wavefunction in momentum space? It means find the description of a system using an unknown function of the continuous variable momentum starting with the description of the system that is a known function of the continuous variable position. It is a change of basis in the realm of continuous variables. Quantum mechanical Fourier transforms are the means to change from position space to momentum space and vice versa for continuous variables.

The de Broglie relation associates wavelength or wavenumber, and momentum,

$$\lambda = \frac{h}{p} \quad \Rightarrow \quad \frac{2\pi}{k} = \frac{h}{p} \quad \Rightarrow \quad k = \frac{p}{\hbar}\,.$$

Substitute $p/\hbar$ for k in both Fourier transforms. This means the differential $dk = dp/\hbar$ in the integral for $f(x)$. Symmetrize the $1/\hbar$ from this differential by placing a factor $1/\sqrt{\hbar}$ in both integrals, and the result is the quantum mechanical form of the Fourier transforms,

$$f(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \widehat{g}(p)\, e^{ipx/\hbar}\, dp, \qquad \widehat{g}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} f(x)\, e^{-ipx/\hbar}\, dx\,.$$

The usual use of these relations is to transform a wave function in position space to momentum space (or vice versa), so will be applied

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \widehat{\psi}(p)\, e^{ipx/\hbar}\, dp, \qquad \widehat{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x)\, e^{-ipx/\hbar}\, dx\,.$$

There are a number of subtleties associated with the quantum mechanical analog of the Fourier transforms, so our heuristic argument should be regarded as nothing more than a useful mnemonic.

Use $\psi(x) = A\, e^{-bx^2}$, $\quad b > 0$, as your position space wavefunction that is Gaussian. Form 3.323.2 from Gradshteyn and Ryzhik,

$$\int_{-\infty}^{\infty} e^{-\alpha^2 x^2 - \beta x}\, dx = \frac{\sqrt{\pi}}{\alpha}\, e^{\beta^2/4\alpha^2},$$

should be useful. Do you recognize the functional form of the momentum space wavefunction?

$$\begin{aligned}\widehat{\psi}(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x)\, e^{-ipx/\hbar}\, dx = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} A\, e^{-bx^2}\, e^{-ipx/\hbar}\, dx \\ &= \frac{A}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-bx^2 - ipx/\hbar}\, dx\,.\end{aligned}$$

Using $\alpha = \sqrt{b}$, and $\beta = ip/\hbar$, our integral is

$$\widehat{\psi}(p) = \frac{A}{\sqrt{2\pi\hbar}}\, \frac{\sqrt{\pi}}{\sqrt{b}}\, \exp\left(\frac{(ip/\hbar)^2}{4\,(\sqrt{b}\,)^2}\right) = \frac{A}{\sqrt{2\hbar b}}\, e^{-p^2/4b\hbar^2}.$$

Postscript: Notice that $\widehat{\psi}(p)$ is also an exponential function with a negative argument that is a constant times the independent variable squared. In fact, $\widehat{\psi}(p)$ is also a Gaussian wavefunction. The Fourier transform of a Gaussian function is another Gaussian function.

---

21. (a) Show that

$$\psi_n(x) = \cos\left(\frac{n\pi x}{l}\right) \quad \text{and} \quad \psi_m(x) = \sin\left(\frac{m\pi x}{l}\right) \quad \text{are orthogonal on the interval} \quad -l < x < l\,.$$

(b) Orthonormalize $\psi_n(x)$ and $\psi_m(x)$.

There are many reasons for this problem at this point. Part (a) is an application of an inner product in position space. Cosines, sines, and orthogonality are intrinsic to the discussion of Fourier series that follows. Finally, we will see these wave functions again when discussing the infinite square well, meaning that they are realistic enough to be useful. A substantial portion of the reason that they are useful is the property of orthonormality addressed in part (b).

To show orthogonality for part (a), you must find that the inner product of each wavefunction with itself is non-zero, *i.e.*, $<\psi_n\,|\,\psi_n> \neq 0$ and $<\psi_m\,|\,\psi_m> \neq 0$. Then you must show that

$$<\psi_{n_i}\,|\,\psi_{n_j}> \;=\; <\psi_{m_i}\,|\,\psi_{m_j}> \;=\; <\psi_n\,|\,\psi_m> = 0.$$

Here are some indefinite integrals taken from the 30th edition of CRC Standard Mathematical Tables and Formulae that you should find useful.

$$\int \cos^2 ax\,dx = \frac{x}{2} + \frac{1}{4a}\sin 2ax\,, \qquad \int \sin^2 ax\,dx = \frac{x}{2} - \frac{1}{4a}\sin 2ax\,,$$

$$\int (\sin ax)(\cos ax)\,dx = \frac{1}{2a}\sin^2 ax\,,$$

and some others from Gradshteyn and Ryzhik,

$$\int (\cos ax)(\cos bx)\,dx = \frac{\sin(a-b)\,x}{2\,(a-b)} + \frac{\sin(a+b)\,x}{2\,(a+b)}\,, \quad a^2 \neq b^2\,,$$

$$\int (\sin ax)(\sin bx)\,dx = \frac{\sin(a-b)\,x}{2\,(a-b)} - \frac{\sin(a+b)\,x}{2\,(a+b)}\,, \quad a^2 \neq b^2\,,$$

$$\int (\sin ax)(\cos bx)\,dx = -\frac{\cos(a-b)\,x}{2\,(a-b)} - \frac{\cos(a+b)\,x}{2\,(a+b)}\,, \quad a^2 \neq b^2\,.$$

The appropriate limits on these integrals are $-l$ and l. Indices are used only to enumerate eigenstates, thus, all n and m are positive integers. Also remember that cosine is an even function and sine is an odd function.

$$\begin{aligned}<\psi_n\,|\,\psi_n> &= \int_{-l}^{l}\left(\cos\left(\frac{n\pi x}{l}\right)\right)^*\left(\cos\left(\frac{n\pi x}{l}\right)\right)dx = \int_{-l}^{l}\cos^2\left(\frac{n\pi x}{l}\right)dx\\
&= \left[\frac{x}{2}+\frac{l}{4n\pi}\sin\,\frac{2n\pi x}{l}\right]_{-l}^{l} = \frac{l}{2}-\frac{-l}{2}+\frac{l}{4n\pi}\sin\,\frac{2n\pi l}{l}-\frac{l}{4n\pi}\sin\,\frac{-2n\pi l}{l}\\
&= l+\frac{l}{4n\pi}\Big(\sin\,2n\pi+\sin\,2n\pi\Big) = l+\frac{l}{2n\pi}\cancel{\sin}\,2n\pi = l\neq 0\,,\end{aligned}$$

where the sine is struck because the sine of any integral multiple of π is zero. Similarly

$$\begin{aligned}<\psi_m\,|\,\psi_m> &= \int_{-l}^{l}\left(\sin\left(\frac{m\pi x}{l}\right)\right)^*\left(\sin\left(\frac{m\pi x}{l}\right)\right)dx = \int_{-l}^{l}\sin^2\left(\frac{m\pi x}{l}\right)dx\\
&= \left[\frac{x}{2}-\frac{l}{4m\pi}\sin\,\frac{2m\pi x}{l}\right]_{-l}^{l} = \frac{l}{2}-\frac{-l}{2}-\frac{l}{4m\pi}\sin\,\frac{2m\pi l}{l}+\frac{l}{4m\pi}\sin\,\frac{-2m\pi l}{l}\\
&= l-\frac{l}{4m\pi}\Big(\sin\,2m\pi+\sin\,2m\pi\Big) = l-\frac{l}{2m\pi}\cancel{\sin}\,2m\pi = l\neq 0\,.\end{aligned}$$

Conjugation is cosmetic when functions are real so is hereafter suppressed. Then

$$\begin{aligned}<\psi_{n_i}\,|\,\psi_{n_j}> &= \int_{-l}^{l}\left(\cos\,\frac{n_i\pi x}{l}\right)\left(\cos\,\frac{n_j\pi x}{l}\right)dx\\
&= \left[\frac{\sin\left(\dfrac{n_i\pi x}{l}-\dfrac{n_j\pi x}{l}\right)}{2\left(\dfrac{n_i\pi}{l}-\dfrac{n_j\pi}{l}\right)}+\frac{\sin\left(\dfrac{n_i\pi x}{l}+\dfrac{n_j\pi x}{l}\right)}{2\left(\dfrac{n_i\pi}{l}+\dfrac{n_j\pi}{l}\right)}\right]_{-l}^{l}\end{aligned}$$

$$= \frac{l}{2\pi}\frac{\sin\,(n_i-n_j)\,\pi}{(n_i-n_j)}-\frac{l}{2\pi}\frac{\sin\,(-n_i+n_j)\,\pi}{(n_i-n_j)}+\frac{l}{2\pi}\frac{\sin\,(n_i+n_j)\,\pi}{(n_i+n_j)}-\frac{l}{2\pi}\frac{\sin\,(-n_i-n_j)\,\pi}{(n_i+n_j)}\,.$$

Remember that $n_i\neq n_j$, because the case $n_i=n_j=n$ is covered by $<\psi_n\,|\,\psi_n>$. The arguments of all the sine functions are sums and differences of integers times π, so are integral multiples of π. The numerators of all terms are therefore zero, so $<\psi_{n_i}\,|\,\psi_{n_j}>=0$. Next

$$\begin{aligned}<\psi_{m_i}\,|\,\psi_{m_j}> &= \int_{-l}^{l}\left(\sin\,\frac{m_i\pi x}{l}\right)\left(\sin\,\frac{m_j\pi x}{l}\right)dx\\
&= \left[\frac{\sin\left(\dfrac{m_i\pi x}{l}-\dfrac{m_j\pi x}{l}\right)}{2\left(\dfrac{m_i\pi}{l}-\dfrac{m_j\pi}{l}\right)}-\frac{\sin\left(\dfrac{m_i\pi x}{l}+\dfrac{m_j\pi x}{l}\right)}{2\left(\dfrac{m_i\pi}{l}+\dfrac{m_j\pi}{l}\right)}\right]_{-l}^{l}\end{aligned}$$

$$= \frac{l}{2\pi}\frac{\sin\,(m_i-m_j)\,\pi}{(m_i-m_j)}-\frac{l}{2\pi}\frac{\sin\,(-m_i+m_j)\,\pi}{(m_i-m_j)}-\frac{l}{2\pi}\frac{\sin\,(m_i+m_j)\,\pi}{(m_i+m_j)}+\frac{l}{2\pi}\frac{\sin\,(-m_i-m_j)\,\pi}{(m_i+m_j)}\,.$$

Again, $m_i\neq m_j$, because the case $m_i=m_j=m$ is addressed by $<\psi_m\,|\,\psi_m>$. Again, the arguments of all the sine functions are sums and differences of integers times π, so are integral multiples of π. The numerators of all terms are again zero, so $<\psi_{m_i} \,|\, \psi_{m_j}> = 0$. Then

$$
\begin{aligned}
<\psi_m \,|\, \psi_n> &= \int_{-l}^{l} \left(\sin \frac{m\pi x}{l}\right)\left(\cos \frac{n\pi x}{l}\right) dx \\
&= \left[-\frac{\cos\left(\frac{m\pi x}{l} - \frac{n\pi x}{l}\right)}{2\left(\frac{m\pi}{l} - \frac{n\pi}{l}\right)} - \frac{\cos\left(\frac{m\pi x}{l} + \frac{n\pi x}{l}\right)}{2\left(\frac{m\pi}{l} + \frac{n\pi}{l}\right)} \right]_{-l}^{l}
\end{aligned}
$$

$$
= -\frac{l}{2\pi}\frac{\cos(m-n)\,\pi}{(m-n)} + \frac{l}{2\pi}\frac{\cos(-m+n)\,\pi}{(m-n)} - \frac{l}{2\pi}\frac{\cos(m+n)\,\pi}{(m+n)} + \frac{l}{2\pi}\frac{\cos(-m-n)\,\pi}{(m+n)}\,.
$$

The cosine is an even function, so $\cos(-m+n)\,\pi = \cos(m-n)\,\pi$ and $\cos(-m-n)\,\pi = \cos(m+n)\,\pi$. Substituting in the last line,

$$
<\psi_m \,|\, \psi_n> = -\frac{l}{2\pi}\frac{\cos(m-n)\,\pi}{(m-n)} + \frac{l}{2\pi}\frac{\cos(m-n)\,\pi}{(m-n)} - \frac{l}{2\pi}\frac{\cos(m+n)\,\pi}{(m+n)} + \frac{l}{2\pi}\frac{\cos(m+n)\,\pi}{(m+n)}\,,
$$

and the first and second terms are identical except they have opposite signs so sum to zero, similarly, the third and fourth terms sum to zero, so we conclude $<\psi_m \,|\, \psi_n> = 0, \quad m \neq n$. The qualification $m \neq n$ applies to the last calculation because $m = n$ is not covered by any previous case. Since both indices are equal for this calculation, let $m = n = k$

$$
<\psi_k \,|\, \psi_k> = \int_{-l}^{l} \left(\sin \frac{k\pi x}{l}\right)\left(\cos \frac{k\pi x}{l}\right) dx = \frac{l}{2k\pi}\,\sin^2 \frac{k\pi x}{l}\bigg|_{-l}^{l}
$$

$$
= \frac{l}{2k\pi}\left(\sin^2 k\pi - \sin^2(-k\pi)\right) = \frac{l}{2k\pi}\left(\sin^2 k\pi - \sin^2 k\pi\right) = 0
$$

since $\sin^2$ is an even function. Also, the sine of an integral multiple of π is zero so the squares of integral multiples of π are zero. This exhausts all possibilities for indices that are positive integers. Therefore,

$$
\psi_n(x) = \cos\left(\frac{n\pi x}{l}\right) \quad \text{and} \quad \psi_m(x) = \sin\left(\frac{m\pi x}{l}\right) \quad \text{are orthogonal on the interval} \quad -l < x < l.
$$

(b) In the first two calculations we found $<\psi_n \,|\, \psi_n> = <\psi_m \,|\, \psi_m> = l$, so

$$
<\psi_n \,|\, \mathrm{A}^* \,\mathrm{A} \,|\, \psi_n> = \left|\,\mathrm{A}\,\right|^2 <\psi_n \,|\, \psi_n> = \left|\,\mathrm{A}\,\right|^2 l = 1 \quad \Rightarrow \quad \mathrm{A} = \frac{1}{\sqrt{l}}
$$

$$
\Rightarrow \quad \psi_n(x) = \frac{1}{\sqrt{l}}\cos\left(\frac{n\pi x}{l}\right) \quad \text{and} \quad \psi_m(x) = \frac{1}{\sqrt{l}}\sin\left(\frac{n\pi x}{l}\right)
$$

are orthonormal on the interval $-l < x < l$.

---

22. (a) Show that if indices of $n < 0$ where to apply, $a_{-n} = a_n$, and that $a_0 = \dfrac{1}{l}\displaystyle\int_{-l}^{l} f(x)\,dx$ for a basic Fourier series.

(b) Show that if indices of $n < 0$ where to apply, $b_{-n} = -b_n$, and that $b_0 = 0$ for a basic Fourier series.

(c) Express the c_n that are the coefficients of the exponential format of the Fourier series in terms of the basic Fourier coefficients a_n and b_n, and show that $c_0 = a_0/2$.

(d) Demonstrate the equivalence of the basic Fourier series and the exponential format of the Fourier series.

---

the same period? A **Fourier series** is a representation of a periodic function as a linear combination of all cosine and sine functions that have the same period. The periodic function $f(x) = f(x + 2l)$, has a period of $2l$, or repeats itself every $2l$. In a basic Fourier series,

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{l}\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{l}\right),$$

where the coefficients are

$$a_n = \frac{1}{l}\int_{-l}^{l} f(x) \cos\left(\frac{n\pi x}{l}\right) dx, \qquad \text{and} \qquad b_n = \frac{1}{l}\int_{-l}^{l} f(x) \sin\left(\frac{n\pi x}{l}\right) dx,$$

given that the period is $2l$. The Fourier series in a exponential format for a function that is periodic over the length $2l$ is

$$f(x) = \sum_{-\infty}^{\infty} c_n\, e^{in\pi x/l}, \qquad \text{where} \qquad c_n = \frac{1}{2l}\int_{-l}^{l} f(x)\, e^{-in\pi x/l}\, dx.$$

Use the defining integrals for a_n, b_n, and c_n, for parts (a), (b), and (c). Cosines are even functions and sines are odd functions. Substituting these results into the summation that defines the exponential form yields the basic Fourier series required for part (d). It is wise to divide the summation into three parts that coincide with the three cases, namely: $n < 0$ implies a summation from $-\infty$ to -1, $n = 0$ is one term of the summation, and $n > 0$ implies a summation from 1 to ∞. The tricky part is switching the indices on the summation from $-\infty$ to -1. It may require some reflection, but realize that

$$\sum_{-\infty}^{-1} \frac{1}{2}(a_n - ib_n)\, e^{in\pi x/l} = \sum_{1}^{\infty} \frac{1}{2}(a_n + ib_n)\, e^{-in\pi x/l}.$$

(a) $$a_{-n} = \frac{1}{l}\int_{-l}^{l} f(x) \cos\left(\frac{-n\pi x}{l}\right) dx = \frac{1}{l}\int_{-l}^{l} f(x) \cos\left(\frac{n\pi x}{l}\right) dx = a_n,$$

$$\text{and} \quad a_0 = \frac{1}{l}\int_{-l}^{l} f(x) \cos(0)\, dx = \frac{1}{l}\int_{-l}^{l} f(x)\, dx.$$

(b) $$b_{-n} = \frac{1}{l}\int_{-l}^{l} f(x) \sin\left(\frac{-n\pi x}{l}\right) dx = -\frac{1}{l}\int_{-l}^{l} f(x) \sin\left(\frac{n\pi x}{l}\right) dx = -b_n,$$

$$\text{and} \quad b_0 = \frac{1}{l} \int_{-l}^{l} f(x) \sin(0)\, dx = 0\,.$$

(c)
$$c_n = \frac{1}{2l} \int_{-l}^{l} f(x)\, e^{-in\pi x/l}\, dx = \frac{1}{2l} \int_{-l}^{l} f(x) \left[\cos\left(\frac{n\pi x}{l}\right) - i \sin\left(\frac{n\pi x}{l}\right) \right] dx$$
$$= \frac{1}{2l} \int_{-l}^{l} f(x) \cos\left(\frac{n\pi x}{l}\right) dx - \frac{i}{2l} \int_{-l}^{l} f(x) \sin\left(\frac{n\pi x}{l}\right) dx = \frac{1}{2} (a_n - ib_n)\,.$$

When $n = 0$, the argument of the exponential is zero, so

$$c_0 = \frac{1}{2l} \int_{-l}^{l} f(x)\, e^0\, dx = \frac{1}{2l} \int_{-l}^{l} f(x)\, dx = \frac{1}{2}\, a_0\,.$$

(d) For convenience, let $\theta = \pi x/l$. The summation is

$$\begin{aligned}
f(x) &= \sum_{-\infty}^{\infty} c_n\, e^{in\theta} = \sum_{-\infty}^{-1} \frac{1}{2} (a_n - ib_n)\, e^{in\theta} + \frac{1}{2}\, a_0 + \sum_{1}^{\infty} \frac{1}{2} (a_n - ib_n)\, e^{in\theta} \\
&= \frac{1}{2}\, a_0 + \sum_{1}^{\infty} \left(\frac{a_n}{2} + \frac{ib_n}{2} \right) e^{-in\theta} + \sum_{1}^{\infty} \left(\frac{a_n}{2} - \frac{ib_n}{2} \right) e^{in\theta} && (1) \\
&= \frac{1}{2}\, a_0 + \sum_{1}^{\infty} \left(\frac{a_n}{2}\, e^{-in\theta} + \frac{ib_n}{2}\, e^{-in\theta} + \frac{a_n}{2}\, e^{in\theta} - \frac{ib_n}{2}\, e^{in\theta} \right) \\
&= \frac{1}{2}\, a_0 + \sum_{1}^{\infty} \Big(\frac{a_n}{2} \cos n\theta - \cancel{\frac{a_n}{2}\, i \sin n\theta} + \cancel{\frac{ib_n}{2} \cos n\theta} + \frac{b_n}{2} \sin n\theta \\
&\qquad\qquad + \frac{a_n}{2} \cos n\theta + \cancel{\frac{a_n}{2}\, i \sin n\theta} - \cancel{\frac{ib_n}{2} \cos n\theta} + \frac{b_n}{2} \sin n\theta \Big) && (2) \\
&= \frac{1}{2}\, a_0 + \sum_{1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta)
\end{aligned}$$

which is identical to the basic Fourier series upon the substitution of $\theta = \pi x/l$. We have used the results of parts (a) and (b) in equation (1). The sign of the argument of the first exponential in equation (1) changes because we have changed the sign of the index of the summation. The signs of the fourth and eighth terms in the summation of equation (2) are correct because both include factors of $i \cdot i = -1$.

Postscript: The cosine and sine functions are orthogonal over a finite interval, say $-l < x < l$, and can be made orthonormal on that interval as demonstrated in problem 21. Orthonormal vectors or orthonormal functions constitute a basis. Any function can be expressed in terms of a linear combination of the vectors, or in this case, the functions that constitute the basis.

The infinite square well, also known as a particle in a box, uses the orthogonal cosine and sine functions as a basis.

The exponential format of the Fourier series is often called **complex Fourier series**. We are interested in the complex Fourier series because it is the form that is most easily generalized to the **Fourier integral**.

23. Sketch the graphs of the following functions and their the Fourier transforms.

(a) $f(x) = \begin{cases} 1\,, & -a < x < a\,, \\ 0\,, & \text{elsewhere,} \end{cases}$ and (b) $f_1(x) = \begin{cases} 1/2\,, & -b < x < b\,, \\ 0\,, & \text{elsewhere,} \end{cases}$ where $b = 2a$.

Fourier integrals are generalizations of Fourier series. A **Fourier integral** is a representation of a non-periodic function that may be regarded as the limit of the Fourier series as the period approaches infinity. Two **Fourier transforms** compose the Fourier integral. Fourier transforms are the objects that we want to develop further for quantum mechanical calculation. We refer the reader to Byron and Fuller[6], or your favorite text on mathematical physics for greater depth. In the limit of $l \to \infty$, the complex Fourier series becomes

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

The function $f(x)$ is the Fourier transform of $g(k)$, and $g(k)$ is the Fourier transform of $f(x)$. Substitution of one transform into the other expressing the combination as one relation, *i.e.*,

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x')\, e^{-ikx'}\, dx' \right) e^{ikx}\, dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} f(x')\, e^{ik(x-x')}\, dx' \quad \text{is the Fourier integral.} \end{aligned}$$

The functions given for parts (a) and (b) are constants for specified intervals and zero elsewhere which means that you can find $g(k)$ and $g_1(k)$ by integrating between $-a$ and a, and $-b$ and b instead of using infinite limits. Express your answer for part (b) in terms of the constant a and look for relations between the graphs of parts (a) and (b). This problem uses constant functions which are among the easiest functions to integrate both as an appropriate place to start and to highlight relations between parts (a) and (b) for which we provide comment in the postscript.

(a)

$$\begin{aligned} g(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} f(x)\, e^{-ikx}\, dx \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{(-ik)}\, e^{-ikx} \Big|_{-a}^{a} = \frac{1}{\sqrt{2\pi}} \frac{e^{-ika} - e^{ika}}{-ik} \\ &= \frac{2}{\sqrt{2\pi}\, k} \frac{e^{-ika} - e^{ika}}{-2i} = \frac{2}{\sqrt{2\pi}\, k} \frac{e^{ika} - e^{-ika}}{2i} = \frac{2}{\sqrt{2\pi}} \frac{\sin(ka)}{k}\,. \end{aligned}$$

The graphs of $f(x)$ and $g(k)$, are

[6] Byron and Fuller, *Mathematics of Classical and Quantum Physics* (Dover Publications, Inc., New York, 1970), pp. 239–253, 566–570.

Graph of $f(x)$ Graph of $g(k)$

(b) The Fourier transform of the given function where $b = 2a$, double the width and half the height, so that the area under the curve is equal for both $f(x)$ and $f_1(x)$, is

$$g_1(k) = \frac{1}{\sqrt{2\pi}} \frac{\sin(kb)}{k} = \frac{1}{\sqrt{2\pi}} \frac{\sin(2ka)}{k}.$$

Graph of $f_1(x)$ Graph of $g_1(k)$

Postscript: The graph of $f_1(x)$ is broader than the graph of $f(x)$, but the graph of $g_1(k)$ is sharper than the graph of $g(k)$. If one of the Fourier transforms is distributed and broad, the other will be localized and sharp. This phenomena is a precursor of the Heisenberg uncertainty relations.

While $f_i(x)$ has a limited domain, the domain of $g_i(k)$ is infinite. This is another general feature of Fourier transforms.

---

24. Derive the Fourier transforms from the complex Fourier series.

The basic Fourier series is useful for introductions and many applications, but the complex Fourier series is the form from which Fourier transforms are most easily attained. This is a challenging problem but provides an opportunity to deepen your understanding of Fourier transforms.

Start with the notation

$$F(y) = \sum_{-\infty}^{\infty} c_n\, e^{in\pi y/l}, \qquad \text{where} \qquad c_n = \frac{1}{2l} \int_{-l}^{l} F(y)\, e^{-in\pi y/l}\, dy.$$

Substitute the integral form of c_n into the summation. Symmetrize the factor of $1/2l$. This means to rearrange the summation so a factor of $1/\sqrt{2l}$ is in both the $F(y)$ and the c_n. You now have $F(y)$ and c_n that are different than originally defined. Change variables by letting $x = \sqrt{\frac{\pi}{l}}\, y$. Also, let $k_n = n\sqrt{\frac{\pi}{l}}$. You can now find that $\Delta k_n = \sqrt{\frac{\pi}{l}}$ and $\frac{n\pi y}{l} = k_n x$. Redefine $F\left(\sqrt{\frac{l}{\pi}}\, x\right) = f(x)$. Introduce a new symbol for the coefficients that indicates the index k_n replaces n, so $c_n \to g_{k_n}$. Finally, let $l \to \infty$, so that the Fourier transforms

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k)\, e^{ikx}\, dk \quad \text{and} \quad g(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx \quad \text{are exposed.}$$

$$F(y) = \sum_{n=-\infty}^{\infty} \left(\frac{1}{2l} \int_{-l}^{l} F(y)\, e^{-in\pi y/l}\, dy \right) e^{in\pi y/l} = \sum_{n=-\infty}^{\infty} \left(\frac{1}{\sqrt{2l}} \int_{-l}^{l} F(y)\, e^{-in\pi y/l}\, dy \right) \frac{1}{\sqrt{2l}}\, e^{in\pi y/l} .$$

We can now identify a symmetrized function and coefficient as

$$F(y) = \sum_{n=-\infty}^{\infty} c_n \frac{1}{\sqrt{2l}}\, e^{in\pi y/l} \quad \text{and} \quad c_n = \frac{1}{\sqrt{2l}} \int_{-l}^{l} F(y)\, e^{-in\pi y/l}\, dy . \tag{1}$$

We are going to change variables to $\sqrt{\frac{\pi}{l}}\, y = x \quad \Rightarrow \quad y = \sqrt{\frac{l}{\pi}}\, x \quad \Rightarrow \quad dy = \sqrt{\frac{l}{\pi}}\, dx$. We are also going to use $k_n = n\sqrt{\frac{\pi}{l}} \quad \Rightarrow \quad \Delta k_n = k_{n+1} - k_n = (n+1)\sqrt{\frac{\pi}{l}} - n\sqrt{\frac{\pi}{l}} = \sqrt{\frac{\pi}{l}}$. These substitutions allow us to rewrite a portion of the argument of the exponentials, $\frac{n\pi y}{l} = n\sqrt{\frac{\pi}{l}}\sqrt{\frac{\pi}{l}}\, y = k_n x$. The difference in the new index is $\Delta k_n = \sqrt{\frac{\pi}{l}} \quad \Rightarrow \quad \frac{\Delta k_n}{\sqrt{2\pi}} = \frac{1}{\sqrt{2l}}$. The limits of the integral become $y = l \quad \Rightarrow \quad x = \sqrt{\frac{\pi}{l}}\, l = \sqrt{\pi l}$, and $y = -l \quad \Rightarrow \quad x = -\sqrt{\pi l}$. We finally introduce a new symbol for the coefficients that indicates the index is now k_n vice n so $c_n \to g_{k_n}$. Using these substitutions and developments in equation (1),

$$F(y) = F\left(\sqrt{\frac{l}{\pi}}\, x\right) = \sum_{k_n=-\infty}^{\infty} g_{k_n} \frac{1}{\sqrt{2\pi}}\, e^{ik_n x}\, \Delta k_n = \frac{1}{\sqrt{2\pi}} \sum_{k_n=-\infty}^{\infty} g_{k_n}\, e^{ik_n x}\, \Delta k_n$$

$$\text{and} \qquad g_{k_n} = \frac{1}{\sqrt{2l}} \int_{-\sqrt{\pi l}}^{\sqrt{\pi l}} F\left(\sqrt{\frac{l}{\pi}}\, x\right) e^{-ik_n x} \sqrt{\frac{l}{\pi}}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{\pi l}}^{\sqrt{\pi l}} F\left(\sqrt{\frac{l}{\pi}}\, x\right) e^{-ik_n x}\, dx$$

Redefine $F\left(\sqrt{\frac{l}{\pi}}\, x\right) = f(x)$. Now, let the period $l \to \infty \quad \Rightarrow \quad \Delta k_n = \sqrt{\frac{\pi}{l}} \to 0$. The difference between successive $k_n \to 0$ so that k_n assumes all real values and thus becomes a continuous variable. The summation over discrete k_n becomes an integral over the continuous variable k, the coefficient g_{k_n} becomes a continuous function $g(k)$, the difference Δk_n becomes the differential dk, and

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

Postscript: The factor of $1/2\pi$ must be in the Fourier integral but it does not need to be symmetrized in the Fourier transforms. Some authors and tables place the entire factor $1/2\pi$ with one integral so that there is no coefficient in front of the other integral. Also, some authors and tables will change the signs of the exponentials from what we have presented. Asymmetric treatment of the factor of $1/2\pi$ and/or use of a different sign convention result in slightly different functional forms for resulting Fourier transforms.

25. Find the momentum space wavefunction corresponding to a position space wavefunction that is Gaussian.

Gaussian wavefunctions are foundational as you will find in chapter 3 and elsewhere. But first, what does it mean to change a wavefunction in position space to a wavefunction in momentum space? It means find the description of a system using an unknown function of the continuous variable momentum starting with the description of the system that is a known function of the continuous variable position. It is a change of basis in the realm of continuous variables. Quantum mechanical Fourier transforms are the means to change from position space to momentum space and vice versa for continuous variables.

The de Broglie relation associates wavelength or wavenumber, and momentum,

$$\lambda = \frac{h}{p} \quad \Rightarrow \quad \frac{2\pi}{k} = \frac{h}{p} \quad \Rightarrow \quad k = \frac{p}{\hbar}\,.$$

Substitute $p/\hbar$ for k in both Fourier transforms. This means the differential $dk = dp/\hbar$ in the integral for $f\,(x)$. Symmetrize the $1/\hbar$ from this differential by placing a factor $1/\sqrt{\hbar}$ in both integrals, and the result is the quantum mechanical form of the Fourier transforms,

$$f\,(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \widehat{g}\,(p)\, e^{ipx/\hbar}\, dp, \qquad \widehat{g}\,(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} f\,(x)\, e^{-ipx/\hbar}\, dx\,.$$

The usual use of these relations is to transform a wave function in position space to momentum space (or vice versa), so will be applied

$$\psi\,(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \widehat{\psi}\,(p)\, e^{ipx/\hbar}\, dp, \qquad \widehat{\psi}\,(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi\,(x)\, e^{-ipx/\hbar}\, dx\,.$$

There are a number of subtleties associated with the quantum mechanical analog of the Fourier transforms, so our heuristic argument should be regarded as nothing more than a useful mnemonic.

Use $\psi\,(x) = A\, e^{-bx^2}, \quad b > 0\,,$ as your position space wavefunction that is Gaussian. Form 3.323.2 from Gradshteyn and Ryzhik,

$$\int_{-\infty}^{\infty} e^{-\alpha^2 x^2 - \beta x}\, dx = \frac{\sqrt{\pi}}{\alpha}\, e^{\beta^2/4\alpha^2},$$

should be useful. Do you recognize the functional form of the momentum space wavefunction?

$$\widehat{\psi}\,(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi\,(x)\, e^{-ipx/\hbar}\, dx = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} A\, e^{-bx^2}\, e^{-ipx/\hbar}\, dx$$
$$= \frac{A}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-bx^2 - ipx/\hbar}\, dx\,.$$

Using $\alpha = \sqrt{b}\,,$ and $\beta = ip/\hbar\,,$ our integral is

$$\widehat{\psi}\,(p) = \frac{A}{\sqrt{2\pi\hbar}}\, \frac{\sqrt{\pi}}{\sqrt{b}}\, \exp\left(\frac{(ip/\hbar)^2}{4\,(\sqrt{b}\,)^2} \right) = \frac{A}{\sqrt{2\hbar b}}\, e^{-p^2/4b\hbar^2}.$$

Postscript: Notice that $\widehat{\psi}\,(p)$ is also an exponential function with a negative argument that is a constant times the independent variable squared. In fact, $\widehat{\psi}(p)$ is also a Gaussian wavefunction. The Fourier transform of a Gaussian function is another Gaussian function.

---

26. Given $|\,\psi>$ such that $<x\,|\,\psi> = \psi\,(x) = \mathrm{A}\,e^{ik_0x}, \quad -\dfrac{4\pi}{k_0} < x < \dfrac{4\pi}{k_0}$, and 0 otherwise,

(a) normalize $\psi\,(x)$,

(b) find $\widehat{\psi}\,(p)$, and

(c) sketch $\psi\,(x)$ and $\widehat{\psi}\,(p)$.

The physical system described by $|\,\psi>$ is a plane wave confined to a one dimensional "box," also known as a one-dimensional infinite square well. Part (a) means to solve for the constant A using the normalization condition on the interval to which the system is confined. In other words, the confinement allows you to use $-4\pi/k_0$ and $4\pi/k_0$ as the limits of integration. Of course, an unconfined plane wave cannot be normalized. Again, you need only to integrate between the limits of $-4\pi/k_0$ and $4\pi/k_0$ for part (b). "Sketch" means to understand the shape of the curves. Draw or plot only the real part of $\psi\,(x)$. Set the normalization constants, k_0, and $\hbar$ equal to 1 to get the shapes of $\psi\,(x)$ and $\widehat{\psi}\,(p)$. You should find that the real part of $\psi\,(x)$ is an evenly distributed cosine curve but that $\widehat{\psi}\,(p)$ is distinctly peaked.

(a) The calculus-based form of the normalization condition on the interval given is

$$1 = \int_{-4\pi/k_0}^{4\pi/k_0} \mathrm{A}^*\,e^{-ik_0x}\,\mathrm{A}\,e^{ik_0x}\,dx = |\,\mathrm{A}\,|^2 \int_{-4\pi/k_0}^{4\pi/k_0} e^0\,dx = |\,\mathrm{A}\,|^2 \int_{-4\pi/k_0}^{4\pi/k_0} dx$$

$$= |\,\mathrm{A}\,|^2 \left(x\Big|_{-4\pi/k_0}^{4\pi/k_0} \right) = |\,\mathrm{A}\,|^2 \left(\frac{4\pi}{k_0} - -\frac{4\pi}{k_0} \right) = |\,\mathrm{A}\,|^2\,\frac{8\pi}{k_0} \quad \Rightarrow \quad \mathrm{A} = \sqrt{\frac{k_0}{8\pi}}$$

$$\Rightarrow \quad \psi\,(x) = \sqrt{\frac{k_0}{8\pi}}\,e^{ik_0x} = \frac{1}{2}\sqrt{\frac{k_0}{2\pi}}\,e^{ik_0x}\,.$$

(b) $$\widehat{\psi}\,(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi\,(x)\,e^{-ipx/\hbar}\,dx = \frac{1}{\sqrt{2\pi\hbar}}\,\frac{1}{2}\sqrt{\frac{k_0}{2\pi}} \int_{-4\pi/k_0}^{4\pi/k_0} e^{ik_0x}\,e^{-ipx/\hbar}\,dx$$

after substituting the results of part (a). So

$$\begin{aligned}
\widehat{\psi}(p) &= \frac{1}{4\pi}\sqrt{\frac{k_0}{\hbar}}\int_{-4\pi/k_0}^{4\pi/k_0} e^{i(k_0-p/\hbar)x}\,dx = \frac{1}{4\pi}\sqrt{\frac{k_0}{\hbar}}\,\frac{1}{i(k_0-p/\hbar)}\,e^{i(k_0-p/\hbar)x}\Big|_{-4\pi/k_0}^{4\pi/k_0}\\
&= \frac{1}{4\pi}\sqrt{\frac{k_0}{\hbar}}\,\frac{1}{i(k_0-p/\hbar)}\left(e^{i(k_0-p/\hbar)(4\pi/k_0)} - e^{i(k_0-p/\hbar)(-4\pi/k_0)}\right)\\
&= \frac{1}{2\pi}\sqrt{\frac{k_0}{\hbar}}\,\frac{1}{2i(k_0-p/\hbar)}\left(e^{i4\pi(1-p/k_0\hbar)} - e^{-i4\pi(1-p/k_0\hbar)}\right)\\
&= \frac{1}{2\pi}\sqrt{\frac{1}{k_0\hbar}}\,\frac{1}{(1-p/k_0\hbar)}\,\frac{e^{i4\pi(1-p/k_0\hbar)} - e^{-i4\pi(1-p/k_0\hbar)}}{2i}\\
&= \frac{1}{2\pi}\sqrt{\frac{1}{k_0\hbar}}\,\frac{\sin 4\pi\,(1-p/k_0\hbar)}{(1-p/k_0\hbar)} \quad\Rightarrow\quad \widehat{\psi}(p) = 2\sqrt{\frac{1}{k_0\hbar}}\;\frac{\sin\left(4\pi - \dfrac{4\pi p}{k_0\hbar}\right)}{\left(4\pi - \dfrac{4\pi p}{k_0\hbar}\right)}\,.
\end{aligned}$$

The region of confinement is $-\infty < p < \infty$ in momentum space.

(c)

Sketch of $\psi(x)$ Sketch of $\widehat{\psi}(p)$

Postscript: The advantage of casting $\widehat{\psi}(p)$ in this form is that a sine divided by the argument of the sine is a **sinc function**, which is used to model numerous systems.

---

27. Given $|\,\psi>$ such that $<p\,|\,\psi> = \widehat{\psi}(p) = \dfrac{\hbar\,\sin\left(\dfrac{ap}{\hbar}\right)}{p}, \quad a\in\mathbf{R}\,,$

(a) normalize $\widehat{\psi}(p)$,

(b) find $\psi(x)$, and

(c) sketch $\widehat{\psi}(p)$ and $\psi(x)$.

The only real difference in this problem from the last is that it emphasizes momentum space. Though most of the developments that you will encounter in this and other books are in position space, momentum space can be a preferred environment.

It is useful to recognize even or odd functions and composite functions. Again, a function such that $f(x) = f(-x)$, like a cosine, is said to be even. A function such that $f(x) = -f(-x)$, like a sine, is said to be odd. An even function times another even function, or an odd function times another odd function is an even composite function. An even function times an odd function is an odd composite function. Most functions and composite functions are neither even nor odd, for instance, e^x is neither even nor odd. However, when this even/odd type of symmetry does exist, there are some useful techniques that are applicable. For instance, the integral of an odd function between symmetric limits is zero. Also, the integral of an even function between symmetric limits is twice the same integral between the limits of zero and the upper limit.

Two integrals are of interest. Form 611 from the 30th edition of the CRC tables,

$$\int_0^\infty \frac{\sin(ax)}{x^2}\, dx = \frac{\pi\,|\,a\,|}{2}$$

for part (a). Assume that $a > 0$ so that you do not have to carry the absolute value symbols. If you change variables to $y = \dfrac{ap}{\hbar}$, you will find that the above integral is useful for application of the normalization condition, given that you have used the even/odd function arguments correctly. The second integral of interest is

$$\int_0^\infty \frac{\sin bx \,\cos cx}{x}\, dx = \begin{cases} 0\,, & c > b > 0\,, \\ \pi/2\,, & b > c > 0\,, \\ \pi/4\,, & b = c > 0\,, \end{cases}$$

from the 30th edition of the CRC tables, form 615. If you correctly apply the quantum mechanical Fourier transform, the Euler equation $e^{i\theta} = \cos\theta + i\sin\theta$, and the even/odd function arguments for the integrals, you should find

$$\psi(x) = \frac{1}{\pi\hbar}\sqrt{\frac{2}{a}} \int_0^\infty \frac{\sin\left(\frac{ap}{\hbar}\right)\,\cos\left(\frac{px}{\hbar}\right)}{p/\hbar}\, dp$$

which can be caste into the desired form by changing variables using $y = p/\hbar$.

(a) The normalization condition in momentum space is

$$1 = \int_{-\infty}^\infty \widehat{\psi}^*(p)\, A^*\, A\, \widehat{\psi}(p)\, dp = |\,A\,|^2 \int_{-\infty}^\infty \frac{\hbar \sin\left(\frac{ap}{\hbar}\right)}{p}\, \frac{\hbar \sin\left(\frac{ap}{\hbar}\right)}{p}\, dp$$

$$= |\,A\,|^2 \int_{-\infty}^\infty \frac{\hbar^2 \sin^2\left(\frac{ap}{\hbar}\right)}{p^2}\, dp = |\,A\,|^2 a^2 \int_{-\infty}^\infty \frac{\sin^2\left(\frac{ap}{\hbar}\right)}{\frac{a^2p^2}{\hbar^2}}\, dp\,.$$

Let $y = \dfrac{ap}{\hbar} \quad\Rightarrow\quad dy = \dfrac{a}{\hbar}\, dp \quad\Rightarrow\quad dp = \dfrac{\hbar}{a}\, dy\,,$ and integral limits are identical,

$$\Rightarrow\quad 1 = |\,A\,|^2 a\hbar \int_{-\infty}^\infty \frac{\sin^2(y)}{y^2}\, dy = |\,A\,|^2\, 2a\hbar \int_0^\infty \frac{\sin^2(y)}{y^2}\, dy = |\,A\,|^2\, 2a\hbar\, \frac{\pi}{2} \quad\Rightarrow\quad A = \frac{1}{\sqrt{a\pi\hbar}}$$

$$\widehat{\psi}(p) = \frac{1}{\sqrt{a\pi\hbar}} \frac{\hbar \sin\left(\frac{ap}{\hbar}\right)}{p}.$$

(b) The wavefunction in position space is

$$\begin{aligned}
\psi(x) &= \frac{1}{\sqrt{2\pi\hbar}} \frac{1}{\sqrt{a\pi\hbar}} \int_{-\infty}^{\infty} \frac{\hbar \sin\left(\frac{ap}{\hbar}\right)}{p} e^{ipx/\hbar}\, dp \\
&= \frac{1}{\pi\hbar} \frac{1}{\sqrt{2a}} \int_{-\infty}^{\infty} \frac{\sin\left(\frac{ap}{\hbar}\right)}{p/\hbar} \left[\cos\left(\frac{px}{\hbar}\right) + i \sin\left(\frac{px}{\hbar}\right)\right] dp \\
&= \frac{1}{\pi\hbar} \frac{1}{\sqrt{2a}} \int_{-\infty}^{\infty} \frac{\sin\left(\frac{ap}{\hbar}\right) \cos\left(\frac{px}{\hbar}\right)}{p/\hbar} dp + \frac{1}{\pi\hbar} \frac{1}{\sqrt{2a}} \int_{-\infty}^{\infty} \frac{\sin\left(\frac{ap}{\hbar}\right) \cancel{\sin}\left(\frac{px}{\hbar}\right)}{\cancel{p}/\hbar} dp
\end{aligned}$$

where the last integral is zero because the product of three odd functions is an odd function, both sines and $1/p$ are odd functions, and the integral of an odd integrand between symmetric limits is zero. The other integrand is the product of two odd and one even function so is an even function. An even integrand between symmetric limits is twice the integral from zero to the upper limit so

$$\psi(x) = \frac{1}{\pi\hbar} \frac{2}{\sqrt{2a}} \int_{0}^{\infty} \frac{\sin\left(\frac{ap}{\hbar}\right) \cos\left(\frac{px}{\hbar}\right)}{p/\hbar} dp.$$

Changing variables to $y = \dfrac{p}{\hbar} \;\Rightarrow\; p = \hbar y \;\Rightarrow\; dp = \hbar\, dy$, so the limits are the same under the change of variables. Then

$$\psi(x) = \frac{1}{\pi\hbar} \frac{2}{\sqrt{2a}} \int_{0}^{\infty} \frac{\sin(ay)\cos(xy)}{y} \hbar\, dy = \begin{cases} 0, & x > a > 0, \\ \dfrac{1}{\sqrt{2a}}, & a > x > 0, \\ \dfrac{1}{2\sqrt{2a}}, & a = x > 0. \end{cases}$$

(c) The constant a has the dimension of length because the argument of a sine is dimensionless, which is pertinent to the graph of $\psi(x)$. Graphs look like

Sketch of $\widehat{\psi}(p)$ Sketch of $\psi(x)$

Postscript: Notice that the momentum space wavefunction given for problem 27 is a sinc function (or can be caste in the form of a sinc function by multiplying top and bottom by a). It shows that the Fourier transform of a sinc function is a constant function. Constants and sinc functions are related by their Fourier transforms in general.

The position space wavefunction given in problem 26 is a plane wave. The only part that survives the integral transform is the periodic cosine portion. The cosine function is closely enough related to a constant function that its Fourier transform on a finite interval is also a sinc function.

Supplementary Problems

29. Form the outer product of

$$|\,v\!> = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \text{and} \qquad <\!w\,| = \begin{pmatrix} 2, & 3, & 4 \end{pmatrix}.$$

This problem is the mechanics of forming an outer product. See problem 1. Notice that an outer product is an operator.

---

30. Show that $\mathcal{P}_i^2 = \mathcal{P}_i$ using Dirac notation.

This problem should help familiarize you with both the projection operator and Dirac notation. First, the meaning of $\mathcal{P}_i^2$ is $\mathcal{P}_i^2 = \mathcal{P}_i\,\mathcal{P}_i$. An operator with an exponent means to apply that operator the exponent number of times consecutively. Secondly, a projection operator is an outer product, so simply form the outer products in Dirac notation. Lastly, recognize the orthonormality condition within the product that you have written. This "proof" is very short. The use of Dirac notation is often simply an exercise in recognizing the meaning of the symbols.

---

31. Show that $<\!f\,|\,g\!>$ is invariant under unitary transformation where

$$|\,f\!> = \begin{pmatrix} i \\ 0 \\ 1+i \end{pmatrix} \qquad \text{and} \qquad |\,g\!> = \begin{pmatrix} 1 \\ 1-i \\ 2i \end{pmatrix},$$

and the unitary transformation is made from the operator $\mathcal{L}_x = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

Invariance is an important property. This problem should provide additional confidence that an inner product is invariant. It should also provide practice in self-consistently transforming vectors though the point of the problem is coordinate independence. This problem is the complex number analogy of problem 6. You should find that $<f\,|\,g> = 2+i$. Problem 24 of part 2 provides the eigenvectors of $\mathcal{L}_x$ which are

$$|-\sqrt{2}> = \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}, \qquad |\,0> = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \qquad \text{and} \qquad |\,\sqrt{2}> = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}.$$

Form the unitary transformation and its adjoint using these eigenvectors. Then find $<f'\,| = <f\,|\,\mathcal{U}$ and $|\,g'> = \mathcal{U}^\dagger\,|\,g>$. The inner product of these vectors is also $2+i$, so $<f'\,|\,g'> = <f\,|\,g>$.

---

32. Evaluate

(a) $\displaystyle\int_{-\pi}^{\pi} \sin(\theta)\,\delta\left(\theta - \frac{\pi}{6}\right) d\theta$

(b) $\displaystyle 2\int_{0}^{\infty} \cos(x)\,e^{-x^2}\,\delta\left(\theta - \frac{\pi}{2}\right) dx$

(c) $\displaystyle\int \frac{\text{A}}{x^2+b^2}\,\delta(x+1)\,dx$

(d) $\displaystyle\int \frac{\sin(p)\,\cos(p)}{p}\,\delta\left(p - \frac{\pi}{4}\right) dp$

Per problem 11, an integral containing a delta function is the value of the integrand evaluated at the point that the argument of the delta function is equal to zero, provided that point is between the limits of the integral. The value of the integral is zero if that point is not within the limits of integration. The integral for parts (c) and (d) have no limits, but are not likely to be indefinite integrals because they contain a delta function. The usual convention is the limits are $-\infty$ and ∞ for a definite integral where the limits are not specified.

---

33. Show that $\displaystyle\int f(x)\,\delta\Big(ax + b(x-c)\Big)\,dx = \frac{1}{a+b}\,f\left(\frac{bc}{a+b}\right), \quad a,\ b > 0.$

This problem is an extension of problem 12. As before, the limits of the integral are negative and positive infinity. Change variables to $y =$ the argument of the delta function, and follow the procedures of problem 12. The qualification that the multiplicative constants are greater than zero means that you have only one case to consider.

There are basically two techniques that are applicable to delta functions—changing variables and integration by parts. Changing variables to make the argument of the delta function a single variable allows the definition and properties of the delta function to be clearly applied.

---

34. Show that $\delta(x-a)\,\delta(x-b) = \delta(a-b)$.

This problem should extend your understanding delta functions. The technique used for this problem is unusual. It is neither changing variables nor integrating by parts. It is posed partially to motivate you to visit the realm of the originator of the Dirac delta function[7].

The delta function is even so $\delta(x-a) = \delta(a-x)$, for instance. Operate on the integral with $\int f(a)\, da$. Switch the order of integration so that you can do the integration over da which will yield an integral over dx. The resulting integral over dx can be true if and only if $x = a$, which is equivalent to integrating over a instead of x, so replace x with a. Compare this with the double integral formed by operating with $\int f(a)\, da$ and the result should emerge. It may be best to see the reference given in footnote 7.

---

35. Find $\displaystyle\int V_0 \left(e^{-2\alpha x} - 2e^{-\alpha x}\right) \delta'(x-a)\, dx\,.$

This problem includes the derivative of a delta function. This potential is realistic and is used to model diatomic molecules where $x = (r - r_0)/r_0$, where r_0 will depend upon the diatomic molecule being modeled. It is known as the **Morse potential**. Calculate the derivative of the function. The integral is the negative of this derivative evaluated at the point where the argument of δ' is zero. You should find that the value of this integral is $2\,V_0\,\alpha\left(e^{-2\alpha a} - e^{-\alpha a}\right)$.

---

36. Show that

$$\mathcal{B}\,|\,a_j\!> \;= \sum_i^3 |\,a_i\!>\, b_j\, \delta_{ij} \quad\Rightarrow\quad \mathcal{B}\,|\,a_j\!> \;= b_j\,|\,a_j\!> \quad \text{using explicit addition.}$$

This problem should increase your familiarity with Dirac notation and convey the fact that a Kronecker delta within a summation has the capability to make a summation "disappear." This technique is used frequently where Dirac notation is employed. Expand the finite summation. The index j can be 1, 2, or 3 in this three dimensional problem. The condition $\delta_{ij} = 1$ only when $i = j$ forces all but one of the three possibilities to be zero for each value of j.

---

37. Show that

$$\sum_i^2 \sum_j^2 |\,a_i\!><\!a_i\,|\,\mathcal{B}\,|\,a_j\!><\!a_j\,|\,\delta_{ij} = \sum_i^2 |\,a_i\!><\!a_i\,|\,\mathcal{B}\,|\,a_i\!><\!a_i\,| \quad \text{using explicit addition.}$$

This is a toy problem intended to provide insight into how a summation that includes a Kronecker delta is simplified. It is similar to the previous problem. Explicit addition means to write out the terms of the finite sum. This problem may be easier than the last problem because both indices, i and j, are explicitly indicated in the summations so are explicitly indicated in the expansion. If you recognize that $<\!a_i\,|\,\mathcal{B}\,|\,a_j\!>$ is a scalar, you can simplify the notation.

[7] Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, England, 1958), 4th ed., pp. 60.

38. Show that if $|\psi>$ is normalized, so is $|\psi'> = e^{i\phi}|\psi>$.

This problem reinforces the meaning of normalization but also introduces the concept of phase. The factor $e^{i\phi}$ is the phase and the scalar ϕ is the phase angle, though at times the scalar ϕ is called the phase because it is the only portion of the exponential that can vary. You are given $<\psi|\psi> = 1$ to show that $<\psi'|\psi'> = 1$. Form the bra $<\psi'|$ and then the braket $<\psi'|\psi'>$ and you should see that the tenet is true.

39. Normalize $\psi(x) = \dfrac{\mathrm{A}}{x^2+b^2}$.

This problem is practice in the mechanics of normalization. The function is known as a Lorentzian function and has numerous applications. Use the normalization condition where A is a normalization constant. Form 3.249.1 on page 294 of Gradshteyn and Ryzhik is

$$\int_0^\infty \frac{dx}{(x^2+a^2)^n} = \frac{(2n-3s)!!}{2\,(2n-2)!!}\,\frac{\pi}{a^{2n-1}}.$$

You need to recognize the even integrand between symmetric limits to use this integral. The double factorial is similar to a single factorial except the factors are reduced by two, that is

$$(2n)!! = (2n)(2n-2)(2n-4)...6\cdot 4\cdot 2, \quad \text{or} \quad (2n+1)!! = (2n+1)(2n-1)(2n-3)...5\cdot 3\cdot 1.$$

The fact that factorials have precedence over multiplication may help with the reduction. You should find that $\psi(x) = \left(\dfrac{2b^3}{\pi}\right)^{1/2} \dfrac{1}{x^2+b^2}$.

40. Show that

(a) $\psi_n(\phi) = e^{in\phi}$ and $\psi_m(\phi) = e^{im\phi}$ are orthogonal on the interval $-\pi < \phi < \pi$ where n and m are integers.

(b) Orthormalize $\psi_n(\phi)$ and $\psi_m(\phi)$.

The intent of this problem is similar to problem 21. You will likely find the exponentials easier than the sines and cosines. You need to show that $<\psi_n|\psi_n> \neq 0$ and $<\psi_m|\psi_m> \neq 0$. You also need to establish that $<\psi_n|\psi_m> = 0$ for all integral n and m. Convert the inner products to integrals for the actual calculation since you have functional forms. The limits of the integration are $-\pi$ and π, because that is the region specified. Euler's equation is $e^{ix} = \cos x + i\sin x$, and if you subtract the same relation for the opposite argument $e^{-ix} = \cos x - i\sin x$, and divide by $2i$ you attain $\sin x = \dfrac{e^{ix}-e^{-ix}}{2i}$, which is an identity that should be useful.

41. Given $<f| = \big(1,\ 2,\ 3\big)$ and $<g| = \big(4,\ 5,\ 6\big)$, show that $<f|g> = 32$ after insertion of the identity two ways.

(a) Show that $<f\,|\left(\sum_1^3 |\,i><i\,|\right)|\,g> = 32\,,$ and then

(b) show that $\sum_1^3 <f\,|\,i><i\,|\,g> = 32\,.$

This problem is a discrete space analog intended to persuade that bras and kets can be moved into or out of an integral as done in problem 18. Part (a) means to assemble the identity operator before multiplying by the vectors. Part (b) means to multiply each of the unit vectors that would compose the outer product in part (a) and sum the terms after multiplication is complete. After understanding the process in a three-dimensional space, imagine the same process in an infinite dimensional space where an integral instead of a summation is appropriate. Does the process change because the dimension of the space changes?

---

22. A system is described by a Hamiltonian $\mathcal{H}$ and by a second observable Ω where

$$\mathcal{H} = \begin{pmatrix} 4 & -i & 0 \\ i & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for which} \quad |\,\psi(0)\!> \,= \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix}.$$

(a) Show that $\mathcal{H}$ and Ω commute, *i.e.*, show that $\big[\mathcal{H},\, \Omega\big] = 0$.

(b) If the energy is measured, what results can be obtained and with what probabilities will these results be obtained? If Ω is measured, what results can be obtained and with what probabilities will these results be obtained?

(c) Calculate the expectation values of the Hamiltonian $<\mathcal{H}>$ and the Omega operator $<\Omega>$ using the initial state vector. Then show that your expectation values agree with your part (b) probabilities and eigenvalues by using the general expression $<\Omega> = \sum_i P(\omega_i)\,\omega_i$.

(d) Calculate the time-dependent state vector $|\,\psi(t)\!>$ in the energy eigenbasis.

(e) Transform to the basis that simultaneously diagonalizes $\mathcal{H}$ and Ω. Calculate the new form of the initial state vector $|\,\psi(0)\!>$ in this diagonal basis.

(f) Repeat parts (b) and (c) in the diagonal basis and compare your calculations.

(g) Calculate the time evolution of the state vector in the diagonal basis. Calculate the possibilities and probabilities of measuring the energy, E_i, and the "omega-ness," ω_j, at time t.

(h) Describe the state vector immediately after each measurement and the result of each measurement if you do a gedanken experiment by alternating $\mathcal{H}$ and Ω measurements starting with an $\mathcal{H}$ measurement, *i.e.*, you measure $\mathcal{H},\ \Omega,\ \mathcal{H},\ \Omega,\ \mathcal{H},\ldots$, for the two possible cases:

1) You find 5 when you first measure $\mathcal{H}$.
2) You find 3 when you first measure $\mathcal{H}$.

(i) Explain how the two-measurement process removes the degeneracy in $\mathcal{H}$.

This problem is a reminder that the members of a complex linear vector space are complex numbers. While the mathematical mechanics is likely clearest when components and elements are from the subset of real numbers, the development is incomplete if calculations are devoid of imaginary numbers. You should find that $\mathcal{H}$ is degenerate so a second operator that commutes with $\mathcal{H}$, like Ω, is required to uniquely determine the state vector after a measurement involving the degenerate eigenvalue of $\mathcal{H}$. Together, $\mathcal{H}$ and Ω form a complete set of commuting observables. These are the same operators used in problem 39 in part 2 of chapter 1.

These operators commute. Solve the eigenvalue/eigenvector problem for the non-degenerate operator Ω first so that you can take advantage of the common eigenbasis and use its eigenvectors in the solution of the eigenvalue/eigenvector problem for $\mathcal{H}$. You should find

$$|\,\omega=0\!> \,= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \qquad |\,\omega=1\!> \,= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad |\,\omega=2\!> \,= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad \text{and}$$

$$|\,E_1=3\!> \,= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \qquad |\,E_2=3\!> \,= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad |\,E=5\!> \,= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}.$$

Remember to conjugate when forming bras and remember also to transform the state vector in part (e) as $\mathcal{U}^\dagger \,|\,\psi\!>$ when preparing to do calculations in the diagonal basis.

(a) The Hermitian operators $\mathcal{H}$ and Ω commute, because

$$\big[\mathcal{H},\,\Omega\big] = \mathcal{H}\,\Omega - \Omega\,\mathcal{H} = \begin{pmatrix} 4 & -i & 0 \\ i & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 4 & -i & 0 \\ i & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 4+1 & -4i-i & 0 \\ i+4i & 1+4 & 0 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} 4+1 & -i-4i & 0 \\ 4i+i & 1+4 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 5 & -5i & 0 \\ 5i & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} 5 & -5i & 0 \\ 5i & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix} = 0\,.$$

(b) Calculate the eigenvalues of $\mathcal{H}$,

$$\begin{aligned}
\det\big(\mathcal{H} - \beta\mathcal{I}\big) &= \det\begin{pmatrix} 4-\beta & -i & 0 \\ i & 4-\beta & 0 \\ 0 & 0 & 3-\beta \end{pmatrix} \\
&= (4-\beta)^2(3-\beta) - (3-\beta) = 0 \quad \Rightarrow \quad (16 - 8\beta + \beta^2)(3-\beta) - 3 + \beta = 0 \\
&\Rightarrow \quad 48 - 24\beta + 3\beta^2 - 16\beta + 8\beta^2 - \beta^3 - 3 + \beta = 0 \\
&\Rightarrow \quad -\beta^3 + 11\beta^2 - 39\beta + 45 = 0 \quad \Rightarrow \quad \beta^3 - 11\beta^2 + 39\beta - 45 = 0 \\
&\Rightarrow \quad (\beta-3)(\beta-3)(\beta-5) = 0 \quad \Rightarrow \quad \beta = 3,\ 3,\ 5 \ \text{ are the eigenvalues of } \mathcal{H}.
\end{aligned}$$

$\mathcal{H}$ is degenerate at $E_i = 3$, so examine Ω in the hope that it is not degenerate.

$$\begin{aligned}
\det\big(\Omega - \omega\mathcal{I}\big) &= \det\begin{pmatrix} 1-\omega & -i & 0 \\ i & 1-\omega & 0 \\ 0 & 0 & 1-\omega \end{pmatrix} = (1-\omega)^3 - (1-\omega) = 0 \\
&\Rightarrow \quad 1 - 3\omega + 3\omega^2 - \omega^3 - 1 + \omega = 0 \quad \Rightarrow \quad -\omega^3 + 3\omega^2 - 2\omega = 0 \\
&\Rightarrow \quad \omega(\omega^2 - 3\omega + 2\omega) = 0 \quad \Rightarrow \quad \omega(\omega-1)(\omega-2) = 0
\end{aligned}$$

$\Rightarrow$ $\omega = 0,\ 1,\ 2$ are the eigenvalues of Ω. Ω is not degenerate so has conventionally determined eigenvectors, so find the eigenvectors of the non-degenerate operator first.

$$\omega = 0 \quad \Rightarrow \quad \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{rcl} a - bi &=& 0 \\ ai + b &=& 0 \\ c &=& 0 \end{array} \quad \Rightarrow \quad \begin{array}{rcl} b &=& -ai \\ b &=& -ai \\ c &=& 0 \end{array}$$

$$\Rightarrow \quad |\,\omega = 0\!> \; = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \quad \text{letting } a = 1 \text{ and normalizing. Next}$$

$$\omega = 1 \quad \Rightarrow \quad \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 1\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{array}{rcl} a - bi &=& a \\ ai + b &=& b \\ c &=& c \end{array} \quad \Rightarrow \quad \begin{array}{rcl} b &=& 0 \\ a &=& 0 \\ c &=& c \end{array}$$

$$\Rightarrow \quad |\,\omega = 1\!> \; = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \text{letting } c = 1\,. \text{ This is already normalized. Then}$$

$$\omega = 2 \quad \Rightarrow \quad \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 2 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{matrix} a - bi & = & 2a \\ ai + b & = & 2b \\ c & = & 2c \end{matrix} \quad \Rightarrow \quad \begin{matrix} b & = & ai \\ ai & = & b \\ c & = & 0 \end{matrix}$$

$\Rightarrow \quad |\,\omega = 2> = \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$, when normalized if $a = 1$. The non-degenerate eigenvalue of $\mathcal{H}$ is

$$E = 5 \quad \Rightarrow \quad \begin{pmatrix} 4 & -i & 0 \\ i & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 5 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{matrix} 4a - bi & = & 5a \\ ai + 4b & = & 5b \\ 3c & = & 5c \end{matrix} \quad \Rightarrow \quad \begin{matrix} b & = & ai \\ ai & = & b \\ c & = & 0 \end{matrix}$$

$\Rightarrow \quad |\,E = 5> = \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$, letting $a = 1$ and normalizing. For $E = 3$,

$$\begin{pmatrix} 4 & -i & 0 \\ i & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 3 \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \Rightarrow \quad \begin{matrix} 4a - bi & = & 3a \\ ai + 4b & = & 3b \\ 3c & = & 3c \end{matrix} \quad \Rightarrow \quad \begin{matrix} b & = & -ai \\ b & = & -ai \\ c & = & c \end{matrix}$$

and we need two eigenvectors from this one set of equations that are orthogonal to each other and to $|\,E = 5>$. Since Ω and $\mathcal{H}$ commute, they have a common set of eigenvectors. Examining the eigenvectors of Ω, $|\,E = 5> = |\,\omega = 2>$, and $|\,\omega = 0>$ and $|\,\omega = 1>$ both satisfy the one set of equations corresponding to the eigenvalue $E = 3$, so choose

$$|\,E_1 = 3> = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \quad \text{and} \quad |\,E_2 = 3> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The possibilities of a measurement are the eigenvalues per postulate 3. Probabilities are

$$P\,(\omega = 0) = \left| \big(1,\ i,\ 0\big) \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} \right|^2 = \left| \frac{1}{2}\,(i) \right|^2 = \frac{1}{4}\,(-i)\,(i) = \frac{1}{4} = P\,(E_1 = 3)\,,$$

$$P\,(\omega = 1) = \left| \big(0,\ 0,\ 1\big) \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} \right|^2 = \frac{1}{2}\,|\,1\,|^2 = \frac{1}{2} = P\,(E_2 = 3)\,,$$

$$P\,(\omega = 2) = \left| \big(1,\ -i,\ 0\big) \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} \right|^2 = \left| \frac{1}{2}\,(i) \right|^2 = \frac{1}{4}\,(-i)\,(i) = \frac{1}{4} = P\,(E = 5)\,,$$

(c) The $t = 0$ expectation values of $\mathcal{H}$ and Ω are $<\psi\,(0)\,|\,\mathcal{A}\,|\,\psi\,(0)>$, or

$$<\mathcal{H}> = \big(-i,\ 0,\ 1\big) \frac{1}{\sqrt{2}} \begin{pmatrix} 4 & -i & 0 \\ i & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2} \big(-i,\ 0,\ 1\big) \begin{pmatrix} 4i \\ -1 \\ 3 \end{pmatrix} = \frac{1}{2}\,(4 + 3) = \frac{7}{2}\,,$$

$$<\Omega> = \big(-i,\ 0,\ 1\big) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2} \big(-i,\ 0,\ 1\big) \begin{pmatrix} i \\ -1 \\ 1 \end{pmatrix} = \frac{1}{2}\,(1 + 1) = 1\,.$$

The sums of the products of the probabilities and the corresponding eigenvalues yield

$$< \mathcal{H} > = \sum_i P\left(E_i\right) E_i = \frac{1}{4}\,(3) + \frac{1}{2}\,(3) + \frac{1}{4}\,(5) = \frac{3+6+5}{4} = \frac{14}{4} = \frac{7}{2}\,,$$

$$\text{and} \qquad < \Omega > = \sum_i P\left(\omega_i\right) \omega_i = \frac{1}{4}\,(0) + \frac{1}{2}\,(1) + \frac{1}{4}\,(2) = \frac{1}{2} + \frac{1}{2} = 1\,.$$

(d) Expanding the initial state in terms of the eigenvectors of $\mathcal{H}$,

$$\begin{aligned}
|\,\psi\,(0)> &= \sum_j |\,j><j\,|\,\psi\,(0)> = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix} \left(1,\ i,\ 0\right) \frac{1}{\sqrt{2}}\,\frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} \\
&\quad + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \left(0,\ 0,\ 1\right) \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \left(1,\ -i,\ 0\right) \frac{1}{\sqrt{2}}\,\frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} \\
&= \frac{1}{2\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}(i) + \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}(1) + \frac{1}{2\sqrt{2}}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}(i) \\
&= \frac{1}{2\sqrt{2}}\begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{2\sqrt{2}}\begin{pmatrix} i \\ -1 \\ 0 \end{pmatrix},
\end{aligned}$$

$$\Rightarrow \qquad |\,\psi\,(t)> = \sum_j |\,j><j\,|\,\psi\,(0)> e^{-iE_j t/\hbar}$$

$$= \frac{1}{2\sqrt{2}}\begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix} e^{-i3t/\hbar} + \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} e^{-i3t/\hbar} + \frac{1}{2\sqrt{2}}\begin{pmatrix} i \\ -1 \\ 0 \end{pmatrix} e^{-i5t/\hbar} = \frac{1}{2\sqrt{2}}\begin{pmatrix} ie^{-i3t/\hbar} + ie^{-i5t/\hbar} \\ e^{-i3t/\hbar} - e^{-i5t/\hbar} \\ 2\,e^{-i3t/\hbar} \end{pmatrix}$$

(e) Per part (a), $\left[\Omega,\ \mathcal{H}\right] = \left[\mathcal{H},\ \Omega\right] = 0$. The unitary matrix $\mathcal{U}$ formed from the eigenvectors of $\mathcal{H}$ by placing them from left to right in order of ascending eigenvalue is

$$\mathcal{U} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 \\ -i & 0 & i \\ 0 & \sqrt{2} & 0 \end{pmatrix} \quad \Rightarrow \quad \mathcal{U}^\dagger = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \end{pmatrix}.$$

The diagonal operators are

$$\mathcal{U}^\dagger \Omega\, \mathcal{U} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \end{pmatrix}\begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 \\ -i & 0 & i \\ 0 & \sqrt{2} & 0 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} 1 & i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 2i \\ 0 & \sqrt{2} & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

$$\mathcal{U}^\dagger \mathcal{H} \mathcal{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \end{pmatrix} \begin{pmatrix} 4 & -i & 0 \\ i & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ -i & 0 & i \\ 0 & \sqrt{2} & 0 \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 5 \\ -3i & 0 & 5i \\ 0 & 3\sqrt{2} & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 10 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

The state vector transforms $\mathcal{U}^\dagger \mid \psi(0) > = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} i \\ \sqrt{2} \\ i \end{pmatrix}.$

(f) Probabilities and expectation values using the diagonal basis are

$$P(\omega = 0) = P(E_1 = 3) = \left| (1,\ 0,\ 0) \frac{1}{2} \begin{pmatrix} i \\ \sqrt{2} \\ i \end{pmatrix} \right|^2 = \frac{1}{4} \mid i \mid^2 = \frac{1}{4} (-i)(i) = \frac{1}{4},$$

$$P(\omega = 1) = P(E_2 = 3) = \left| (0,\ 1,\ 0) \frac{1}{2} \begin{pmatrix} i \\ \sqrt{2} \\ i \end{pmatrix} \right|^2 = \frac{1}{4} \left| \sqrt{2} \right|^2 = \frac{2}{4} = \frac{1}{2},$$

$$P(\omega = 2) = P(E = 5) = \left| (0,\ 0,\ 1) \frac{1}{2} \begin{pmatrix} i \\ \sqrt{2} \\ i \end{pmatrix} \right|^2 = \frac{1}{4} \mid i \mid^2 = \frac{1}{4} (-i)(i) = \frac{1}{4},$$

$$< \Omega > = (-i,\ \sqrt{2},\ -i) \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \frac{1}{2} \begin{pmatrix} i \\ \sqrt{2} \\ i \end{pmatrix} = \frac{1}{4} (-i,\ \sqrt{2},\ -i) \begin{pmatrix} 0 \\ \sqrt{2} \\ 2i \end{pmatrix} = \frac{1}{4}(2+2) = 1,$$

$$< \mathcal{H} > = (-i,\ \sqrt{2},\ -i) \frac{1}{2} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} \frac{1}{2} \begin{pmatrix} i \\ \sqrt{2} \\ i \end{pmatrix} = \frac{1}{4} (-i,\ \sqrt{2},\ -i) \begin{pmatrix} 3i \\ 3\sqrt{2} \\ 5i \end{pmatrix} = \frac{3+6+5}{4} = \frac{7}{2}.$$

All results are identical to those calculated previously.

(g) The time dependence of the state vector in the diagonal basis is

$$\begin{aligned} \mid \psi(t) > &= \sum_j \mid j >< j \mid \psi(0) > e^{-iE_j t/\hbar} \\ &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} (1,\ 0,\ 0) \frac{1}{2} \begin{pmatrix} i \\ \sqrt{2} \\ i \end{pmatrix} e^{-i3t/\hbar} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} (0,\ 1,\ 0) \frac{1}{2} \begin{pmatrix} i \\ \sqrt{2} \\ i \end{pmatrix} e^{-i3t/\hbar} \\ &\quad + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} (0,\ 0,\ 1) \frac{1}{2} \begin{pmatrix} i \\ \sqrt{2} \\ i \end{pmatrix} e^{-i5t/\hbar} \end{aligned}$$

$$= \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} (i)\, e^{-i3t/\hbar} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \sqrt{2}\, e^{-i3t/\hbar} + \frac{1}{2} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} (i)\, e^{-i5t/\hbar} = \frac{1}{2} \begin{pmatrix} i\, e^{-i3t/\hbar} \\ \sqrt{2}\, e^{-i3t/\hbar} \\ i\, e^{-i5t/\hbar} \end{pmatrix}.$$

The probabilities $P(E_1 = 3) = P(\omega = 0)$ are

$$\left| \left(1,\ 0,\ 0\right) \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} i\, e^{-i3t/\hbar} \right|^2 = \frac{1}{4} \left| i\, e^{-i3t/\hbar} \right|^2 = \frac{1}{4} \left(-i\, e^{+i3t/\hbar} \right) \left(i\, e^{-i3t/\hbar} \right) = \frac{1}{4} .$$

The probabilities $P(E_2 = 3) = P(\omega = 1)$ are

$$\left| \left(0,\ 1,\ 0\right) \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \sqrt{2}\, e^{-i3t/\hbar} \right|^2 = \frac{1}{4} \left| \sqrt{2}\, e^{-i3t/\hbar} \right|^2 = \frac{1}{4} \left(\sqrt{2}\, e^{+i3t/\hbar} \right) \left(\sqrt{2}\, e^{-i3t/\hbar} \right) = \frac{1}{2} .$$

The final set of probabilities $P(E = 5) = P(\omega = 2)$ are

$$\left| \left(0,\ 0,\ 1\right) \frac{1}{2} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} i\, e^{-i5t/\hbar} \right|^2 = \frac{1}{4} \left| i\, e^{-i5t/\hbar} \right|^2 = \frac{1}{4} \left(-i\, e^{+i5t/\hbar} \right) \left(i\, e^{-i5t/\hbar} \right) = \frac{1}{4} ,$$

identical to $t = 0$ probabilities. We have used only the pertinent component from $|\,\psi\,(t)>$ which is a convenience inherent in the eigenstates being unit vectors, that is, in a diagonal basis.

(h) When you measure $E = 5$ the state vector is $|\,\psi(t > 0)> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ in the diagonal basis so the probability of measuring $\omega = 2$ is 1 with a zero probability of measuring $\omega = 0$ or $\omega = 1$. The state vector remains $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ for all $t > 0$ so a subsequent measurement of $\mathcal{H}$ is necessarily $E = 5$, a subsequent measurement of Ω is necessarily $\omega = 2$. If the result of an initial measurement is $E = 3$, the state vector is

$$|\,\psi(t > 0)> = c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \Rightarrow \quad |\,\psi(t > 0)> \rightarrow A \begin{pmatrix} i \\ \sqrt{2} \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad \left(-i,\ \sqrt{2},\ 0 \right) A^* A \begin{pmatrix} i \\ \sqrt{2} \\ 0 \end{pmatrix} = |\,A\,|^2 \left(1{+}2 \right) = 1 \quad \Rightarrow \quad A = \frac{1}{\sqrt{3}} \quad \Rightarrow \quad |\,\psi(t > 0)> = \frac{1}{\sqrt{3}} \begin{pmatrix} i \\ \sqrt{2} \\ 0 \end{pmatrix} .$$

$$\Rightarrow \quad P(\omega = 0) = \left| \left(1,\ 0,\ 0\right) \frac{1}{\sqrt{3}} \begin{pmatrix} i \\ \sqrt{2} \\ 0 \end{pmatrix} \right|^2 = \frac{1}{3} \left| \, i \, \right|^2 = \frac{1}{3} \left(-\,i \right) \left(i \right) = \frac{1}{3} ,$$

$$P(\omega = 1) = \left| \left(0,\ 1,\ 0\right) \frac{1}{\sqrt{3}} \begin{pmatrix} i \\ \sqrt{2} \\ 0 \end{pmatrix} \right|^2 = \frac{2}{3} .$$

A subsequent measurement of $\omega = 0$ puts the system into the unique state $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow E = 3$, $\omega = 0$, $E = 3$, $\omega = 0$, ad infinitum for further measuremements of $\mathcal{H}$ and Ω, or a subsequent measurement of $\omega = 1$ puts the system into the unique state $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow \; E = 3\,, \; \omega = 1\,,$ $E = 3\,, \; \omega = 1\,,$ ad infinitum for further measuremements of $\mathcal{H}$ and Ω.

(i) The degeneracy at $E = 3$ is removed because

$$E = 3 \;\text{ and }\; \omega = 0 \;\; \Rightarrow \;\; |\,\psi\!> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \;\; \text{and } \; E = 3 \;\text{ and }\; \omega = 1 \;\; \Rightarrow \;\; |\,\psi\!> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

The state vector is not uniquely determined by a measurement of $E = 3$ but the state vector is uniquely determined by a measurement of $E = 3$ and $\omega = 0$ or 1.

Postscript: The finality of the "ad infinitum" of part (h) is moderated by the fact that the system under study can and likely will interact with another system, collision with another particle for instance, so that the state vector can change. Until it does interact with another system, however, it will remain in the state that corresponds to the eigenvalue measured during any measurement.

---

23. (a) Find the position space form of an inner product,

(b) the momentum space form of an inner product, and

(c) comment on the consequence of the two kets of part (a) being dual state vectors.

Dirac notation is independent of basis so is particularly useful for exploiting the advantages of abstract Hilbert space. Conversion to a specific subspace, such as position space, is often the most expedient way to address specific physical problems. The key techniques in transitioning from Hilbert space to a specific representation are the insertion of the identity and the choice of the form for that identity. Representations described in problems 9 and 10 in part 3 of chapter 1 are pertinent. Limits are not used on integrals where the integration is to be done over all space, meaning from $-\infty$ to ∞ in each dimension. Excluding limits over all space is a common practice.

(a)

$$\begin{aligned} <\psi\,|\,\psi'> &= <\psi\,|\,\mathcal{I}\,|\,\psi'> && (1)\\ &= <\psi\,|\left(\int |\,x\!><\!x\,|\,dx\right)|\,\psi'> && (2)\\ &= \int <\psi\,|\,x\!><\!x\,|\,\psi'>\,dx, && (3) \end{aligned}$$

$$\text{but} \quad <\psi\,|\,x> = \psi^*(x) \;\text{ and }\; <x\,|\,\psi'> = \psi'(x) \;\; \Rightarrow \; <\psi\,|\,\psi'> = \int \psi^*(x)\,\psi'(x)\,dx$$

is the position space form of an inner product. The identity is inserted in line (1) and the position space form of the identity is explicitly denoted in line (2). The bra and ket may be moved as indicated in line (3) since the variable of integration is not inherent in either abstract vector.

(b) Proceeding as in the previous example but choosing a momentum space form of the identity,

$$<\psi\,|\,\psi'> = <\psi\,|\,\mathcal{I}\,|\,\psi'> = <\psi\,|\int |\,p><p\,|\,dp\,|\,\psi'> = \int <\psi\,|\,p><p\,|\,\psi'>\;dp = \int \widehat{\psi}^*(p)\,\widehat{\psi}'(p)\,dp.$$

(c) If the two vectors of part (a) are dual vectors, the result becomes

$$<\psi\,|\,\psi> = \int \psi^*(x)\,\psi(x)\,dx, \quad \text{and since} \quad <\psi\,|\,\psi> = 1 \quad \text{for a normalized vector,}$$

$\int \psi^*(x)\,\psi(x)\,dx = 1$ is the normalization condition in one spatial dimension in position space.

Postscript: Integral forms are used to describe continuous systems or systems populated so heavily that they may be treated as continuous.

---

24. Show that (a) $<x\,|\,\mathcal{X}\,|\,x'> = x\,\delta\,(x-x')$ and (b) $<x\,|\,\mathcal{P}\,|\,x'> = -i\hbar\,\delta'\,(x-x')$.

These relations are so useful in developing representation in position space that they have been stated as portions of the postulates of quantum mechanics[2].

Part (a) is a one-line application of the eigenvalue/eigenvector problem where you must recognize a delta function. Postulate 2 requires that $\mathcal{X}$ is Hermitian. Hermitian operators can operate to the right or to the left. The eigenvalue/eigenvector relation is $<x\,|\,\mathcal{X} = <x\,|\,x$ when $\mathcal{X}$ operates to the left. Unfortunately, the eigenvalue/eigenvector relation does not apply to part (b) because neither $<x\,|$ nor $|\,x'>$ is an eigenbasis of $\mathcal{P}$.

For part (b), use the quantum mechanical Fourier transforms

$$\psi\,(x) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\widehat{\psi}\,(p)\,e^{ipx/\hbar}\,dp, \qquad \widehat{\psi}\,(p) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\psi\,(x)\,e^{-ipx/\hbar}\,dx$$

to form a Fourier integral for $\psi\,(x)$. You can write $\psi\,(x) = \int_{-\infty}^{\infty}\delta\,(x-x')\,\psi\,(x')\,dx'$, set the two expressions for $\psi\,(x)$ equal, and develop relations for $\delta\,(x-x')$ and finally $\delta'\,(x-x')$. Then

$$<x\,|\,\mathcal{P}\,|\,x'> = <x\,|\,\mathcal{I}\,|\,\mathcal{P}\,|\,x'> = <x\,|\left(\int |\,p><p\,|\,dp\right)|\,\mathcal{P}\,|\,x'>,$$

and using the relations of $<x\,|\,p>$ and $<p\,|\,x>$ seen in the postscript of problem 28 from part 3 of chapter 1, you eventually arrive at expressions that you can recognize as $\delta'\,(x-x')$.

(a) Attaining the eigenvalue by letting $\mathcal{X}$ operate to the left and recognizing $<x\,|\,x'> = \delta\,(x-x')$,

$$<x\,|\,\mathcal{X}\,|\,x'> = <x\,|\,x\,|\,x'> = x<x\,|\,x'> = x\,\delta\,(x-x')\,.$$

[2] Shankar, *Principles of Quantum Mechanics* (Plenum Press, New York, 1994), 2nd ed., pp. 115-116.

(b) The quantum mechanical Fourier integral is

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int \left(\frac{1}{\sqrt{2\pi\hbar}} \int \psi(x') \, e^{-ipx'/\hbar} \, dx' \right) e^{ipx/\hbar} \, dp = \frac{1}{2\pi\hbar} \int dp \int \psi(x') \, e^{ip(x-x')/\hbar} \, dx' .$$

The wavefunction in position space can be written $\psi(x) = \int \delta(x - x') \, \psi(x') \, dx'$ so

$$\int \delta(x - x') \, \psi(x') \, dx' = \frac{1}{2\pi\hbar} \int dp \int \psi(x') \, e^{ip(x-x')/\hbar} \, dx' .$$

$$\Rightarrow \quad \delta(x - x') \, \psi(x') = \frac{1}{2\pi\hbar} \int \psi(x') \, e^{ip(x-x')/\hbar} \, dp = \frac{1}{2\pi\hbar} \, \psi(x') \int e^{ip(x-x')/\hbar} \, dp$$

$$\Rightarrow \quad \delta(x - x') = \frac{1}{2\pi\hbar} \int e^{ip(x-x')/\hbar} \, dp$$

$$\Rightarrow \quad \delta'(x - x') = \frac{1}{2\pi\hbar} \int \frac{i}{\hbar} \, p \, e^{ip(x-x')/\hbar} \, dp . \tag{1}$$

$$\begin{aligned}
<x \,|\, \mathcal{P} \,|\, x'> &= <x \,|\, \mathcal{I} \,|\, \mathcal{P} \,|\, x'> = <x \,| \left(\int |\, p><p \,|\, dp \right) |\, \mathcal{P} \,|\, x'> \\
&= \int <x \,|\, p><p \,|\, \mathcal{P} \,|\, x'> \, dp \\
&= \int <x \,|\, p><p \,|\, p \,|\, x'> \, dp = \int p \; < x \,|\, p><p \,|\, x'> \, dp && (2) \\
&= \int p \left(\frac{1}{\sqrt{2\pi\hbar}} \, e^{ipx/\hbar} \right) \left(\frac{1}{\sqrt{2\pi\hbar}} \, e^{-ipx'/\hbar} \right) dp && (3) \\
&= \frac{1}{2\pi\hbar} \int p \, e^{ip(x-x')/\hbar} \, dp = \frac{\hbar}{i} \, \delta'(x - x') = -i\hbar \, \delta'(x - x') . && (4)
\end{aligned}$$

Line (1) is differentiation with respect to x. The first expression in line (2) is the result of the eigenvalue/eigenvector equation, $< p \,|\, \mathcal{P} = < p \,|\, p$. Line (3) employs the expressions developed for $< x \,|\, p >$ and $< p \,|\, x >$ in the postscript of problem 28 in part 3 of chapter 1. The second expression in line (4) is attained by comparing the first expression in line (4) with equation (1).

Postscript: The relations above are in the position basis. Analogous relations in the momentum basis are $<p \,|\, \mathcal{P} \,|\, p'> = p \, \delta(p - p')$ and $<p \,|\, \mathcal{X} \,|\, p'> = i\hbar \, \delta'(p - p')$. The absence of a negative sign in the last equation is a choice of phase. A choice of phase is necessary.

---

25. Find the form of an expectation value of the position operator in position space.

The expectation value of position is $<\psi \,|\, \mathcal{X} \,|\, \psi>$ in Dirac notation. Insert the identity in terms of $|\, x'>$ on the left and $|\, x>$ on the right of the abstract position operator. Rearrange the expression to find a factor of $<x' \,|\, \mathcal{X} \,|\, x> = x' \, \delta(x' - x)$. One of the two integrations can be done using this delta function to find $<\psi \,|\, \mathcal{X} \,|\, \psi> = \int \psi^*(x) \, x \, \psi(x) \, dx$.

$$
\begin{aligned}
<\psi\,|\,\mathcal{X}\,|\,\psi> \;&= <\psi\,|\,\mathcal{I}\,|\,\mathcal{X}\,|\,\mathcal{I}\,|\,\psi> && (1)\\
&= <\psi\,|\,\left(\int |\,x'><x'\,|\,dx'\right)|\,\mathcal{X}\,|\left(\int |\,x><x\,|\,dx\right)|\,\psi> && (2)\\
&= \int\!\!\int <\psi\,|\,x'><x'\,|\,\mathcal{X}\,|\,x><x\,|\,\psi>\;dx'\,dx\\
&= \int\!\!\int \psi^*(x')\,x'\,\delta(x'-x)\,\psi(x)\,dx'\,dx && (3)\\
&= \int\left(\int \psi^*(x')\,x'\,\delta(x'-x)\,\psi(x)\,dx'\right)dx = \int \psi^*(x)\,x\,\psi(x)\,dx\,.
\end{aligned}
$$

The identity is inserted twice in line (1), two different forms of the identity are used in line (2), inner products and the braket (per problem 24 a)) are represented in position space in line (3).

Postscript: We stated earlier that $\mathcal{X} \rightarrow x$ in position space and problem 25 demonstrates this fact. Similarly, if more than one spatial dimension is required, $\mathcal{Y} \rightarrow y$, and $\mathcal{Z} \rightarrow z$.

---

26. Find the form of the expectation value of the momentum operator in position space.

This problem is similar to the last problem except, of course, it addresses momentum. Start with $<\psi\,|\,\mathcal{P}\,|\,\psi>$, insert identities in the same form as the last problem but you will need to use the relation $<x\,|\,\mathcal{P}\,|\,x'> = -i\hbar\,\delta'(x-x')$ addressed in problem 24. The delta function is an even function but the derivative of the delta function is an odd function so $<x'\,|\,\mathcal{P}\,|\,x> = i\hbar\,\delta'(x'-x)$ is the form that you want to use. Problem 17 of part 3 of chapter 1 may help you to evaluate the derivative of the delta function. You should find

$$
<\psi\,|\,\mathcal{P}\,|\,\psi> = -i\hbar\int \psi^*(x)\,\frac{d}{dx}\,\psi(x)\,dx\,.
$$

$$
\begin{aligned}
<\psi\,|\,\mathcal{P}\,|\,\psi> \;&= <\psi\,|\,\mathcal{I}\,|\,\mathcal{P}\,|\,\mathcal{I}\,|\,\psi>\\
&= <\psi\,|\,\left(\int |\,x'><x'\,|\,dx'\right)|\,\mathcal{P}\,|\left(\int |\,x><x\,|\,dx\right)|\,\psi>\\
&= \int\!\!\int <\psi\,|\,x'><x'\,|\,\mathcal{P}\,|\,x><x\,|\,\psi>\;dx'\,dx\\
&= \int\!\!\int \psi^*(x')\Big(i\hbar\,\delta'(x'-x)\Big)\,\psi(x)\,dx'\,dx && (1)\\
&= \int \psi^*(x)\Big(-i\hbar\,\frac{d}{dx}\Big)\,\psi(x)\,dx = -i\hbar\int \psi^*(x)\,\frac{d}{dx}\,\psi(x)\,dx\,. && (2)
\end{aligned}
$$

The inner products and the derivative of the delta function are both represented in position space in line (1). Evaluating the derivative of the delta function results in the first expression in line (2).

Postscript: This problem demonstrates that $\mathcal{P} \rightarrow -i\hbar \dfrac{d}{dx}$ in position space as we have indicated in part 3 of chapter 1. The parentheses in line (2) are placed to identify this explicitly.

This position space relation and $\mathcal{X} \rightarrow x$ are easily generalized to two or three dimensions by working two or three one-dimensional problems. If more than one spatial component is required,

$$\mathcal{P} \rightarrow -i\hbar\nabla = -i\hbar\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}\right), \quad \text{in three spatial dimensions, for instance.}$$

The analogous relations in momentum space in one dimension are $\mathcal{X} \rightarrow i\hbar \dfrac{d}{dp}$ and $\mathcal{P} \rightarrow p$, and in three dimensions are

$$\mathcal{X} \rightarrow i\hbar\frac{d}{dp_x}, \quad \mathcal{Y} \rightarrow i\hbar\frac{d}{dp_y}, \quad \mathcal{Z} \rightarrow i\hbar\frac{d}{dp_z}, \quad \text{and} \quad \mathcal{P}_x \rightarrow p_x, \quad \mathcal{P}_y \rightarrow p_y, \quad \mathcal{P}_z \rightarrow p_z .$$

The absence of a negative sign in the relations for spatial operators in momentum space reflects the popular choice of phase, also mentioned in the postscript to problem 24.

---

27. Derive the position space form of the time-independent Schrodinger equation in one dimension. Assume that potential energy is a function of position only.

Basis-independent statements of the Schrodinger equation offer generality and flexibility and can provide physical insight, but they are not often useful to attain functional or numeric solutions to real problems. Representations of the Schrodinger equation in position space and momentum space are the tools generally used to address specific systems.

Start with the eigenvalue/eigenvector equation for the time-independent, quantum mechanical Hamiltonian, $\mathcal{H}\,|\,\psi\rangle = E\,|\,\psi\rangle$. Form the quantum mechanical Hamiltonian,

$$H_{\text{classical}} = T + V = \frac{p^2}{2m} + V \quad \longrightarrow \quad \mathcal{H} = \frac{\mathcal{P}^2}{2m} + \mathcal{V}.$$

Potential energy is given to be a function of position only $\Rightarrow$ $\mathcal{V} \rightarrow \mathcal{V}(\mathcal{X})$ and $\mathcal{X} \rightarrow x$ $\Rightarrow$ $\mathcal{V}(\mathcal{X}) \rightarrow V(x)$ for this problem. Form a braket of both sides using $\langle x\,|$. Assume that this bra commutes with the Hamiltonian. You should find

$$-\frac{\hbar^2}{2m}\frac{d^2}{d^2x}\,\psi(x) + V(x)\,\psi(x) = E\,\psi(x).$$

$$\begin{aligned} \mathcal{H}\,|\,\psi\rangle = E\,|\,\psi\rangle \quad &\Rightarrow \quad \left(\frac{\mathcal{P}^2}{2m} + \mathcal{V}\right)|\,\psi\rangle = E\,|\,\psi\rangle \\ &\Rightarrow \quad \left(\frac{\mathcal{P}^2}{2m} + \mathcal{V}(\mathcal{X})\right)|\,\psi\rangle = E\,|\,\psi\rangle \qquad (1) \end{aligned}$$

$$\Rightarrow \;<x\,|\left(\frac{\mathcal{P}^2}{2m}+\mathcal{V}\,(\mathcal{X})\right)|\,\psi> \;=\; <x\,|\,E\,|\,\psi> \tag{2}$$

$$\Rightarrow \;\left(\frac{1}{2m}\,\mathcal{P}^2+\mathcal{V}\,(\mathcal{X})\right) <x\,|\,\psi> \;=\; E<x\,|\,\psi> \tag{3}$$

$$\Rightarrow \;\left[\frac{1}{2m}\left(-i\hbar\,\frac{d}{dx}\right)\left(-i\hbar\,\frac{d}{dx}\right)+V(x)\right]\psi\,(x) = E\,\psi\,(x) \tag{4}$$

$$\Rightarrow \;-\frac{\hbar^2}{2m}\,\frac{d^2}{d^2x}\,\psi\,(x)+V\,(x)\,\psi\,(x) \;=E\,\psi\,(x)\,.$$

Equation (1) imbeds the assumption that potential energy is a function of position only. Brakets are formed with $<x\,|$ so as to eventually attain a position space representation in equation (2). The E on the right side of equation (3) is an eigenvalue and any constant can be removed from a braket. The expression in parenthesis on the left side of equation (3) is treated as an eigenvalue. The position space representations of the momentum operator, the potential energy operator, and the inner products are substituted in equation (4).

Postscript: The assumption that $<x\,|$ commutes with the Hamiltonian is founded by the requirement for eigenstates to be unique, a property that results from the linear independence of the eigenstates. There are an infinite number of $\psi\,(x)$'s or $\psi_i\,(x)$'s. The right side of equation (2) becomes a constant times an eigenfunction. In order for the left side of equation (2) to be the same in an infinite number of cases, it also must become a constant times a function which is why $<x\,|$ commutes with $\mathcal{H}$ in equation (3). In general, an arbitrary bra does not commute with a Hermitian operator.

---

28. Consider a particle described by the wavefunction

$$\psi\,(x) = \mathrm{A} \quad \text{when} \quad -10\text{ Å} \leq \; x \; \leq +10\text{ Å}, \qquad\qquad \psi\,(x) = 0 \quad \text{otherwise.}$$

(a) Calculate the normalization constant A .

(b) Calculate the probability that the particle will be found between $x = -3$ Å and $x = +5$ Å .

(c) What are the possible results and the probability of each possible result of a measurement of position?

(d) A measurement of position is made and the particle is found at $x = 3.5$ Å. Sketch the wavefunction immediately after the measurement.

(e) Calculate the momentum space wavefunction $\widehat{\psi}\,(p)$ that corresponds to the position space wavefunction $\psi\,(x)$ by calculating the Fourier transform of $\psi\,(x)$. Sketch $\widehat{\psi}\,(p)$.

(f) The momentum of the particle described by $\psi\,(x)$ is measured. What results can be obtained, and with what probabilities will they be obtained?

(g) Suppose the momentum is measured and $p = 0$ is obtained. Sketch $\widehat{\psi}\,(p)$ immediately after the measurement.

Integral forms are used to describe continuous systems or systems populated so heavily that they may be treated as continuous systems. This problem introduces the application of the postulates to such systems. It also introduces the methods of calculating the possible results of a measurement and the probabilities of these possible results for continuous systems. The term **probability density** is introduced. This problem also introduces the concept of the units of a wavefunction.

The normalization condition is given by

$$< \psi \,|\, \psi > = \int_{-\infty}^{\infty} \psi^*(x)\, \psi\,(x)\, dx = 1\,,$$

as you found in problem 23. The wavefunction is zero everywhere except between -10Å and $+10\text{Å}$, and between these limits it is a constant. Substitute the constant wavefunction into this integral using limits where $\psi\,(x)$ is non-zero and solve for A to complete part (a). Integrate the probability, $P = |\,\psi\,(x)\,|\, dx = \psi^*\,(x)\,\psi\,(x)\,dx$, between the given limits for part (b). It is same integration as part (a) except for the limits.

Write down the interval where the wavefunction is non-zero and you are done with part (c). This, however, is deeper than it might appear. The eigenvalues of the position operator are positions, x's in one dimension. Any possible x_i is an eigenvalue in the same sense that an α_i is an eigenvalue of a matrix operator and the result of a measurement can be only an individual eigenvalue per postulate 3. There are two aspects to the physics of this situation. First, any measurement of position is limited by the quality of your equipment. A position measured to be 3.5 Å means only that you have measured to 3.5 Å $\pm\, 0.05$ Å. This is an interval. Secondly, the probability of measuring any given eigenvalue x_i is zero because there are an infinite number of points on any continuous curve. It is like asking what is one divided by infinity? The mathematics answers this question since $\displaystyle\int_a^a f\,(x)\,dx = 0$, any integral with the same upper and lower limit is zero. That said, assume that you had ideal instruments that could measure individual eigenvalues (which is both mathematically and physically impossible, but assume it nevertheless). Assume that you have ideal instruments for part (d) also. Use postulate 5. Only ideal instruments could measure a delta function. If you measure $x = 3.5$ Å, then the state vector is the eigenvector of the position operator with eigenvalue $3.5\mathring{A}$, which is a delta function centered at $x = 3.5$ Å. For part (e), you must calculate the Fourier transform

$$\widehat{\psi}\,(p) = \frac{1}{\sqrt{2\pi\,\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar}\, \psi\,(x)\, dx = \frac{1}{\sqrt{2\pi\,\hbar}} \int_{-10\mathring{A}}^{+10\mathring{A}} e^{-ipx/\hbar}\, \psi\,(x)\, dx\,,$$

where the integration is over only the region where the wavefunction is non-zero. The sinc function

$$\widehat{\psi}\,(p) = \frac{1}{\sqrt{10\pi\hbar\,\text{Å}}} \frac{\sin\left(\frac{p}{\hbar}\,(10\text{Å})\right)}{p/\hbar}\,,$$

is the result. The eigenvalues of the momentum operator are momenta for part (f). These are eigenvalues in the same sense as the eigenvalues of a matrix operator per postulate 3 so the possible results are momenta between $-\infty$ and $+\infty$. In contrast to part (c), all momenta are possible results of a measurement. Notice also that the **probability density**, $|\,\widehat{\psi}\,(p)\,|^2$ diminishes as you go away from $p = 0$. Part (g) is very similar to part (d). If you could measure an individual eigenvalue (again you really cannot), you would get a delta function at $p = 0\ \ldots$ so sketch that.

(a) Calculate the normalization constant

$$\int_{-\infty}^{\infty} \psi^*(x)\,\psi(x)\,dx = 1 \quad\Rightarrow\quad \int_{-10\,\text{Å}}^{+10\,\text{Å}} \mathrm{A}^*\mathrm{A}\,dx = 1 \quad\Rightarrow\quad \int_{-10\,\text{Å}}^{+10\,\text{Å}} |\,\mathrm{A}\,|^2\,dx = 1$$

$$\Rightarrow\quad |\,\mathrm{A}\,|^2 \int_{-10\,\text{Å}}^{+10\,\text{Å}} dx = 1 \quad\Rightarrow\quad |\,\mathrm{A}\,|^2\, x\Big|_{-10\,\text{Å}}^{+10\,\text{Å}} = 1 \quad\Rightarrow |\,\mathrm{A}\,|^2\big(+10\,\text{Å} - (-10\,\text{Å})\big) = 1$$

$$\Rightarrow\quad 20\,\text{Å}\,|\,\mathrm{A}\,|^2 = 1 \quad\Rightarrow\quad |\,\mathrm{A}\,|^2 = \frac{1}{20\,\text{Å}} \quad\Rightarrow\quad \mathrm{A} = \frac{1}{\sqrt{20\,\text{Å}}}\,.$$

The normalization constant will always have units of $1/\sqrt{\text{length}}$ in one dimension in position space. A popular procedure is to minimize clutter by excluding units during calculation, and then carefully placing the units back into the final answer.

(b) Calculate the probability

$$P\,(-3\text{Å} < x < 5\text{Å}) = \int_{-3}^{5} \psi^*(x)\,\psi(x)\,dx = \int_{-3}^{5} |\,\mathrm{A}\,|^2\,dx$$

$$= \frac{1}{20\,\text{Å}}\; x\Big|_{-3\,\text{Å}}^{+5\,\text{Å}} = \frac{1}{20\,\text{Å}}\Big(+5\,\text{Å} - -3\,\text{Å}\Big) = \frac{8}{20} = \frac{2}{5} = 0.4\,.$$

(c) The possible results of a position measurement are any x in the interval $-10\,\text{Å} \le x \le 10\,\text{Å}$.

(d) If we measure the position, and we find the particle at $x = 3.5\,\text{Å}$, then the wavefunction is a delta function centered at $x = 3.5\,\text{Å}$, which is sketched at the right.

(e) The wavefunction in momentum space is given by the Fourier transform

$$\widehat{\psi}\,(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar}\,\psi(x)\,dx = \frac{1}{\sqrt{2\pi\hbar}} \int_{-10}^{10} e^{-ipx/\hbar}\,\mathrm{A}\,dx$$

$$= \frac{\mathrm{A}}{\sqrt{2\pi\hbar}} \int_{-10}^{10} e^{-ipx/\hbar}\,dx = \frac{\mathrm{A}}{\sqrt{2\pi\hbar}}\,\frac{\hbar}{(-ip)}\Big(e^{-ip10/\hbar} - e^{+ip10/\hbar}\Big)$$

$$= \frac{\mathrm{A}\sqrt{2}}{\sqrt{\pi\hbar}}\,\frac{\hbar}{p}\left(\frac{e^{+ip10/\hbar} - e^{-ip10/\hbar}}{2i}\right) = \frac{\sqrt{2}}{\sqrt{20\,\text{Å}}}\,\frac{1}{\sqrt{\pi\hbar}}\,\frac{\hbar}{p}\,\sin\left(\frac{p}{\hbar}10\,\text{Å}\right)$$

$$= \frac{1}{\sqrt{10\pi\hbar\,\text{Å}}}\,\frac{\sin\left(\frac{p}{\hbar}\,(10\,\text{Å})\right)}{p/\hbar}.$$

(f) The possible results of a momentum measurement are any p in the interval $-\infty \le p \le \infty$. The probability density in momentum space is $|\,\widehat{\psi}\,(p)\,|^2$, and

$$P\,(p_a < p < p_b) = \frac{1}{10\pi\hbar\,\text{Å}} \int_{p_a}^{p_b} \frac{\sin^2\left(\frac{p}{\hbar}\,(10\,\text{Å})\right)}{p^2/\hbar^2}\,dp$$

is a relation to calculate probability.

(g) If you find $p = 0$, the momentum wavefunction will be a delta function located at $p = 0$.

Postscript: The probability of measuring any individual x_i in parts (c) or (d) is zero. The probability of attaining a value between the limits a and b is $P(a < x < b) = \frac{1}{20\,\text{Å}} \int_a^b dx$, for this problem where a and b are in angstroms. The region of interest can be very small, for instance if $a = 3.499999\,\text{Å}$ and $b = 3.500001\,\text{Å}$. There is a small probability of finding the particle between these limits, and this is realistically what is meant by a measurement of $x = 3.5\,\text{Å}$.

---

29. For $\psi(x) = \mathrm{A}\, e^{-x^2/4}$,

(a) normalize $\psi(x)$,

(b) find $P(-\infty < x < 0)$,

(c) find $P(0 < x < 1)$, and

(d) find $P(1 < x < \infty)$.

The given wavefunction is **Gaussian** because the exponent of Euler's number contains only the negative of the square of the independent variable and the "A" and 4 are constants. The Gaussian wavefunction or "**wave packet**" is dominantly the most popular description of the localized probability density that is associated with a particle. It is a primary topic of chapter three. Many realistic problems, including part (c) of this one, do not have a closed form solution. The solution to part (c) may remind you of a solution to a statistics problem using a normal distribution. A normal distribution is the same thing as a Gaussian distribution.

Refer to problem 20 of part 3 of chapter 1 to normalize $\psi(x)$. Part (b) is an application of form 15.3.29 from the *Handbook of Mathematical Formulas and Integrals* by Jeffrey,

$$\int_0^\infty e^{-q^2t^2}\, dt = \frac{\sqrt{\pi}}{2q}, \quad q > 0\,.$$

You need to recognize that the given $\psi(x)$ is an even function to use this integral. Part (c) requires

$$\int_0^b e^{-at^2}\, dt = \frac{1}{2}\sqrt{\frac{\pi}{a}}\ \text{erf}\left(b\sqrt{a}\,\right), \quad \text{where} \quad \text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\, dt\,,$$

and is known as the **error function**. The last integral is solved numerically and is used frequently enough that it is tabulated in standard tables, for instance,

x	0.5	$1/\sqrt{2}$	1.0	$\sqrt{2}$	1.5	2.0	2.5	3.0
erf(x)	0.5205	0.6827	0.8427	0.9546	0.9331	0.9772	0.9937	0.9986

is an abbreviated table. There are related functions such as the **normal probability function**. Values for the error function can be derived from tables of the normal probability function. Ensure that you read the definitions of the functions before you start using tabulated data because different authors employ different conventions and different notation. Keep good track of your constants.

(a) Problem 20 of part 3 of chapter 1 found the normalization constant of

$$f(x) = \mathrm{A}\, e^{-\mathrm{B}x^2} \quad \text{to be} \quad \mathrm{A} = \left(\frac{2\mathrm{B}}{\pi}\right)^{1/4}. \qquad \text{For} \quad \psi(x) = \mathrm{A}\, e^{-x^2/4}, \quad \text{we see that} \quad \mathrm{B} = \frac{1}{4}$$

$$\Rightarrow \quad \psi(x) = \left(\frac{1}{2\pi}\right)^{1/4} e^{-x^2/4} \quad \text{is the normalized wavefunction.}$$

(b) Probability for the given wavefunction is

$$P(a < x < b) = \int_a^b \psi^*(x)\,\psi(x)\,dx = \int_a^b \left(\frac{1}{2\pi}\right)^{1/4} e^{-x^2/4} \left(\frac{1}{2\pi}\right)^{1/4} e^{-x^2/4}\,dx = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\,dx$$

$$\Rightarrow \quad P(-\infty < x < 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-x^2/2}\,dx$$

because $\psi(x)$ is an even function. Then

$$\int_0^\infty e^{-q^2t^2}\,dt = \frac{\sqrt{\pi}}{2q} \quad \Rightarrow \quad \psi(x) = \frac{1}{\sqrt{2\pi}}\,\frac{\sqrt{2\pi}}{2} = \frac{1}{2} \quad \text{since} \quad q = \frac{1}{\sqrt{2}} \quad \text{for our integral.}$$

(c)
$$P(0 < x < 1) = \frac{1}{\sqrt{2\pi}} \int_0^1 e^{-x^2/2}\,dx,$$

$$\int_0^b e^{-at^2}\,dt = \frac{1}{2}\sqrt{\frac{\pi}{a}}\,\mathrm{erf}\left(b\sqrt{a}\,\right) \quad \text{and} \quad a = \frac{1}{2}, \quad b = 1, \quad \text{for this integral,}$$

$$\Rightarrow \quad P(0 < x < 1) = \frac{1}{\sqrt{2\pi}} \int_0^1 e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}}\,\frac{1}{2}\sqrt{\frac{2\pi}{1}}\,\mathrm{erf}\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{2}(0.6827) = 0.3414\,.$$

(d) Since $P(-\infty < x < \infty) = 1$,

$$P(1 < x < \infty) = 1 - P(-\infty < x < 0) - P(0 < x < 1) = 1 - 0.5 - 0.3414 = 0.1586\,.$$

Postscript: You may recognize the answers to part (c) to be the portion of a normal distribution that is one standard deviation greater than the mean and part (d) to be the percentage greater than one standard deviation above the mean. A general normal or Gaussian distribution is

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\,\exp\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right), \quad \text{where } \mu \text{ is the mean and } \sigma \text{ is the standard deviation.}$$

The probability distribution for $\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/4}$ is a Gaussian distribution with a mean of zero and a standard deviation of 1.

---

30. Consider the spin-1 angular momentum operators:

$$\mathcal{L}_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathcal{L}_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \text{and} \quad \mathcal{L}_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(a) What are the possible values one can obtain if $\mathcal{L}_z$ is measured?

(b) Consider the state for which $\mathcal{L}_z = 1$. What are $<\mathcal{L}_x>$, $<\mathcal{L}_x^2>$, and $\Delta\mathcal{L}_x$ in this state?

(c) Find the eigenvalues and eigenvectors of the $\mathcal{L}_x$ operator in the $\mathcal{L}_z$ basis.

(d) If the particle is in the state with $\mathcal{L}_z = -1$ and $\mathcal{L}_x$ is measured, what are the possible results of the measurement and with what probabilities will they be obtained?

(e) Consider the initial state $|\psi> = \frac{1}{\sqrt{21}} \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}$ in the $\mathcal{L}_z$ basis. If $\mathcal{L}_z^2$ is measured and the result $+1$ is obtained, what is the state immediately after the measurement? How probable was this result? What are the possible outcomes and their respective probabilities for a measurement of $\mathcal{L}_z$ immediately after attaining a result of 1 for a measurement of $\mathcal{L}_z^2$?

(f) Find the state vector for which the $\mathcal{L}_z$ probabilities are

$$P(\mathcal{L}_z = 1) = \frac{1}{4}, \quad P(\mathcal{L}_z = 0) = \frac{1}{2}, \quad \text{and} \quad P(\mathcal{L}_z = -1) = \frac{1}{4}.$$

The given matrices are the spin-1 angular momentum operators (less a factor of $\hbar$). This problem highlights the postulates of quantum mechanics using these commonly encountered operators. Parts (c) and (e) of this problem introduce a consequence of operators that do not commute.

Notice first that postulate 2 is satisfied because all three operators are Hermitian. Part (a) is an application of postulate 3. Notice that it concerns only $\mathcal{L}_z$. If an eigenvalue of $\mathcal{L}_z$ has been measured, you can ascertain the system's eigenstate from postulate 5 for part (b). The calculation of the expectation values $<\mathcal{L}_z = 1 \,|\, \mathcal{L}_x \,|\, \mathcal{L}_z = 1>$, $<\mathcal{L}_z = 1 \,|\, \mathcal{L}_x \, \mathcal{L}_x \,|\, \mathcal{L}_z = 1>$, and

$$\sqrt{<\mathcal{L}_z = 1 \,|\, (\mathcal{L}_x \, \mathcal{L}_x - <\mathcal{L}_x> \mathcal{I})^2 \,|\, \mathcal{L}_z = 1>} \quad \text{are extensions of postulate 4.}$$

None of the three spin-1 angular momentum operators commute. You can diagonalize any one of them (see problems 19 and 24 of part 2 of chapter 1) and express the other two in that basis. $\mathcal{L}_z$ is the usual choice of the three operators to be expressed as a diagonal matrix. The given $\mathcal{L}_z$ is diagonal which means that the operators are given in the $\mathcal{L}_z$ basis. Part (c) directs you only to solve the eigenvalue/eigenvector problem for the given $\mathcal{L}_x$ for which problem 24 of part 2 of chapter 1 may be of interest. Postulate 5 provides the eigenstate corresponding to a measurement of $\mathcal{L}_z = -1$ for part (d). Postulate 3 and the results of part (c) provide the possible results of a measurement of $\mathcal{L}_x$. Postulate 4 tells you how to calculate probabilities for each possibility.

Construct the operator $\mathcal{L}_z^2$ to start part (e). You will find that it is two-fold degenerate at $\mathcal{L}_z^2 = 1$, therefore, the state vector following the measurement is a superposition of the two possible eigenstates per postulate 1. The projection operator $\mathcal{P} = |\,\alpha\!><\!\alpha\,| + |\,\beta\!><\!\beta\,|$, where $|\,\alpha\!>$ and $|\,\beta\!>$ represent the two $\mathcal{L}_z^2 = 1$ eigenstates, operating on the given state vector, $\mathcal{P}\,|\,\psi\!>$, provides the state vector that is the superposition. Normalize this state vector so that an equality may be used in applying postulate 4 to attain probabilities. You have the eigenvectors and the probabilities but the state vector is unknown for part (f), for instance $|\,\psi\!> = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$. Use postulate 4 to solve for each component.

(a) The eigenvalues are on the principal diagonal in the diagonal matrix $\mathcal{L}_z$ so are 1, 0, and -1.

(b) Measurement forces the system into the eigenstate corresponding to the measured eigenvalue. Therefore $\mathcal{L}_z = 1 \;\Rightarrow\; |\,\psi\!> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. The expectation values and the uncertainty are

$$< \mathcal{L}_x >_\psi = \big(1,\ 0,\ 0\big)\,\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\big(1,\ 0,\ 0\big)\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\,(0) = 0\,.$$

$$\mathcal{L}_x^2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$\Rightarrow\quad < \mathcal{L}_x^2 >_\psi = \big(1,\ 0,\ 0\big)\,\frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\big(1,\ 0,\ 0\big)\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2}\,,$$

$$\Delta\mathcal{L}_x = <\big(\,\mathcal{L}_x - <\mathcal{L}_x>\big)^2>^{1/2} = <\big(\,\mathcal{L}_x - 0\big)^2>^{1/2} = <\mathcal{L}_x^2>^{1/2} = <\frac{1}{2}>^{1/2} = \frac{1}{\sqrt{2}}\,.$$

(c) This is the standard eigenvalue/eigenvector problem for $\mathcal{L}_x$ because $\mathcal{L}_z$ is diagonal meaning that the three operators are given in the $\mathcal{L}_z$ basis. Referring to problem 24 of part 2 of chapter 1,

$$\mathcal{U}^\dagger\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\mathcal{U} = \begin{pmatrix} -\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

$$\Rightarrow\quad \mathcal{U}^\dagger\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\mathcal{U} = \frac{1}{\sqrt{2}}\begin{pmatrix} -\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

so $-1,\ 0$ and 1 are the eigenvalues. The eigenvectors in the original $\mathcal{L}_z$ basis were found to be

$$|\,-1\!> = \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}, \quad |\,0\!> = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \text{and} \quad |\,1\!> = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}.$$

(d) The measurement $\mathcal{L}_z = -1 \quad \Rightarrow \quad |\,\psi\!> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ per postulate 5. Then per postulate 4,

$$P\,(\mathcal{L}_x = -1) = \Big|\,<\!\mathcal{L}_x = -1\,|\,\psi\!>\Big|^2 = \left|\frac{1}{2}\big(1,\ -\sqrt{2},\ 1\big)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right|^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}\,,$$

$$P\,(\mathcal{L}_x = 0) = \Big|\,<\!\mathcal{L}_x = 0\,|\,\psi\!>\Big|^2 = \left|\frac{1}{\sqrt{2}}\big(1,\ 0,\ -1\big)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right|^2 = \left(\frac{-1}{\sqrt{2}}\right)^2 = \frac{1}{2}\,,$$

$$P\,(\mathcal{L}_x = 1) = \Big|\,<\!\mathcal{L}_x = 1\,|\,\psi\!>\Big|^2 = \left|\frac{1}{2}\big(1,\ \sqrt{2},\ 1\big)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right|^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}\,.$$

(e) The operator $\mathcal{L}_z^2$ is

$$\mathcal{L}_z^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\,,$$

and the eigenvalues of $1,\ 0,$ and 1 are on the principal diagonal. A measurement of $\mathcal{L}_z^2 = +1$ forces the system into a superposition of *both* eigenvectors that have eigenvalue 1, meaning

$$|\,\psi\!> = \alpha\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \\ \beta \end{pmatrix}.$$

The projection operator constructed from the states existing in the superposition is

$$\mathcal{P} = |\,1\!><\!1\,| + |\,1'\!><\!1'\,| = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

So the unnormalized state vector immediately after a measurement that yields $\mathcal{L}_z^2 = 1$ is

$$\mathcal{P}\,|\,\psi\!> = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix}.$$

The new state vector will have to be normalized if eigenstates are missing from the original state vector so the original normalization constant is an unnecessary complication and is excluded. Then

$$\big(4,\ 0,\ 2\big)\,\mathrm{A}^*\,\mathrm{A}\begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix} = |\,\mathrm{A}\,|^2\,(16+4) = |\,\mathrm{A}\,|^2\,(20) = 1 \quad \Rightarrow \quad \mathrm{A} = \frac{1}{2\sqrt{5}}$$

$$\Rightarrow \quad |\,\psi\!> = \frac{1}{2\sqrt{5}}\begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix} \quad \text{is normalized, and} \quad \alpha = \frac{2}{\sqrt{5}}\,, \quad \beta = \frac{1}{\sqrt{5}}\,, \quad \text{if you want}$$

to resolve the state into individual vectors. To find the probability of measuring $\mathcal{L}_z^2 = +1$, add the probabilities corresponding to each of the ways that we can obtain $\mathcal{L}_z^2 = 1$,

$$P\left(\mathcal{L}_z^2 = +1\right) = \left| <1\,|\,\psi> \right|^2 + \left| <1'\,|\,\psi> \right|^2$$
$$= \left| (1,\ 0,\ 0)\frac{1}{\sqrt{21}}\begin{pmatrix}4\\1\\2\end{pmatrix}\right|^2 + \left|(0,\ 0,\ 1)\frac{1}{\sqrt{21}}\begin{pmatrix}4\\1\\2\end{pmatrix}\right|^2$$
$$= \left|\frac{4}{\sqrt{21}}\right|^2 + \left|\frac{2}{\sqrt{21}}\right|^2 = \frac{16}{21} + \frac{4}{21} = \frac{20}{21}.$$

If $\mathcal{L}_z$ is measured for the state vector prepared by measuring $\mathcal{L}_z^2 = 1$, the possible outcomes are 1 or -1 because the new state vector is a linear combination of the $|\,\mathcal{L}_z = +1>$ and $|\,\mathcal{L}_z = -1>$ eigenvectors. The $|\,\mathcal{L}_z = 0>$ eigenvector is absent, so there is no possibility of measuring $\mathcal{L}_z = 0$. The probabilities of measurements of $\mathcal{L}_z$ following a measurement of $\mathcal{L}_z^2 = 1$ are

$$P\left(\mathcal{L}_z = 1\right) = \left| <\mathcal{L}_z = 1\,|\,\psi> \right|^2 = \left|(1,\ 0,\ 0)\,\frac{1}{2\sqrt{5}}\begin{pmatrix}4\\0\\2\end{pmatrix}\right|^2 = \left|\frac{4}{2\sqrt{5}}\right|^2 = \frac{4}{5},$$
$$P\left(\mathcal{L}_z = -1\right) = \left| <\mathcal{L}_z = -1\,|\,\psi> \right|^2 = \left|(0,\ 0,\ 1)\,\frac{1}{2\sqrt{5}}\begin{pmatrix}4\\0\\2\end{pmatrix}\right|^2 = \left|\frac{2}{2\sqrt{5}}\right|^2 = \frac{1}{5}.$$

(f) The eigenvectors are the unit vectors in the $\mathcal{L}_z$ basis and the state vector is unknown, so

$$P\left(\mathcal{L}_z = 1\right) = \left|\ <\mathcal{L}_z = 1\,|\,\psi>\ \right|^2 = \left|(1,\ 0,\ 0)\ \begin{pmatrix}\alpha\\\beta\\\gamma\end{pmatrix}\right|^2 = \left|\,\alpha\,\right|^2 = \frac{1}{4} \quad\Rightarrow\quad \alpha = \frac{1}{2},$$
$$P\left(\mathcal{L}_z = 0\right) = \left|\ <\mathcal{L}_z = 0\,|\,\psi>\ \right|^2 = \left|(0,\ 1,\ 0)\ \begin{pmatrix}\alpha\\\beta\\\gamma\end{pmatrix}\right|^2 = \left|\,\beta\,\right|^2 = \frac{1}{2} \quad\Rightarrow\quad \beta = \frac{1}{\sqrt{2}},$$
$$P\left(\mathcal{L}_z = -1\right) = \left|\ <\mathcal{L}_z = -1\,|\,\psi>\ \right|^2 = \left|(0,\ 0,\ 1)\ \begin{pmatrix}\alpha\\\beta\\\gamma\end{pmatrix}\right|^2 = \left|\,\gamma\,\right|^2 = \frac{1}{4} \quad\Rightarrow\quad \gamma = \frac{1}{2},$$

and the state vector is $\quad |\,\psi> = \begin{pmatrix}1/2\\1/\sqrt{2}\\1/2\end{pmatrix}$ which is normalized.

---

31. Use the normalized state vector of part (f) of problem 30 to demonstrate that components of a state vector having different phases lead to variable probabilities.

Phase is not a focal issue because results are independent of multiplication by the same $e^{i\phi}$. It can, however, be an important consideration. For instance, if the components of a state vector differ in phase, the probabilities are generally phase dependent.

A general form of the phase-dependent state vector from part (f) of the last problem is

$$| \psi> = \frac{e^{i\delta_1}}{2} \, | \mathcal{L}_z = 1> + \frac{e^{i\delta_2}}{\sqrt{2}} \, | \mathcal{L}_z = 0> + \frac{e^{i\delta_3}}{2} \, | \mathcal{L}_z = -1> = \begin{pmatrix} \frac{1}{2} \, e^{i\delta_1} \\ \frac{1}{\sqrt{2}} \, e^{i\delta_2} \\ \frac{1}{2} \, e^{i\delta_3} \end{pmatrix}$$

Show that the state vector with components having different phases is still normalized. Then calculate the probability of measuring $\mathcal{L}_x = 0$, for instance. You will get non-vanishing exponential terms which depend on the difference between the phases.

The phase-dependent form of the state vector is normalized because

$$\Big| <\psi \, | \, \psi> \Big|^2 = \left| \left(\frac{e^{-i\delta_1}}{2}, \; \frac{e^{-i\delta_2}}{\sqrt{2}}, \; \frac{e^{-i\delta_3}}{2} \right) \begin{pmatrix} \frac{1}{2} \, e^{i\delta_1} \\ \frac{1}{\sqrt{2}} \, e^{i\delta_2} \\ \frac{1}{2} \, e^{i\delta_3} \end{pmatrix} \right|^2 = \left| \frac{1}{4} + \frac{1}{2} + \frac{1}{4} \right|^2 = 1 \, ,$$

so component phase dependence is not consequential to normalization. However,

$$\begin{aligned} P \left(\mathcal{L}_x = 0 \right) &= \Big| <\mathcal{L}_x = 0 \, | \, \psi> \Big|^2 = \left| \frac{1}{\sqrt{2}} \big(1, \; 0, \; -1 \big) \begin{pmatrix} \frac{1}{2} \, e^{i\delta_1} \\ \frac{1}{\sqrt{2}} \, e^{i\delta_2} \\ \frac{1}{2} \, e^{i\delta_3} \end{pmatrix} \right|^2 \\ &= \left| \frac{1}{\sqrt{2}} \left(\frac{e^{i\delta_1}}{2} - \frac{e^{i\delta_3}}{2} \right) \right|^2 = \frac{1}{2} \left(\frac{e^{-i\delta_1}}{2} - \frac{e^{-i\delta_3}}{2} \right) \left(\frac{e^{i\delta_1}}{2} - \frac{e^{i\delta_3}}{2} \right) \\ &= \frac{1}{8} \, \left(e^{-i\delta_1 + i\delta_1} - e^{-i\delta_1 + i\delta_3} - e^{-i\delta_3 + i\delta_1} + e^{-i\delta_3 + i\delta_3} \right) \\ &= \frac{1}{8} \left(2 - e^{i(\delta_1 - \delta_3)} - e^{-i(\delta_1 - \delta_3)} \right) = \frac{1}{4} \left(\frac{2}{2} - \frac{e^{i(\delta_1 - \delta_3)} + e^{-i(\delta_1 - \delta_3)}}{2} \right) \\ &= \frac{1}{4} \Big(1 - \cos(\delta_1 - \delta_3) \Big) \, . \end{aligned}$$

The probability of measuring $\mathcal{L}_x = 0$ varies from a maximum of $1/2$ to a minimum of 0 as a function of $\delta_1 - \delta_3$. A relative phase difference between the components of the state vector is, therefore, physically significant.

Postscript: We expect all of the probabilities for any of the eigenvalues of $\mathcal{L}_x$ and $\mathcal{L}_y$ to be phase dependent. This is because the spin-1 angular momentum operators do not commute. In fact, we expect this sort of behavior whenever the operators under consideration do not commute. It is a consequence of multiple operators describing the same system that do not commute. Picture a vector representing angular momentum in three dimensions. In establishing the three spin-1 angular momentum operators in the $\mathcal{L}_z$ basis, we have effectively fixed the z-component of angular momentum but since the operators do not commute, the projection of the angular momentum vector onto the xy plane varies as the angular momentum vector precesses around the z-axis. We will address this "semi-classical" picture when discussing angular momentum.

Practice Problems

32. (a) What are the dimensions of $\psi(x, y)$?

(b) What are the dimensions of $\psi(x, y, z)$?

(c) What are the dimensions of $\psi(r, \theta, \phi)$?

(d) What are the dimensions of $\widehat{\psi}(p)$?

(e) What are the dimensions of $\widehat{\psi}(p_x, p_y)$?

(f) What physical quantity does the units of $\hbar$ represent?

Dimensional consistency is a valuable tool in physics and the field of quantum mechanics is no exception. It is helpful to have a simple procedure to derive the dimensions of a state vector or wavefunction. A general statement of probability useful for extension to larger dimension is

$$P = \int_V |\psi|^2 \, dV$$

so probability is $|\psi(x, y)|^2 \, dx \, dy$ in two dimensions, and $|\psi(x, y, z)|^2 \, dx \, dy \, dz$ in three dimensions. In spherical coordinates, probability is $|\psi(r, \theta, \phi)|^2 \, r^2 \sin\theta \, dr \, d\theta \, d\phi$, where the additional factors are from the differential volume element. Probability is a dimensionless number, so ψ must have the dimensions of $\left(1/dV\right)^{1/2}$. Planck's constant is the fundamental constant of quantum mechanics. If an expression contains some form of Planck's constant, it is necessarily a quantum mechanical result. Planck's constant has dimensions of energy multiplied by time. What physical quantity does energy multiplied by time represent?

---

33. Consider a system described by the Hamiltonian

$$\mathcal{H} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad \text{which at} \quad t = 0 \quad \text{is in the state} \quad |\psi(0)> = \frac{1}{\sqrt{17}} \begin{pmatrix} 2 \\ 3 \\ -2 \end{pmatrix}.$$

(a) What results can be obtained for a measurement of energy and with what probabilities will these results be obtained?

(b) Calculate the expectation value of the Hamiltonian $<\mathcal{H}>$ for the state $|\,\psi\,(0)>$ using $<\mathcal{H}> = <\psi\,(0)\,|\,\mathcal{H}\,|\,\psi\,(0)>$. Show that your expectation value agrees with your calculations from part (a) using $<\mathcal{H}> = \sum_i P\,(E_i)\,E_i$.

(c) Expand the initial state vector $|\,\psi\,(0)>$ in the energy eigenbasis to calculate the time dependent state vector $|\,\psi\,(t)>$.

(d) If the energy is measured at time t what results can be obtained and with what probabilities will these results be obtained? Compare your probabilities with the $t=0$ case in part (a).

(e) Suppose that you measure the energy of the system at $t=0$ and you find $E=-1$. What is the state vector of the system immediately after your measurement? Now let the system evolve without any additional measurements until $t=10$. What is the state vector $|\,\psi\,(10)>$ at $t=10$? What energies will you measure if you repeat the energy measurement at $t=10$?

See problem 19.

---

34. Consider a system described by the Hamiltonian

$$\mathcal{H} = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}, \quad \text{which at} \quad t=0 \quad \text{is in the state} \quad |\,\psi\,(0)> = \frac{1}{\sqrt{29}}\begin{pmatrix} 2 \\ 5 \end{pmatrix}.$$

(a) Is $\mathcal{H}$ Hermitian?

(b) Attain the eigenvalues and eigenvectors of $\mathcal{H}$ by solving the eigenvalue/eigenvector problem.

(c) If the energy is measured what results can be obtained and with what probabilities will these results be obtained?

(d) Calculate the expectation value of the Hamiltonian using both

$$<\mathcal{H}>_{\psi\,(0)} = <\psi\,(0)\,|\,\mathcal{H}\,|\,\psi\,(0)> \quad \text{and} \quad <\mathcal{H}> = \sum_i P\,(E_i)\,E_i\,.$$

(e) Expand the initial state vector $|\,\psi\,(0)>$ in the energy eigenbasis and use the time evolution of the energy eigenvectors, *i.e.*, $|\,E_i\,(t)> = e^{-iE_i t/\hbar}\,|\,E_i\,(0)>$, to calculate the time dependent state vector $|\,\psi\,(t)>$.

(f) If the energy is measured at time t, what results can be obtained and with what probabilities will these results be obtained? Compare your answer with the $t=0$ case in part (c).

(g) Diagonalize $\mathcal{H}$ using a unitary transformation. Transform $|\,\psi\,(0)>$ and both eigenvectors to be consistent with this unitary transformation.

(h) Calculate $|\,\psi\,(t)>$ and both $t>0$ probabilities in the basis in which $\mathcal{H}$ is diagonal.

(i) Suppose that you measure the energy of the system at $t=0$ and you find $E=2$. Find the state vector of the system immediately after your measurement and at time $t=10$ in both the basis in which $\mathcal{H}$ is diagonal and in the basis of part (b). What energies will you measure if you repeat the energy measurement at $t=10$?

See problem 20.

---

35. Given the arbitrary diagonal operator and the arbitrary state vector

$$\Omega = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \quad \text{and} \quad |\,\psi\!> = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \text{show that} \quad <\psi\,|\,\Omega\,|\,\psi> = \sum_i P\,(\omega_i)\,\omega_i\,.$$

This problem reinforces problem 9. The two methods of calculating expectation values yield identical results. First, normalize the state vector to find that $1/\sqrt{|\,x\,|^2 + |\,y\,|^2 + |\,z\,|^2}$ is the normalization constant. Then calculate $<\psi\,|\,\Omega\,|\,\psi>$. Calculate the probabilities for each eigenvalue using the unit vectors appropriate to a diagonal operator. You then have the probabilities and the eigenvalues to calculate $\sum_i P\,(\omega_i)\,\omega_i$. This problem is relatively simple in the basis of a diagonal operator. Think about how you would do this in the basis of an arbitrary operator… which may make you appreciate more fully the utility of diagonal operators.

---

36. Consider the operator

$$\Omega = \begin{pmatrix} 0 & 4i & 0 \\ -4i & 0 & -3i \\ 0 & 3i & 0 \end{pmatrix} \quad \text{and the state vector} \quad |\,\psi\!> = \begin{pmatrix} 1 \\ i \\ -1 \end{pmatrix}.$$

(a) Verify that the eigenvalues and eigenvectors of Ω are

$$|\,\omega = -5\!> = \frac{4}{5\sqrt{2}}\begin{pmatrix} 1 \\ i5/4 \\ 3/4 \end{pmatrix}, \quad |\,\omega = 0\!> = \frac{3}{5}\begin{pmatrix} 1 \\ 0 \\ -4/3 \end{pmatrix}, \quad \text{and} \quad |\,\omega = 5\!> = \frac{4}{5\sqrt{2}}\begin{pmatrix} 1 \\ -i5/4 \\ 3/4 \end{pmatrix}.$$

Using Ω, $|\,\psi>$, and the given eigenvalues and eigenvectors,

(b) find the probabilities of each possibility,

(c) find the expectation value using both $<\psi\,|\,\Omega\,|\,\psi>$ and $\sum_i P\,(\omega_i)\,\omega_i$,

(c) calculate the uncertainty using $\Delta\Omega_\psi = <\psi\,|\,\Big(\Omega - <\Omega>\mathcal{I}\Big)^2\,|\,\psi>^{1/2}$,

(e) and calculate the uncertainty using $\Delta\Omega_\psi = <\psi\,|\,\Omega^2 - <\Omega>^2\,\mathcal{I}\,|\,\psi>^{1/2}$. Compare this answer with your result from part (c).

Uncertainty is the primary point of this problem. What is called "standard deviation" in many fields is called uncertainty in quantum mechanics. This is uncertainty in the same sense of the word used in the term "Heisenberg uncertainty principle." Also, you can never leave the possibilities and probabilities behind so this problem is practice in a three dimensional case where selected elements and components are imaginary.

First, normalize the state vector. Remember to conjugate components when forming bra's for probabilities. You have already attained the eigenvalues and eigenvectors if you completed problem 34 of part 2 of chapter 1. You also diagonalized Ω in that problem. Using the diagonal basis will be the easier path and we address that method in the next problem. Your probabilities, $P(\omega_i) = |<\omega_i|\psi>|^2$, must sum to 1 and the expectation values must be the same using both methods. You do not have to carry the square root in your calculations if you compute variance and then take a square root to attain uncertainty. We get $\Delta\Omega = \sqrt{74}/3$.

---

37. The diagonal form of the Ω operator used in the last problem is $\Omega = \begin{pmatrix} -5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{pmatrix}$.

(a) Transform $|\psi> = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ i \\ -1 \end{pmatrix}$ used in the previous problem to the diagonal basis.

(b) Calculate the probability of each possibility of a measurement of Ω.

(c) Calculate the expectation value using both $<\psi|\Omega|\psi>$ and $\sum_i P(\omega_i)\,\omega_i$.

(d) Calculate the uncertainty using $\Delta\Omega_\psi = <\psi|\left(\Omega - <\Omega> \mathcal{I}\right)^2|\psi>^{1/2}$.

(e) Calculate the uncertainty using $\Delta\Omega_\psi = <\psi|\Omega^2 - <\Omega>^2\,\mathcal{I}\,|\psi>^{1/2}$. Compare this answer with your result from part (d).

You will likely find the calculations of the last problem easier in the diagonal basis. You need to understand the concepts surrounding diagonal operators to understand the techniques that are employed frequently in realistic problems.

You need to calculate $\mathcal{U}^\dagger|\psi>$ to transform the state vector to the same basis as the diagonal Ω. Calculations yield nonsense if the operator and state vector are not in the same basis. You calculated $\mathcal{U}^\dagger$ using the eigenvectors of in problem 34 of part 2 of chapter 1,

$$\mathcal{U}^\dagger = \begin{pmatrix} 4/5\sqrt{2} & -i/\sqrt{2} & 3/5\sqrt{2} \\ 3/5 & 0 & -4/5 \\ 4/5\sqrt{2} & i/\sqrt{2} & 3/5\sqrt{2} \end{pmatrix}.$$

The probabilities of measuring any eigenvalue, the expectation value, and the uncertainty must be the same as the last problem.

---

38. Find the possible outcomes of a measurement, the probabilities of each possible outcome, the expectation value, and the uncertainty for

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{given that} \quad |\psi> = \begin{pmatrix} 2 \\ i \end{pmatrix}.$$

Support your answers with alternative calculations or consistency checks where possible.

This problem features one of the Pauli spin matrices. This problem is posed without parts because realistic problems do not come with parts though it still leads you down the necessary path. Would you know to travel this path at this point if asked only for uncertainty?

You want to normalize the state vector, solve the eigenvalue/eigenvector problem, and use the eigenvectors and the state vector to attain probabilities. The sum of the probabilities must be 1. Calculate $<\sigma_y> = <\psi\,|\,\sigma_y\,|\,\psi>$ and check this result using $<\sigma_y> = \sum P\,(\sigma_{y_i})\,\sigma_{y_i}$. Calculate the uncertainty using $\Delta\sigma_{y_\psi} = <\psi\,|\,\left(\sigma_y - <\sigma_y> \mathcal{I}\right)^2\,|\,\psi>^{1/2}$ and check the uncertainty using $\Delta\sigma_{y_\psi} = <\psi\,|\,\sigma_y^2 - <\sigma_y>^2\,\mathcal{I}\,|\,\psi>^{1/2}$.

---

39. Consider the operator $\Lambda = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$, which has eigenvalues and normalized eigenvectors

$$|\,\lambda = 2> = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad |\,\lambda = 3> = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}. \quad \text{Given the state vector} \quad |\,\psi> = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

the inner product $<\lambda\,|\,\psi>$ is $9/10$ using the eigenvector that corresponds to 2, and $16/25$ using the eigenvector that corresponds to 3. These should be probabilities but their sum is not 1. There is, in fact, an error present. Can you determine what the error is?

This counter example is meant to highlight an error and thus increase the chance that you will avoid it. If you perform the mathematical mechanics you will find that there is not an error in the given information. Consider postulate 2. Also, are these eigenvectors orthogonal?

---

40. (a) Represent the expectation value of the momentum operator in momentum space.

(b) Represent the expectation value of the position operator in momentum space.

Use procedures similar to those seen in problems 25 and 26. Use the relations

$$<p\,|\,\mathcal{P}\,|\,p'> = p\,\delta\,(p - p') \quad \text{and} \quad <p\,|\,\mathcal{X}\,|\,p'> = i\hbar\,\delta'\,(p - p')$$

discussed in the postscript to problem 24. You should find that

$$<\psi\,|\,\mathcal{P}\,|\,\psi> = \int \widehat{\psi}^*\,(p)\,p\,\widehat{\psi}\,(p)\,dp \quad \Rightarrow \quad \mathcal{P} \;\rightarrow\; p \;\text{ in momentum space, and}$$

$$<\psi\,|\,\mathcal{X}\,|\,\psi> = \int \widehat{\psi}^*\,(p)\left(i\hbar\,\frac{d}{dp}\right)\widehat{\psi}\,(p)\,dp \quad \Rightarrow \quad \mathcal{X} \;\rightarrow\; i\hbar\,\frac{d}{dp} \;\text{ in momentum space.}$$

---

41. For $\psi\,(x) = \mathrm{A}\,x\,,\;\; 0 \le\; x\; \le 1\,,$

(a) Find the normalization constant A,

(b) calculate $P\,(1/2 < x < 3/4)$,

(c) calculate $P\,(0 < x < 1/2)$,

(d) calculate $P\,(3/4 < x < 1)$.

This is a straightforward application of the normalization procedure and probability calculation for a continuous wavefunction that is easy to integrate.

---

42. For $\psi(x) = \mathrm{A}\left(2 - x^2\right), \quad -1 < x < 1\,,$

(a) find the normalization constant,

(b) calculate $P(-1 < x < 0)\,,$

(c) calculate $P(0 < x < 1/2)\,,$

(d) calculate $P(1/2 < x < 1)\,.$

This is another toy problem with the same intent as the last. The sum of the probabilities over all space where the wavefunction is non-zero must be 1 as in problem 41.

---

43. Consider a particle described by the wavefunction $\psi(x) = \dfrac{\mathrm{A}}{x^2 + 1}\,.$

(a) Find the normalization constant A.

(b) Calculate $P(0 < x < 1)\,.$

(c) What are the possible results and the probability of each possible result of a measurement of position?

(d) A measurement of position is made and the particle is found at $x = \pi$. Sketch the wavefunction immediately after the measurement.

(e) Calculate the momentum space wavefunction $\widehat{\psi}(p)$ that corresponds to the position space wavefunction $\psi(x)$ by calculating the Fourier transform of $\psi(x)$. Sketch $\widehat{\psi}(p)$.

(f) The momentum of the particle described by $\psi(x)$ is measured. What results can be obtained, and with what probabilities will they be obtained?

(g) Suppose the momentum is measured and $p = 0$ is obtained. Sketch $\widehat{\psi}(p)$ immediately after the measurement.

Follow the procedures of problem 28. The wavefunction is a Lorentzian curve which is used frequently in spectral analysis. Problem 39 of part 3 of chapter 1 may be of interest. The integrals

$$\int \frac{dx}{(a^2 + b^2x^2)^2} = \frac{x}{2a^2\left(a^2 + b^2x^2\right)} + \frac{1}{2a^3b}\arctan\left(\frac{bx}{a}\right) \quad \text{and} \quad \int_{-\infty}^{\infty} \frac{e^{-iqx}}{x^2 + a^2}\,dx = \frac{\pi}{a}\,e^{-|aq|}$$

will also be of interest. You should find

$$P(0 < x < 1) = \frac{1}{2\pi} + \frac{1}{4} \approx 0.4092 \;\text{ for part (b), and }\; \widehat{\psi}(p) = \frac{e^{-|p|/\hbar}}{\sqrt{\hbar}} \;\text{ for part (e).}$$

---

44. For $\psi(x) = e^{-x^2}\,,$

(a) normalize $\psi(x)$,

(b) find $P(-\infty < x < 0)$,

(c) find $P(0 < x < 1)$, and

(d) find $P(1 < x < \infty)$.

This is also a Gaussian wavefunction so there is little to add to the comments for problem 29. The interpretation of $<\psi\,|\,\psi>$ as a probability is due to Max Born and exists in postulate 4. This intepretation defies explanation in terms of macroscopic experience which is essentially why it is one of the postulates.
