3 Operators (I)

3.1 Linear operators

As was mentioned in Sections 1.2 and 1.3, measurable physical quantities correspond to certain mathematical operators in Quantum Mechanics. You have encountered a number of such operators in the Term 1 course, including the Hamiltonian operator, the position operator, the momentum operator and various angular momentum operators. You have also encountered operators such as ladder operators, which do not correspond to measurable quantities.

These operators are mathematical objects which transform elements of a certain vector space into elements of the same vector space. In other words, they map vectors to vectors.

Examples

- As we will see later in the course, the matrix

$$S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{3.1}$$

  is an operator related to spin measurements. This operator transforms 2-component column vectors into 2-component column vectors. For instance, it transforms the column vector

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} \tag{3.2}$$

  into the column vector

$$S_y\chi = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} -ib \\ ia \end{pmatrix}. \tag{3.3}$$

- As another example, take the Hamiltonian of a linear harmonic oscillator of mass m,

$$H = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x), \tag{3.4}$$

  where $V(x)$ is the potential energy. In terms of the oscillator's angular frequency ω,

$$V(x) = m\omega^2x^2/2. \tag{3.5}$$

  This operator transforms functions into functions. For instance, it transforms the function

$$\psi(x) = \exp(-\alpha^2x^2/2), \tag{3.6}$$

where α is a real constant, into the function

$$\begin{aligned} H\psi(x) &= -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V(x)\psi(x) & (3.7)\\ &= \left[-\frac{\hbar^2}{2m}\left(\alpha^4 x^2 - \alpha^2\right) + \frac{m\omega^2}{2}x^2\right]\exp(-\alpha^2 x^2/2). & (3.8) \end{aligned}$$

Domain of an operator

This last example illustrates the need for a careful definition of the vector space in which an operator acts. Since it includes a term in $\mathrm{d}^2/\mathrm{d}x^2$, the Hamiltonian operator of Eq. (3.4) is defined only when acting on functions which can be differentiated twice. Such functions form a subspace of the vector space of all functions of x. Thus H maps twice-differentiable functions to functions (the latter may or may not be twice-differentiable themselves, hence the mapping is not from the space of all twice-differentiable functions to the space of all twice-differentiable functions).

More generally, an operator is a mapping from a subspace V' of a vector space V to the vector space V itself. The vector space V' in which the operator acts is called the domain of this operator.

It might be that $V' = V$, in which case the domain of the operator is the whole of V. (By definition of a subspace, V' is a subset of V; however, as a subset of a set can be this set itself, V' can be the whole of V.) This is the case of the operator S_y defined in the first example: any 2-component column vector can be multiplied by a 2×2 matrix, hence the domain of this operator is the whole of the vector space of 2-component column vectors. However, Quantum Mechanics also makes use of operators whose domain is smaller than the vector space they map to — typically, these are operators acting on wave functions, such as the Hamiltonian operator H of Eq. (3.4).

Linear operators

Operator representing measurable physical quantities and most of the other operators used in Quantum Mechanics have the important mathematical property to be linear. An operator A is said to be linear if it fulfils the following two conditions:

1. If w is the sum of the vectors v_1 and v_2, then $Aw = Av_1 + Av_2$.
2. If w is the product of a vector v by a scalar c, then $Aw = cAv$. (The right-hand side is meant to represent the product of the vector Av by the scalar c.)

These two conditions can be summarized into the single condition that for any vector v_1 and v_2 and any scalar c_1 and c_2,

$$A(c_1 v_1 + c_2 v_2) = c_1 A v_1 + c_2 A v_2. \tag{3.9}$$

For example, the differential operator $\mathrm{d}/\mathrm{d}x$ is linear since, if c_1 and c_2 are constants,

$$\frac{\mathrm{d}}{\mathrm{d}x}[c_1 f_1(x) + c_2 f_2(x)] = c_1 \frac{\mathrm{d}f_1}{\mathrm{d}x} + c_2 \frac{\mathrm{d}f_2}{\mathrm{d}x}. \tag{3.10}$$

Throughout the rest of the course we will always assume that the operators we are talking about are linear operators.

- ☞ Not all operators are linear. For example, consider an operator O which would multiply vectors by their norm, i.e., such that

  $$Ov = ||v||\, v \tag{3.11}$$

  for any vector v of the space in which this operator acts. This operator violates the conditions operators must fulfil to qualify as linear operators. In particular, it is not the case that $O(cv) = c\,Ov$ for *any* scalar c and *any* vector v, since

  $$O(cv) = ||cv||\, cv, \tag{3.12}$$

  and $||cv||\, cv \neq c\,||v||\, v$ unless $c = 0$, $|c| = 1$ or $v = 0$. Hence O is not a linear operator.

- ☞ Linear operators are particular instances of more general mappings called linear transformations.

The identity operator

The identity operator, which is usually denoted by the letter I, is the operator which “transforms” any vector into itself:

$$Iv = v \tag{3.13}$$

for any vector v. We will use this operator from time to time.

3.2 Matrix representation of an operator

We start by the case of operators acting in a finite-dimensional Hilbert space, e.g., spin operators. Recall that an operator A acting on a vector v transforms this vector into a vector denoted Av. Let us set

$$w = Av. \tag{3.14}$$

Let us also assume that these vectors belong to a finite-dimensional Hilbert space of dimension N. We can therefore write them as linear combinations of N orthonormal basis vectors u_n:

$$\begin{aligned} v &= c_1 u_1 + c_2 u_2 + \cdots + c_N u_N, & (3.15)\\ w &= d_1 u_1 + d_2 u_2 + \cdots + d_N u_N, & (3.16) \end{aligned}$$

where the coefficients $c_1,\ldots, c_N$ and $d_1,\ldots, d_N$ are in general complex numbers. Recall that saying that the vectors u_n are orthornormal means that

$$(u_i, u_j) = \delta_{ij}, \tag{3.17}$$

and that $c_n = (u_n, v)$ and $d_n = (u_n, w)$, $n = 1, 2,\ldots, N$. (See Section 2.10 for further information about orthonormal bases.) Eq. (3.14) can be directly written as a relation between the coefficients c_n and the the coefficients d_n: One can organise these two sets of coefficients into two column vectors, c and d, such that

$$\mathsf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \qquad \text{and} \qquad \mathsf{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{pmatrix}, \tag{3.18}$$

and in terms these two column vectors Eq. (3.14) reads

$$\mathsf{d} = \mathsf{A}\,\mathsf{c} \tag{3.19}$$

where A is the $N \times N$ matrix

$$\mathsf{A} = \begin{pmatrix} (u_1, Au_1) & (u_1, Au_2) & \cdots & (u_1, Au_N) \\ (u_2, Au_1) & (u_2, Au_2) & \cdots & (u_2, Au_N) \\ \vdots & \vdots & \ddots & \vdots \\ (u_N, Au_1) & (u_N, Au_2) & \cdots & (u_N, Au_N) \end{pmatrix}. \tag{3.20}$$

Proof: Written in terms of the basis set expansions of the vectors v and w, Eq. (3.14) reads

$$\sum_{j=1}^{N} d_j u_j = \sum_{j=1}^{N} c_j A u_j. \tag{3.21}$$

Taking the inner product of this equation with u_1, we see that

$$\sum_{j=1}^{N} d_j (u_1, u_j) = \sum_{j=1}^{N} c_j (u_1, A u_j). \tag{3.22}$$

In view of Eq. (3.17), this equation reduces to

$$d_1 = \sum_{j=1}^{N} c_j (u_1, A u_j). \tag{3.23}$$

Likewise, taking the inner product of Eq. (3.21) with u_2 yields

$$d_2 = \sum_{j=1}^{N} c_j (u_2, A u_j), \tag{3.24}$$

and similarly for all the other coefficients d_n. In general,

$$d_i = \sum_{j=1}^{N} A_{ij} c_j, \quad i = 1, 2, \ldots, N, \tag{3.25}$$

with $A_{ij} = (u_i, A u_j)$. Eq. (3.19) follows. $\square$

The matrix A is said to represent the operator A in the basis $\{u_n\}$. Its elements — i.e., the inner products $(u_i, A u_j)$ — are called the matrix elements of A in that basis.

It is clear that the column vectors c and d representing the vectors v and w and the matrix A representing the operator A all depend on the basis: changing the basis from a set of orthonormal vectors $\{u_n\}$ to a set of orthonormal vectors $\{u'_n\}$ changes the column vectors c and d into column vectors c' and d' of elements $c'_n = (u'_n, v)$ and $d'_n = (u'_n, w)$ and changes the matrix A into a matrix A' of elements $(u'_i, A v'_j)$. As long as the set $\{u'_n\}$ is also an orthonormal basis, however, the column vector d' is related to A' and to c' in the same way as d is related to A and to c:

$$\mathsf{d}' = \mathsf{A}'\,\mathsf{c}'. \tag{3.26}$$

Many different sets of units vectors can form an orthonormal basis.[1] Therefore, a same operator can be represented by many different matrices.

Examples

- As mentioned in Section 2.11, the spherical harmonics $Y_{lm}(\theta, \phi)$ are orthonormal in the sense that

$$\int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi\, Y^*_{lm}(\theta, \phi)\, Y_{l'm'}(\theta, \phi) = \delta_{ll'} \delta_{mm'}. \tag{3.27}$$

[1] Actually, infinitely many different sets of units vectors can form an orthonormal basis if the dimension of the vector space is 2 or higher. There is no choice of basis set possible in spaces of dimension 1.

Therefore the three $l = 1$ spherical harmonics $Y_{1-1}(\theta, \phi)$, $Y_{10}(\theta, \phi)$ and $Y_{11}(\theta, \phi)$, constitute an orthonormal basis for the vector space of all linear combinations of the form

$$c_{-1}\, Y_{1-1}(\theta, \phi) + c_0\, Y_{10}(\theta, \phi) + c_1\, Y_{11}(\theta, \phi),$$

where c_{-1}, c_0 and c_1 are complex numbers (see Section 2.4 for this vector space). You may remember from the Term 1 course that the spherical harmonics are eigenfunctions of the angular momentum operator L_z. More precisely,

$$L_z = -i\hbar \frac{\partial}{\partial \phi} \tag{3.28}$$

and

$$L_z Y_{lm}(\theta, \phi) = m\hbar Y_{lm}(\theta, \phi). \tag{3.29}$$

Hence, when L_z acts on a linear combination of spherical harmonics with $l = 1$, the result is also a linear combination of spherical harmonics with $l = 1$; in fact,

$$\begin{aligned} L_z[c_{-1}\, Y_{1-1}(\theta, \phi) + &c_0\, Y_{10}(\theta, \phi) + c_1\, Y_{11}(\theta, \phi)] \\ &= [-\hbar\, c_{-1}\, Y_{1-1}(\theta, \phi) + \hbar\, c_1\, Y_{11}(\theta, \phi)]. \end{aligned} \tag{3.30}$$

One can therefore represent the operator L_z by a 3×3 matrix L_z in the basis formed by the spherical harmonics $Y_{1-1}(\theta, \phi)$, $Y_{10}(\theta, \phi)$ and $Y_{11}(\theta, \phi)$, the elements of this matrix being the integrals

$$\int_0^\pi \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi\, Y^*_{lm}(\theta, \phi)\, L_z\, Y_{l'm'}(\theta, \phi). \tag{3.31}$$

These integrals are easy to calculate in view of Eqs. (3.27) and (3.29). The calculation gives

$$\mathsf{L}_z = \begin{pmatrix} -\hbar & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hbar \end{pmatrix}. \tag{3.32}$$

Written in terms of this matrix, Eq. (3.30) reads

$$\begin{pmatrix} -\hbar & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hbar \end{pmatrix} \begin{pmatrix} c_{-1} \\ c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} -\hbar c_{-1} \\ 0 \\ \hbar c_1 \end{pmatrix}. \tag{3.33}$$

It is worth noting that the column vectors representing the linear combinations of spherical harmonics and the matrix representing the operator L_z depend on the order of the basis functions in the basis set. Eqs. (3.32) and (3.33) apply to the case where the basis is the ordered set

$$\{Y_{1-1}(\theta, \phi), Y_{10}(\theta, \phi), Y_{11}(\theta, \phi)\}.$$

If we had taken the basis to be the ordered set

$$\{Y_{10}(\theta,\phi), Y_{11}(\theta,\phi), Y_{1-1}(\theta,\phi)\},$$

these two equations would have been

$$\mathsf{L}_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \hbar & 0 \\ 0 & 0 & -\hbar \end{pmatrix} \tag{3.34}$$

and

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & \hbar & 0 \\ 0 & 0 & -\hbar \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_{-1} \end{pmatrix} = \begin{pmatrix} 0 \\ \hbar c_1 \\ -\hbar c_{-1} \end{pmatrix}. \tag{3.35}$$

Other choices of basis functions are also possible. For instance, we can work with the functions $Y_{1x}(\theta,\phi)$, $Y_{1y}(\theta,\phi)$ and $Y_{1z}(\theta,\phi)$ defined as follows:

$$Y_{1x}(\theta,\phi) = \ [Y_{1-1}(\theta,\phi) - Y_{11}(\theta,\phi)]/\sqrt{2}, \tag{3.36}$$

$$Y_{1y}(\theta,\phi) = i[Y_{1-1}(\theta,\phi) + Y_{11}(\theta,\phi)]/\sqrt{2}, \tag{3.37}$$

$$Y_{1z}(\theta,\phi) = \ Y_{10}(\theta,\phi). \tag{3.38}$$

It is not particularly difficult to show that these three functions also form an orthonormal basis for the vector space spanned by the spherical harmonics $Y_{1-1}(\theta,\phi)$, $Y_{10}(\theta,\phi)$ and $Y_{11}(\theta,\phi)$, and that in the basis $\{Y_{1x}(\theta,\phi), Y_{1y}(\theta,\phi), Y_{1z}(\theta,\phi)\}$ the operator L_z is represented by the matrix

$$\mathsf{L}'_z = \begin{pmatrix} 0 & i\hbar & 0 \\ -i\hbar & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.39}$$

(The calculation is left as an exercise.)

- The five spherical harmonics $Y_{2-2}(\theta,\phi)$, $Y_{2-1}(\theta,\phi)$, $Y_{20}(\theta,\phi)$, $Y_{21}(\theta,\phi)$ and $Y_{22}(\theta,\phi)$ constitute an orthonormal basis for the vector space of all linear combinations of the form

$$c_{-2}\,Y_{2-2}(\theta,\phi) + c_{-1}\,Y_{2-1}(\theta,\phi) + c_0\,Y_{20}(\theta,\phi) + c_1\,Y_{21}(\theta,\phi) + c_2\,Y_{22}(\theta,\phi),$$

where the coefficients c_n are complex numbers. In the basis formed by these five spherical harmonics, the angular momentum operator is represented by the matrix

$$\mathsf{L}_z = \begin{pmatrix} -2\hbar & 0 & 0 & 0 & 0 \\ 0 & -\hbar & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \hbar & 0 \\ 0 & 0 & 0 & 0 & 2\hbar \end{pmatrix}. \tag{3.40}$$

☞ Although its analytical form [Eq. (3.28)] is the same as in the previous example, this operator is represented by a 5×5 matrix here, not by a 3×3 matrix. It may seem bizarre that a same operator can be represented by matrices of different sizes. However, technically, the L_z operator considered here is not the same operator as the L_z operator considered above: The mathematical definition of an operator includes a specification of the vector space in which the operator acts, and this vector space differs between these two examples.

- Since $(u_i, Iu_j) = (u_i, u_j) = \delta_{ij}$ if I is the identity operator and the vectors u_i and u_j are orthonormal, this operator is always represented by the unit matrix in any orthonormal basis (specifically, by the $N \times N$ unit matrix if the space is of dimension N).[2]

Operators in infinite-dimensional vector spaces

All what we have seen above for the case of operators acting in finite-dimensional vector spaces generalizes to the case of infinite-dimensional vector spaces, although the rigorous mathematical theory of the latter is more difficult. Ignoring (as usual) various mathematical subtleties, any linear operator acting on vectors belonging to an infinite-dimensional Hilbert space can be represented, at least formally, by "square" matrices of an infinite number of elements. Such matrices can be constructed as in the finite-dimensional case: given an orthonormal basis spanning the space, $\{u_1, u_2, u_3, \ldots\}$, an operator A is represented in that basis by a matrix A whose elements A_{ij} are the inner products (u_i, Au_j). (A_{ij} is the element of A located on the i-th row and in the j-th column.) Likewise, a vector v of this Hilbert space is represented by a column vector c whose i-th component ($i = 1, 2, 3, \ldots$) is the inner product (u_i, v). In this representation, the vector Av is calculated as the product of the column vector c by the matrix A.

Take, for example, the Hilbert space of all square-integrable functions of the polar angles θ and ϕ, which is infinite-dimensional. (We are talking about the Hilbert space of *all* square-integrable functions of these two angles, not about a finite-dimensional Hilbert space of functions that can be written as a linear combination of spherical harmonics with same l values as in the previous examples.) One can show that the spherical harmonics $Y_{lm}(\theta, \phi)$ form an orthonormal basis for this Hilbert space. In the basis

$$\{Y_{00}(\theta,\phi), Y_{1-1}(\theta,\phi), Y_{10}(\theta,\phi), Y_{11}(\theta,\phi), Y_{2-2}(\theta,\phi), Y_{2-1}(\theta,\phi), \ldots\},$$

[2]Recall that the unit matrix is the diagonal matrix whose diagonal elements are 1 and off diagonal elements are 0.

the angular momentum operator L_z is represented by the infinite matrix

$$\mathsf{L}_z = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & -\hbar & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \hbar & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & -2\hbar & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & -\hbar & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.41}$$

Likewise, a function $f(\theta,\phi)$ which can be expanded as

$$f(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm} Y_{lm}(\theta,\phi) \tag{3.42}$$

is represented, in that basis, by the column vector

$$\mathsf{c} = \begin{pmatrix} c_{00} \\ c_{1-1} \\ c_{10} \\ c_{11} \\ c_{2-2} \\ c_{2-1} \\ \vdots \end{pmatrix}. \tag{3.43}$$

Warning: The mathematical theory of finite matrices does not extend straightforwardly to infinite matrices. For example, a column vector of infinitely many components can be multiplied by a row vector of infinitely many components only if the resulting sum of products of components converges; the issue does not arise in the finite-dimensional case.

3.3 Adding and multiplying operators

Sum of two operators

Operators acting on the same vectors can be added. The sum $A + B$ of an operator A and an operator B is the operator $A + B$ such that

$$(A + B)v = Av + Bv \tag{3.44}$$

for any vector v the operators A and B can both act on.

☞ For example, the linear harmonic oscillator Hamiltonian given by Eq. (3.4) can be seen as being the sum of the operator

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}$$

and the operator $V(x)$, representing, respectively, the kinetic energy and the potential energy of the oscillator. The latter is a multiplicative operator: seen as an operator, $V(x)$ transforms a function $\psi(x)$ into the function $V(x)\psi(x)$ [e.g., transforms the function $\exp(-\alpha^2x^2/2)$ into the function $(m\omega^2x^2/2)\exp(-\alpha^2x^2/2)$].

Not surprisingly, the matrix representing the sum of two operators is the sum of the corresponding matrices: E.g., if the operators A and A' are represented by the matrices

$$\mathsf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \mathsf{A}' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}, \tag{3.45}$$

then the operator $A + A'$ is represented by the matrix

$$\mathsf{A} + \mathsf{A}' = \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix}. \tag{3.46}$$

Multiplication by a scalar

Operators can also be multiplied by scalars. To state the obvious, if α is a number, the operator αA is defined as the operator such that

$$(\alpha A)v = \alpha(Av) \tag{3.47}$$

for any vector v the operator A can act on. Therefore one can make linear combinations of operators in the same way as one can make linear combinations of vectors. Examples of such linear combinations are the ladder operators $\hat{J}_+$ and $\hat{J}_-$ used in the theory of the angular momentum, which are defined in terms of the angular momentum operators $\hat{J}_x$ and $\hat{J}_y$ as $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$ (you may have encountered these operators in Term 1, and we will come back to them later in this course).

In terms of matrix representations, multiplying an operator by a number amounts to multiplying the corresponding matrix by this number, which also amounts to multiplying each of the elements of that matrix by this number:

$$\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{pmatrix}. \tag{3.48}$$

Products of operators

Less obvious perhaps is that operators acting in the same space can also be multiplied: If A and B are two operators, then, by definition, their product AB is the operator such that

$$(AB)v = A(Bv) \tag{3.49}$$

for any vector v (or more precisely, for any vector v for which the right-hand side of this equation is defined). I.e., operating on a vector v with the product operator AB is operating on v with B and then operating on the resulting vector with A. Recall that the order of the operators in such products often matters: in many cases, AB and BA are different operators (see Section 3.4 for further information about non-commuting operators).

Clearly, an operator can be multiplied by itself to form the square of that operator, and this process can be iterated to form higher powers. For example, if A is an operator, the operator A^2 is the product AA, A^3 is the product AA^2 (which can also be written AAA and A^2A), etc.

In terms of matrix representations, the matrix representing a product operator AB in a given basis is the product of the matrix representing A with the matrix representing B in the same basis. (These matrices do not commute if the operators A and B do not commute, in which case they must be multiplied in the same order as the corresponding operators.)

Proof: The elements of the matrix representing a product operator AB in an orthonormal basis $\{u_1, u_2, \ldots, u_N\}$ are the inner products (u_n, ABu_m), $n, m = 1, \ldots, N$. Now, by definition of the product of two operators, $ABu_m = A(Bu_m)$. Since the vectors $\{u_1, u_2, \ldots, u_N\}$ form a basis, it is always possible to write Bu_m as a linear combination of these basis vectors, for each u_m:

$$Bu_m = \sum_{i=1}^{N} c_i^{(m)} u_i, \tag{3.50}$$

where the coefficients $c_i^{(m)}$ are scalars. Thus

$$(u_n, ABu_m) = \sum_{i=1}^{N} c_i^{(m)} (u_n, Au_i). \tag{3.51}$$

However, $c_i^{(m)} = (u_i, Bu_m)$ since the basis is orthogonal. Therefore

$$(u_n, ABu_m) = \sum_{i=1}^{N} (u_i, Bu_m)(u_n, Au_i). \tag{3.52}$$

Since the matrix elements (u_n, Au_i) and (u_i, Bu_m) are numbers and numbers commute, Eq. (3.52) can also be written as

$$(u_n, ABu_m) = \sum_{i=1}^{N} (u_n, Au_i)(u_i, Bu_m). \tag{3.53}$$

Eq. (3.53) says that the element on the n-th row and m-th column of the matrix representing AB is obtained by multiplying the n-th row of the matrix representing A by the m-th column of the matrix representing B. Hence, the matrix representing AB is the product of these two matrices. □

☞ You may have noticed that what is written above is a bit imprecise in regards to the domain of the operators concerned. A product operator AB may act only on vectors v which are in the domain of B *and* such that Bv is in the domain of A.

Exponentials of operators

More complicated functions of operators can also be defined. One which often crops up in Quantum Mechanics is the exponential of an operator. Recall that if z is a number,

$$\exp(z) = 1 + z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!} z^n. \tag{3.54}$$

Similarly, the exponential of an operator A is defined by the following equation:

$$\exp(A) = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!} A^n. \tag{3.55}$$

☞ Warning: Whereas $\exp(a+b) = \exp(a)\exp(b)$ if a and b are two numbers, it is in general not the case that $\exp(A+B) = \exp(A)\exp(B)$ if A and B are two operators (this equation is correct if A and B commute, though).

3.4 Commutators

Two operators A and B are said to commute if $ABv = BAv$ for any vector v these operators may act on. (Recall from the previous section that ABv is the vector obtained by first transforming v with B and then with A, while BAv is the vector obtained by first transforming v with A and then with B.) If A and B commute then $AB = BA$ and $AB - BA = 0$. (The right-hand side of this last equation is the zero operator, i.e., the operator which transforms any vector into the zero vector. Acting with this operator on a vector amounts to a multiplication by the scalar 0.)

☞ More precisely, A and B commute if $ABv = BAv$ for all the vectors v which are in the domain of AB as well as in the domain of BA. It is possible for these different operators to have different domains if they act in an infinite-dimensional vector space.

The commutator of two operators A and B is the operator $AB - BA$. This operator is usually represented by the symbol $[A, B]$:

$$[A, B] = AB - BA. \tag{3.56}$$

Therefore

$$[A, B]v = ABv - BAv. \tag{3.57}$$

Clearly, the commutator of two commuting operators is zero.

For example, take the matrix operators

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

which are used to represent spin operators (see later in the course). These two operators do not commute since $\sigma_z \sigma_x$ and $\sigma_x \sigma_z$ are different matrices:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.58}$$

whereas

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.59}$$

Clearly,

$$[\sigma_z, \sigma_x] = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}. \tag{3.60}$$

☞ Note: You may remember to have seen the equation $[x, p_x] = i\hbar$, where x and p_x are, respectively, the position and momentum operators for the x-direction. Since the left-hand side of this equation is an operator, the correct (but pedantic) way of writing its right-hand side is $i\hbar I$, where I is the identity operator. Writing $[x, p_x] = i\hbar$ is completely acceptable, though, unless there would be a risk of confusion.

A few properties of commutators worth remembering: For any A, B,, C,

- $[B, A] = -[A, B]$;
- $[A, I] = 0$, where I is the identity operator;
- $[A, A] = 0$ (an operator always commutes with itself);
- $[A, A^2] = 0$, and more generally $[A, f(A)] = 0$ for any function $f(A)$ of the operator A [e.g., $\exp(A)$];
- $[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0$ (the Jacobi identity for commutators).

3.5 The inverse of an operator

If A is an operator and there exists an operator B such that $AB = I$ and $BA = I$, where I is the identity operator, then one says that A is invertible and that B is its inverse. The inverse of an operator is usually denoted by the superscript $^{-1}$. Thus A^{-1} is the inverse of A if and only if

$$AA^{-1} = A^{-1}A = I. \tag{3.61}$$

> ☞ In finite-dimensional spaces, that $AA^{-1} = I$ implies that $A^{-1}A = I$, and the other way round. However, the two conditions $AA^{-1} = I$ and $A^{-1}A = I$ are not equivalent in infinite-dimensional spaces.

The following theorems are easily proven:

1. If the operators A and B are both invertible, then the product AB is also invertible and
$$(AB)^{-1} = B^{-1}A^{-1}. \tag{3.62}$$
2. The inverse of the inverse of an operator is the operator itself:
$$(A^{-1})^{-1} = A. \tag{3.63}$$
3. If the operator A is invertible, then Av is the zero vector only if v is the zero vector.
4. If the operator A is invertible and is represented by a matrix A, then this matrix has an inverse, A^{-1}, and this inverse matrix represents the inverse of the operator A.

Note quite all operators are invertible. Recall, for example, that a matrix whose determinant is zero is not invertible. One says that an operator is singular if it not invertible.

3.6 Eigenvalues and eigenvectors

Suppose that v is a non-zero vector such that

$$Av = \lambda v \tag{3.64}$$

for a certain scalar λ. One then says that λ is an eigenvalue and v an eigenvector of the operator A. Or, more specifically,one says that v is an "eigenvector of A with eigenvalue λ" or an "eigenvector of A belonging to the eigenvalue λ". (Don't be confused by this equation: the left-hand side represents the vector resulting from the action of A on v while the right-hand side represents the vector obtained by multiplying v by the number λ.)

Clearly, if the operator A and the vector v appearing in Eq. (3.64) are represented by a matrix A and a column vector c, then one also has

$$\mathsf{A}\mathsf{c} = \lambda\mathsf{c}. \tag{3.65}$$

I.e., the column vector c is an eigenvector of the matrix A.

One often uses the term eigenfunction instead of eigenvector if the operator considered acts on functions, as in the case, for example, of Eq. (3.29) of Section 3.2. It may be worth stressing that the words “eigenfunction” and “wave function” do not mean the same thing. An eigenfunction is what we just defined. A wave function is a function representing a quantum state. An eigenfunction may or may not represent a certain quantum state, and therefore may or may not also be a wave function. Similarly, a wave function may or may not be an eigenfunction of some interesting operator. For example, the time-dependent function

$$\Psi(r,\theta,\phi,t) = [\psi_{100}(r,\theta,\phi)\exp(-iE_1t/\hbar) + \psi_{211}(r,\theta,\phi)\exp(-iE_2t/\hbar)]/\sqrt{2} \tag{3.66}$$

is a wave function representing a linear superposition of the 1s and $2\mathrm{p}_{m=1}$ states of atomic hydrogen. Although a valid wave function, solution of the time-dependent Schrödinger equation, $\Psi(r,\theta,\phi,t)$ is *not* an eigenfunction of the Hamiltonian, the angular momentum operator $\mathbf{L}^2$ or the angular momentum operator L_z.

☞ It should be noted that only vectors belonging to the space in which the operator is defined are regarded as being eigenvectors of this operator (eigenvectors in the sense normally given to this term in Mathematics). For example, the differential operator $\mathrm{d}/\mathrm{d}x$ has infinitely many eigenfunctions if regarded as an operator acting on *any* differentiable function, since

$$\frac{\mathrm{d}}{\mathrm{d}x}\exp(\lambda x) = \lambda\exp(\lambda x) \tag{3.67}$$

for any real or complex λ. But $\mathrm{d}/\mathrm{d}x$ has no eigenfunction if regarded as an operator acting *only* on differentiable *square-integrable* functions on $(-\infty,\infty)$, since all the solutions of the equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \lambda y(x) \tag{3.68}$$

are of the form $C\exp(\lambda x)$, where C is a constant, and none of these solutions is square-integrable on $(-\infty,\infty)$. We will come back to this issue in a later part of these notes.

Degenerate eigenvalues

It may happen that several *linearly-independent* vectors belong to a same eigenvalue. This eigenvalue is said to be degenerate in that case.

In particular, one says that the eigenvalue λ is M-fold degenerate (or that its degree of degeneracy is M) if there exist M linearly-independent vectors v_1, v_2,…, v_M such that

$$Av_n = \lambda v_n, \quad n = 1, 2, \ldots, M, \tag{3.69}$$

and if any other eigenvector belonging to that eigenvalue is necessarily a linear combination of these M linearly-independent vectors.

For example, ignoring spin, the $n = 2$ eigenenergy of the non-relativistic Hamiltonian of atomic hydrogen is 4-fold degenerate since (1) the 2s, $2\mathrm{p}_{m=0}$, $2\mathrm{p}_{m=1}$ and $2p_{m=-1}$ wave functions all belong to this eigenenergy, (2) these four wave functions are mutually orthogonal and therefore linearly independent, and (3) it is not possible to find a fifth $n = 2$ energy eigenfunction that would be orthogonal to all these four functions.

The words "linearly independent" are an important part of the definition above. If v is an eigenvector of an operator A with eigenvalue λ, then any multiple of v is also an eigenvector of A with that same eigenvalue since, for any scalar c,

$$A(cv) = cAv = c\lambda v = \lambda(cv). \tag{3.70}$$

Therefore it is always the case that infinitely many eigenvectors belong to a same eigenvalue. However, vectors multiple of each other are not linearly independent. (As we will see, in Quantum Mechanics vectors multiple of each other represent the same quantum state, while linearly independent vectors necessarily represent different states. An eigenvalue is degenerate if it corresponds to several *different* quantum states.)

> ☞ Suppose that the operator A appearing in Eq. (3.69) acts in a vector space V. The M linearly-independent eigenvectors v_1, v_2,…, v_M then span a M-dimensional subspace of V, called an invariant subspace, whose elements are transformed by A into elements of the same subspace.

The spectrum of an operator

In Physics, the set of all the eigenvalues of an operator is usually called the spectrum (or the eigenvalue spectrum) of that operator.

> ☞ In Mathematics, however, the spectrum of an operator is defined as the set of the scalars λ for which the operator $A - \lambda I$ is singular, where I is the identity operator. One can show that these two definitions are equivalent for finite-dimensional vector spaces. However, for operators acting in infinite-dimensional spaces, it is possible that the operator $A - \lambda I$ is singular at values of λ which are not eigenvalues of A. (This operator is always singular at the eigenvalues of A, since, if v is an eigenvector of A with eigenvalue λ, $(A - \lambda I)v = 0$ although $v \neq 0$.)

3.7 The adjoint of an operator

The adjoint of an operator A is the operator $A^\dagger$ such that

$$(v, Aw) = (w, A^\dagger v)^* \tag{3.71}$$

for any vector v and w. The symbol †, pronounced "dagger", is traditionally used in Physics to denote the adjoint of an operator.

> ☞ Saying, as above, that Eq. (**??**) must apply to any vector v and w of the Hilbert space in which the operator A acts is unproblematic for finite-dimensional Hilbert spaces. For infinite-dimensional spaces, however, a mathematically sound definition of the adjoint requires a careful specification of the domains of the operators A and $A^\dagger$. One says that $A^\dagger$ is the adjoint of A if $(v, Aw) = (w, A^\dagger v)^*$ for any vector w in the domain of A and any vector v in the domain of $A^\dagger$, the latter being defined as the set of all the vectors v such that there exists a vector v_A for which $(v, Aw) = (v_A, w)$ for any vector w in the domain of A.
>
> ☞ It can be shown that any operator has one and only one adjoint.
>
> ☞ The adjoint of an operator is also called the Hermitian conjugate of this operator.

The following theorems are easily proven:

1. The adjoint of a sum is the sum of the adjoints:
$$(A + B)^\dagger = A^\dagger + B^\dagger. \tag{3.72}$$
2. Scalars multiplying operators get complex conjugated in the adjoints: if α is a complex number,
$$(\alpha A)^\dagger = \alpha^* A^\dagger. \tag{3.73}$$
3. The adjoint of a product of two operators is the product of their adjoints *in reverse order*:
$$(AB)^\dagger = B^\dagger A^\dagger. \tag{3.74}$$
4. The adjoint of the adjoint of an operator is this operator:
$$(A^\dagger)^\dagger = A. \tag{3.75}$$

☞ The ladder operators a_+ and a_- used in the Term 1 course for calculating the energy levels of a linear harmonic oscillator are examples of an operator and its adjoint: As we will see later on, $a_- = a_+^\dagger$ and $a_+ = a_-^\dagger$.

☞ If the operator A is represented by the matrix A in an orthonormal basis $\{|u_1\rangle, |u_2\rangle, \ldots, |u_N\rangle\}$, its adjoint is represented in this basis by the conjugate tranpose matrix, $\mathsf{A}^\dagger$. (The proof of this assertion is left as an exercise.) For example, if an operator is represented by the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

its adjoint is represented by the matrix

$$\begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}.$$

Of particular importance in Quantum Mechanics are operators that are identical to their adjoint. We will explore their properties later in these notes.