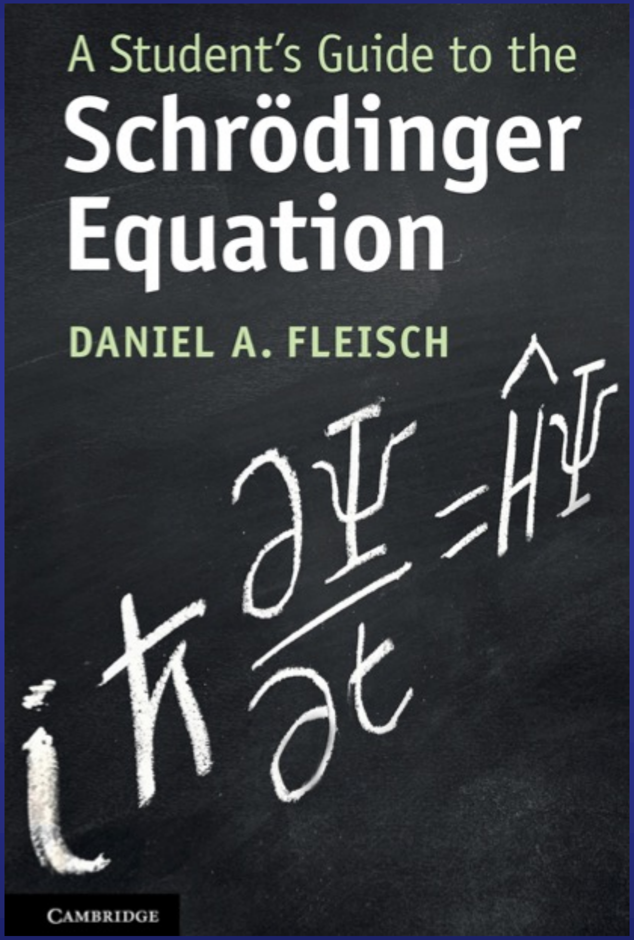
Operators and eigenfunctions

Javier Junquera

UC
UNIVERSIDAD
DE CANTABRIA

CITIMAC

The most important reference followed in this lecture



ISBN 978-1-108-83473-5 Hardback
ISBN 978-1-108-81978-7 Paperback

An EXTREMELY RECOMMENDED textbook for a new comer in the field

What is exactly an operator?

An instruction to perform a certain process on a number, vector or function

Examples of operators:

Operator	Acts on	Action
$\sqrt{}$	On a number	Takes the square root of whatever appears under the roof of the symbol
$\frac{d()}{dx}$	On a function	Takes the first derivative with respect to x of whatever appears inside the parenthesis

Linear operators

Those are the operators that you will find in quantum mechanics

Applying them to a sum of vectors or functions gives the same result as applying to the individual vectors or functions and them summing the results

$$\underset{\substack{\uparrow\\ \text{Operator}}}{\hat{O}}(\underset{\nwarrow}{f_1} + \underset{\nearrow}{f_2}) = \hat{O}(f_1) + \hat{O}(f_2)$$

Operator **Functions**

Multiplying a function by a scalar and then applying the operator gives the same result as first applying the operator and then multiplying the result by a scalar

$$\hat{O}(cf) = c\hat{O}(f)$$

Representing operators as square matrices

Result of multiplying a matrix $\overline{\overline{R}}$ by a column vector $\vec{A}$

$$\bar{\bar{R}}\vec{A} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} R_{11}A_1 + R_{12}A_2 \\ R_{21}A_1 + R_{22}A_2 \end{pmatrix}$$

This type of multiplication can be done only when the number of columns of the matrix equals the number of rows of the vector
(two in this case)

The result of multiplying the matrix by a vector produces another vector
It transforms one vector into another

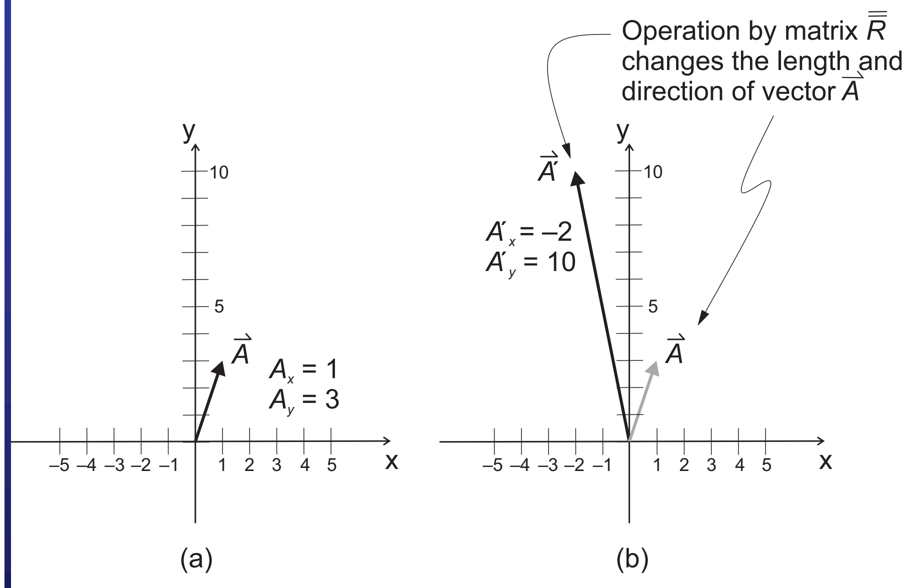
Representing operators as square matrices

The effect of the operation depends on the matrix

$$\overline{\overline{R}} = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}$$

$$\vec{A} = \hat{\imath} + 3\hat{\jmath},$$

$$\overline{\overline{R}}\vec{A} = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} (4)(1) + (-2)(3) \\ (-2)(1) + (4)(3) \end{pmatrix} = \begin{pmatrix} -2 \\ 10 \end{pmatrix}$$



(a) (b)

Figure 2.1 Vector $\vec{A}$ before (a) and after (b) operation of matrix $\overline{\overline{R}}$.

The matrix operating on the vector, generally changes the direction of the vector

The x-component of the new vector is a weighted combination of both components of the original vector.
The weighting coefficients are provided by the first row of the matrix

The same happens for the y-component
The weighting coefficients are provided by 2nd row

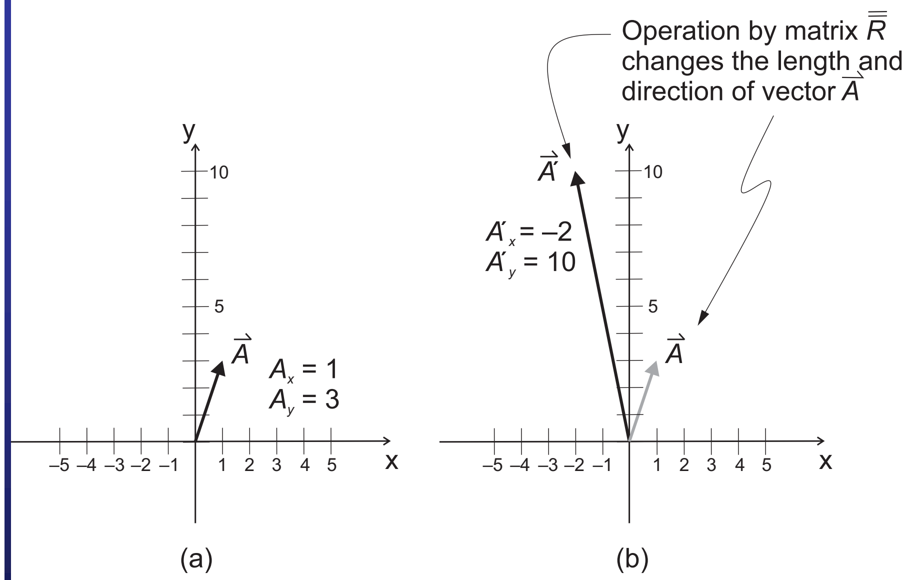
Representing operators as square matrices

The effect of the operation depends on the matrix

$$\bar{\bar{R}} = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}$$

$$\vec{A} = \hat{\imath} + 3\hat{\jmath},$$

$$\bar{\bar{R}}\vec{A} = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} (4)(1) + (-2)(3) \\ (-2)(1) + (4)(3) \end{pmatrix} = \begin{pmatrix} -2 \\ 10 \end{pmatrix}$$


Operation by matrix $\bar{\bar{R}}$ changes the length and direction of vector $\vec{A}$

(a)

(b)

Figure 2.1 Vector $\vec{A}$ before (a) and after (b) operation of matrix $\bar{\bar{R}}$.

Depending on the values of the matrix elements and the components of the original vector, the weighted combinations will endow the new vector:

A new magnitude

And if the ratio of the new components differs from the ratio of the original ones:

A new direction

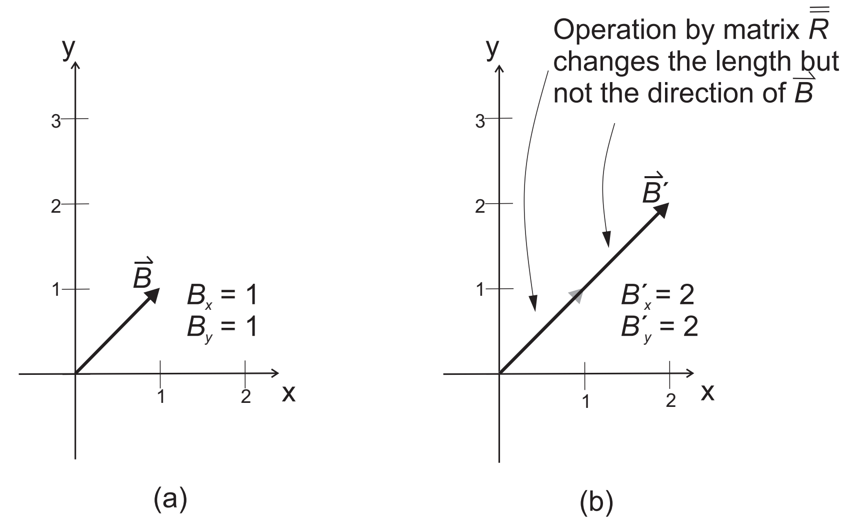
Eigenvectors and eigenvalues of a matrix

Let us try with a different vector

$$\bar{\bar{R}} = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}$$

$$\vec{B} = \hat{\imath} + \hat{\jmath}$$

$$\begin{aligned} \bar{\bar{R}}\vec{B} &= \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} (4)(1) + (-2)(1) \\ (-2)(1) + (4)(1) \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\vec{B}. \end{aligned}$$



Operation by matrix $\bar{\bar{R}}$ changes the length but not the direction of $\vec{B}$

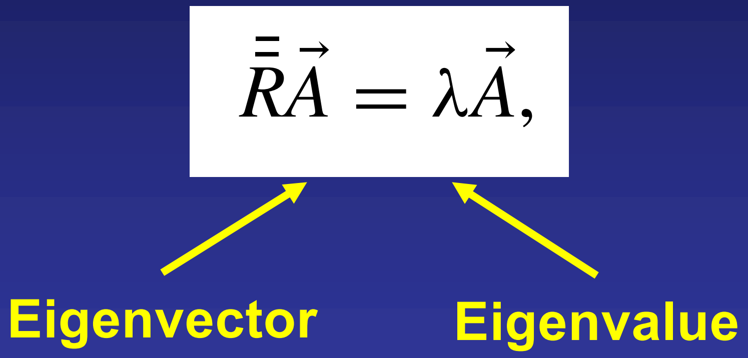
Figure 2.2 Vector $\vec{B}$ before (a) and after (b) operation of matrix $\bar{\bar{R}}$.

Now, the application of the matrix streches the length of the vector to twice its original magnitude, but it does not change the direction

A vector for which the direction is not changed after multiplication by a matrix is called an “eigenvector” of that matrix

And the factor by which the length of the vector is scaled is called the “eigenvalue” for that eigenvector

General form of an eigenvalue equation



$$\bar{\bar{R}}\vec{A} = \lambda\vec{A},$$

Eigenvector

Eigenvalue

Hints for the matrices you might found in Quantum Mechanics

The sum of the eigenvalues of a matrix is equal to the trace of the matrix (that is, the sum of the diagonal elements of the matrix)

The product of the eigenvalues is equal to the determinant of the matrix

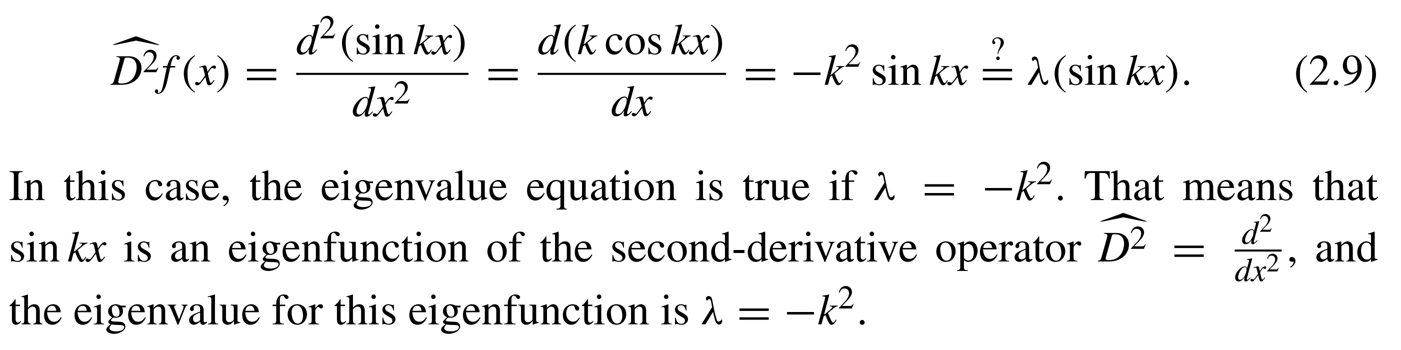
Eigenvalues of an operator

There are mathematical processes that act as operators on functions to produce new functions

If the new function is a scalar multiple of the original function, that function is called an “eigenfunction” of the operator

$$\hat{O}\underset{\substack{\nearrow\\ \textbf{Eigenfunction}}}{\psi} = \underset{\substack{\nwarrow\\ \textbf{Eigenvalue}}}{\lambda}\psi$$

Problem: find the eigenfunctions of the operator second-derivative



$$\widehat{D^2}f(x) = \frac{d^2(\sin kx)}{dx^2} = \frac{d(k\cos kx)}{dx} = -k^2\sin kx \overset{?}{=} \lambda(\sin kx). \qquad (2.9)$$

In this case, the eigenvalue equation is true if $\lambda = -k^2$. That means that $\sin kx$ is an eigenfunction of the second-derivative operator $\widehat{D^2} = \frac{d^2}{dx^2}$, and the eigenvalue for this eigenfunction is $\lambda = -k^2$.

Main Ideas of This Section

- A linear operator may be represented as a matrix that transforms a vector into another vector.
- If that new vector is a scaled version of the original vector, that vector is an eigenvector of the matrix, and the scaling factor is the eigenvalue for that eigenvector.
- An operator may also be applied to a function, producing a new function; if that new function is a multiple of the original function, then that function is an eigenfunction of the operator.

Relevance to Quantum Mechanics

- In quantum mechanics, every physical observable such as position, momentum, and energy is associated with an operator, and the state of a system may be expressed as a linear combination of the eigenfunctions of that operator.

- The eigenvalues for those eigenfunctions represent possible outcomes of measurements of that observable.

Operators in Dirac notation

$$\hat{O}|\psi\rangle = \lambda|\psi\rangle$$

Operator **Eigenvalue** **Eigenket**

Forming the inner product with the ket $|\phi\rangle$ with both sides of this equation
This is equivalent to multiply by the bra $\langle\phi|$ from the left

$$\langle\phi|\hat{O}|\psi\rangle = \langle\phi|\lambda|\psi\rangle$$

It represents a scalar (not a vector or operator)

This inner producto is proportional to the projection of the ket $\hat{O}|\psi\rangle$ onto the direction of ket $|\phi\rangle$

Expressions like this will determine one of the most useful quantities in quantum mechanics:
expectation values of measurements in quantum mechanics

Matrix representation of an operator

Sandwiching an operator between pairs of basis vectors allows you to determine the elements of the matrix representation of that operator in the basis

Example:

Consider an operator $\hat{A}$ which can be represented by a 2×2 matrix

$$\bar{\bar{A}} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

$A_{11}, A_{12}, A_{21}, \ \text{and} \ A_{22}$ (collectively defind as A_{ij})

depend on the basis system
In the same way as the components of a vector depend on the basis vectors
to which the components apply

Matrix representation of an operator

$$\bar{\bar{A}} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

How to compute the matrix elements A_{ij} ?

Apply the operator to each of the basis vectors of that system

Orthonormal basis vectors	**Represented by kets**
$\hat{\epsilon}_1$	$\|\epsilon_1\rangle$
$\hat{\epsilon}_2$	$\|\epsilon_2\rangle$

Applying the operator $\hat{A}$ to each of the orthonormal basis vectors determine the "amount" of each basis vector in the result

$$\widehat{A}\,|\epsilon_1\rangle = A_{11}\,|\epsilon_1\rangle + A_{21}\,|\epsilon_2\rangle$$
$$\widehat{A}\,|\epsilon_2\rangle = A_{12}\,|\epsilon_1\rangle + A_{22}\,|\epsilon_2\rangle$$

The columns of the matrix determine the amount of each vector

Matrix representation of an operator

$$\bar{\bar{A}} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

How to compute the matrix elements A_{ij} ?

$$\begin{aligned} \widehat{A}\,|\epsilon_1\rangle &= A_{11}\,|\epsilon_1\rangle + A_{21}\,|\epsilon_2\rangle \\ \widehat{A}\,|\epsilon_2\rangle &= A_{12}\,|\epsilon_1\rangle + A_{22}\,|\epsilon_2\rangle \end{aligned}$$

The columns of the matrix determine the amount of each vector

Taking the inner product of the first of these equations with the first basis ket $|\epsilon_1\rangle$

$$\begin{aligned} \langle\epsilon_1|\,\widehat{A}\,|\epsilon_1\rangle &= \langle\epsilon_1|\,A_{11}\,|\epsilon_1\rangle + \langle\epsilon_1|\,A_{21}\,|\epsilon_2\rangle \\ &= A_{11}\,\langle\epsilon_1|\epsilon_1\rangle + A_{21}\,\langle\epsilon_1|\epsilon_2\rangle = A_{11}, \end{aligned}$$

$$\langle\epsilon_1|\epsilon_1\rangle = 1 \;\text{ and }\; \langle\epsilon_1|\epsilon_2\rangle = 0$$

in an orthonormal basis set

$$A_{11} = \langle\epsilon_1|\hat{A}|\epsilon_1\rangle$$

Taking the inner product of the second of these equations with the first basis ket $|\epsilon_1\rangle$

$$\begin{aligned} \langle\epsilon_1|\,\widehat{A}\,|\epsilon_2\rangle &= \langle\epsilon_1|\,A_{12}\,|\epsilon_1\rangle + \langle\epsilon_1|\,A_{22}\,|\epsilon_2\rangle \\ &= A_{12}\,\langle\epsilon_1|\epsilon_1\rangle + A_{22}\,\langle\epsilon_1|\epsilon_2\rangle = A_{12} \end{aligned}$$

$$A_{12} = \langle\epsilon_1|\hat{A}|\epsilon_2\rangle$$

Matrix representation of an operator

$$\bar{\bar{A}} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

How to compute the matrix elements A_{ij} ?

$$\bar{\bar{A}} = \begin{pmatrix} \langle\epsilon_1|\widehat{A}|\epsilon_1\rangle & \langle\epsilon_1|\widehat{A}|\epsilon_2\rangle \\ \langle\epsilon_2|\widehat{A}|\epsilon_1\rangle & \langle\epsilon_2|\widehat{A}|\epsilon_2\rangle \end{pmatrix}$$

$$A_{ij} = \langle\epsilon_i|\hat{A}|\epsilon_j\rangle$$

Commutation

Two operators $\hat{A}$ and $\hat{B}$ are said to “commute” if the order of application can be switched without changing the results

$$\hat{A}(\hat{B}\,|\psi\rangle) = \hat{B}(\hat{A}\,|\psi\rangle) \quad \text{if } \widehat{A} \text{ and } \widehat{B} \text{ commute}$$

or

$$\hat{A}\hat{B}(|\psi\rangle) - \hat{B}\hat{A}(|\psi\rangle) = 0$$

$$\left(\hat{A}\hat{B} - \hat{B}\hat{A}\right)|\psi\rangle = 0.$$

↑

Commutator, usually written as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}.$$

The bigger the change in the result caused by switching the order of operation, the larger the commutator

Main Ideas of This Section

- The elements of the matrix representation of an operator in a specified basis may be determined by sandwiching the operator between pairs of basis vectors
- Two operators for which changing the order of operation does not change the result are said to commute

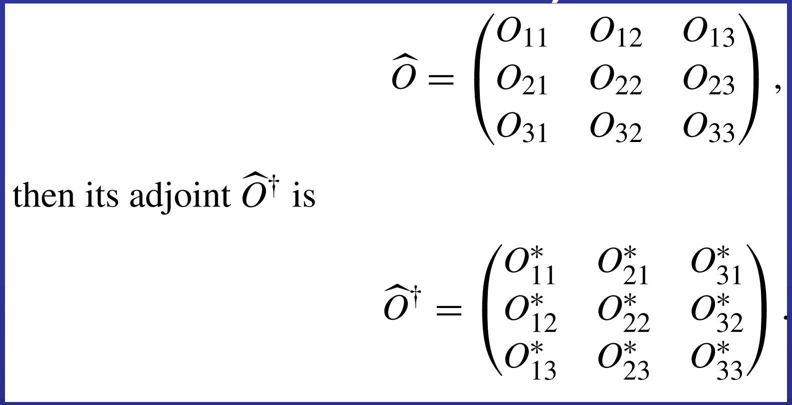
Relevance to Quantum Mechanics

- The elements of the matrix representing the important quantum operator called the "projection operator" will be found by sandwiching that operator between pairs of basis vectors
- The expression $\langle\psi|\hat{O}|\psi\rangle$ can be used to determine the expectation value of measurements of the quantum observable corresponding to operator $\hat{O}$ for a system in state $|\psi\rangle$
- Every quantum observable has an associated operator, and if two operators commute, the measurements associated with those two operators may be done in either order with the same result. That means that those two observables may be simultaneously known with precision limited only by the experimental configuration and instrumentation
- The Heisenberg Uncertainty Principle limits the precision with which two observables whose operators do not commute may be simultaneously known

Adjoint (alias “transpose conjugate, alias “Hermitian conjugate”)

The adjoint of operator $\hat{O}$ is written as $\hat{O}^\dagger$

Finding the adjoint of an operator in matrix form is straightforward:
Take the complex conjugate of each element of the matrix
And then form the transpose matrix (i.e. interchange the rows and the columns of the matrix)

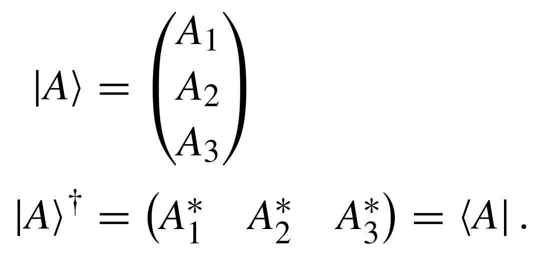


$$\widehat{O} = \begin{pmatrix} O_{11} & O_{12} & O_{13} \\ O_{21} & O_{22} & O_{23} \\ O_{31} & O_{32} & O_{33} \end{pmatrix},$$

then its adjoint $\widehat{O}^\dagger$ is

$$\widehat{O}^\dagger = \begin{pmatrix} O_{11}^* & O_{21}^* & O_{31}^* \\ O_{12}^* & O_{22}^* & O_{32}^* \\ O_{13}^* & O_{23}^* & O_{33}^* \end{pmatrix}.$$

Apply the conjugate-transpose process to a column vector

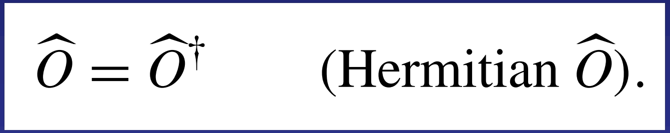


$$|A\rangle = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}$$

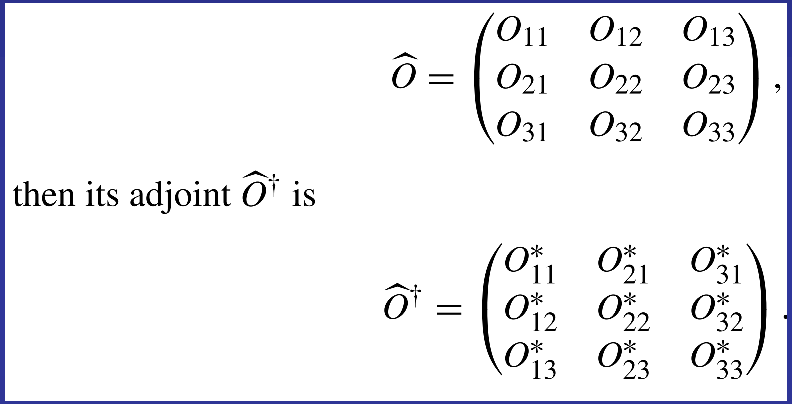
$$|A\rangle^\dagger = \begin{pmatrix} A_1^* & A_2^* & A_3^* \end{pmatrix} = \langle A| \, .$$

Hermitian operators

Hermitian operators equal their own adjoints



$$\widehat{O} = \widehat{O}^{\dagger} \qquad \text{(Hermitian } \widehat{O}\text{).}$$



$$\widehat{O} = \begin{pmatrix} O_{11} & O_{12} & O_{13} \\ O_{21} & O_{22} & O_{23} \\ O_{31} & O_{32} & O_{33} \end{pmatrix},$$

then its adjoint $\widehat{O}^{\dagger}$ is

$$\widehat{O}^{\dagger} = \begin{pmatrix} O_{11}^{*} & O_{21}^{*} & O_{31}^{*} \\ O_{12}^{*} & O_{22}^{*} & O_{32}^{*} \\ O_{13}^{*} & O_{23}^{*} & O_{33}^{*} \end{pmatrix}.$$

- **Diagonal elements of the matrix must be real**
- **Every off-diagonal element must equal the complex conjugate of the corresponding element on the other side of the diagonal**

Why Hermitian operators are important

- Hermitian operators must have real eigenvalues

- Eigenfunctions of an Hermitian operator with different eigenvalues must be ortogonal

- The eigenfuctions of a Hermitian operator: they form a complete set.
  That means that any function in the abstract vector space containing the eigenfunctions of a Hermitian operator may be made up of a linear combination of those eigenfunctions.

Main Ideas of this Section

- Hermitian operators may be applied to either member of an inner producto and the result will be the same.
- Hermitian operators have real eigenvalues, and the nondegenerate eigenfunctions of a Hermitian operator are orthogonal and form a complete set

Relevance to Quantum Mechanics

- The discussion of the solutions to the Schrödinger equation will show that every quantum observable (such as position, momentum, and energy) is associated with an operator.
- The possible results of any measurement are given by the eigenvalues of that operator.
- Since the results of measurements must be real, operators associated with observables must be Hermitian.
- The eigenfunctions of Hermitian operators are (or can be combined to be) orthogonal, and the orthogonality of those eigenfunctions has a profound impact on our ability to construct solutions to the Schrödinger equation and to use those solutions to determine the probability of various measurement outcomes.

Projection operators

The general expression for a projector operator is

$$\hat{P}_i = |\epsilon_i\rangle\langle\epsilon_i|$$

where $|\epsilon_i\rangle$ is a normalized vector

$$\widehat{P}_1 \, |A\rangle = |\epsilon_1\rangle \, \langle\epsilon_1|A\rangle = A_1 \, |\epsilon_1\rangle$$

So applying the projection operator to $|A\rangle$ produces the new ket $A_1|\epsilon_1\rangle$ The magnitude of that new ket is the (scalar) projection of the ket that you feed into the operator (in this case, $|A\rangle$) onto the direction of the ket you use to define the operator (in this case, $|\epsilon_1\rangle$).

But here's an important step:

that magnitude is then multiplied by the ket you use to define the operator.

So the result of applying the projection operator to a ket is not just the (scalar) component (such as A_1) of that ket along the direction of $|\epsilon_1\rangle$, it's a new ket in that direction

Closure or completeness relations

$$\sum_{n=1}^{N} \widehat{P}_n = \sum_{n=1}^{N} |\epsilon_n\rangle \langle\epsilon_n| = \underset{\uparrow}{\widehat{I}}$$

Identity operator

Main Ideas of This Section

- The projection operator is a Hermitian operator that projects one vector onto the direction of another and forms a new vector in that direction

- Operating on a vector with the projection operators for all of the basis vectors of that space reproduces the original vector

- That means that the sum of the projection operators for all the basis vectors equals the identity operator (completeness relation)

Relevance to Quantum Mechanics

- The projection operator is useful in determining the probability of measurement outcomes for a quantum observable by projecting the state of a system onto the eigenstates of the operator for that observable

Expectation values

Specific predictions about *average* measurement outcomes are posible in Quantum Mechanics, provided you know two things:

- the operator ($\hat{O}$) corresponding to the measurement you plan to make,
- the state of the system, represented by the ket $|\psi\rangle$ prior to the measurement

Predictions come in the form of expectation value of an observable

$$\langle o\rangle = \langle\psi|\hat{O}|\psi\rangle$$

the angle brackets on the left side signify the expectation value:
the average value of the outcome of a number of measurements
of the observable associated with the operator $\hat{O}$

(average over many systems all prepared to be on the same state prior to the measurement)

Expectation values

The ket representing the state of a system can be written as the weighted combination of the kets representing the eigenvectors of an operator

$$|\psi\rangle = c_1\,|\psi_1\rangle + c_2\,|\psi_2\rangle + \cdots + c_N\,|\psi_N\rangle = \sum_{n=1}^{N} c_n\,|\psi_n\rangle\,,$$

$c_1, \ldots, c_n$ represent the amount of each orthonormal eigenfunction $|\psi_n\rangle$ in $|\psi\rangle$

Applying the operator to the state

$$\widehat{O}\,|\psi\rangle = \widehat{O}\sum_{n=1}^{N} c_n\,|\psi_n\rangle = \sum_{n=1}^{N} c_n\widehat{O}\,|\psi_n\rangle = \sum_{n=1}^{N} \lambda_n c_n\,|\psi_n\rangle$$

λ_n eigenvalues of the operator applied to the eigenket $|\psi_n\rangle$

The bra corresponding to the state is

$$\langle\psi| = \langle\psi_1|\,c_1^* + \langle\psi_2|\,c_2^* + \cdots + \langle\psi_N|\,c_N^* = \sum_{m=1}^{N} \langle\psi_m|\,c_m^*,$$

Expectation values

$$\begin{aligned}\langle\psi|\,\widehat{O}\,|\psi\rangle &= \sum_{m=1}^{N}\langle\psi_m|\,c_m^*\sum_{n=1}^{N}\lambda_n c_n\,|\psi_n\rangle\\ &= \sum_{m=1}^{N}\sum_{n=1}^{N}c_m^*\lambda_n c_n\,\langle\psi_m|\psi_n\rangle\,.\end{aligned}$$

Since the eigenkets are orthonormal

$$\langle\psi|\,\widehat{O}\,|\psi\rangle = \sum_{n=1}^{N}\lambda_n c_n^* c_n = \sum_{n=1}^{N}\lambda_n|c_n|^2 = \langle o\rangle$$

$\langle\psi|\hat{O}|\psi\rangle$ **will produce the expectation value as long as the square magnitude of** c_n **represents the probability of obtaining result** λ_n

Uncertainties

Square of the expectation value

$$\Delta x = \sqrt{\left\langle x^2 \right\rangle - \langle x \rangle^2}$$

Expectation value of the square

$$\text{Variance of } x = (\Delta x)^2 \equiv \left\langle (x - \langle x \rangle)^2 \right\rangle$$