3 Linear operators

Throughout this section I assume that \( V \) is an \( n \)-dimensional inner-product space and that \( |e_1\rangle, \ldots, |e_n\rangle \) are an orthonormal basis for this space. Recall that linear operators are mappings from a vector space \( V \) to itself that satisfy the conditions (1.23). Let \( A \) be a linear operator and suppose that

\[
|b\rangle = A|a\rangle. \tag{3.1}
\]

In §1.4 we saw how this is equivalent to the matrix equation \( b = Aa \), where \( a \) and \( b \) are the column vectors that represent \( |a\rangle \) and \( |b\rangle \) respectively. Armed with our inner product we can now obtain an explicit expression for the elements of the matrix representing the operator \( A \). Expanding \( |a\rangle = \sum_{i=1}^n a_i |e_i\rangle \) and \( |b\rangle = \sum_{i=1}^n b_i |e_i\rangle \), equation (3.1) becomes

\[
\sum_{i=1}^n b_i |e_i\rangle = \sum_{i=1}^n a_i A|e_i\rangle. \tag{3.2}
\]

Now choose any bra \( \langle e_j | \) and apply it to both sides:

\[
\langle e_j | \sum_{i=1}^n b_i |e_i\rangle = \langle e_j | \sum_{i=1}^n a_i A|e_i\rangle \tag{3.3}
\]

Therefore equation (3.1) can be represented by the matrix equation

\[
b_j = \sum_{k=1}^n A_{jk} a_k, \tag{3.4}
\]

where \( A_{jk} = \langle e_j | A |e_k\rangle \) are the matrix elements of the operator \( A \) in the \( |e_1\rangle, \ldots, |e_n\rangle \) basis.

3.1 The identity operator

The identity operator \( I \) defined through \( I|v\rangle = |v\rangle \) for all \( |v\rangle \in V \) is clearly a linear operator. Less obviously, it can be written as

\[
I = \sum_{i=1}^n |e_i\rangle \langle e_i|. \tag{3.5}
\]

This is sometimes known as resolution of the identity.

Proof: Any \( |v\rangle \in V \) can be expressed as \( |v\rangle = \sum_{i=1}^n \alpha_i |e_i\rangle \). Using the expression (3.5) for \( I \), we have that

\[
I|v\rangle = \sum_{i=1}^n |e_i\rangle \sum_{j=1}^n \alpha_j \langle e_j | e_i \rangle = \sum_{i=1}^n \sum_{j=1}^n \alpha_j \langle e_j | e_i \rangle |e_i\rangle.
\]

\[
I|v\rangle = \sum_{i=1}^n |e_i\rangle \sum_{j=1}^n \alpha_j |e_j\rangle = \sum_{i=1}^n |e_i\rangle \alpha_i |e_i\rangle = |v\rangle.
\]

The individual terms \( |e_i\rangle \langle e_i| \) appearing in the sum (3.5) are known as projection operators: if we apply \( P_i = |e_i\rangle \langle e_i| \) to a vector \( |a\rangle = \sum_j a_j |e_j\rangle \) the result \( P_i |a\rangle = a_i |e_i\rangle \). Similarly, \( \langle a | P_i |b\rangle = \langle e_i | a |e_i\rangle \). We have already seen a use of projection operators in the Gram–Schmidt procedure earlier (equation 2.21).

3.2 Combining operators

The composition of two linear operators \( A \) and \( B \) is another linear operator. In case it is not obvious how to show this, let us write \( C = AB \) for the result of applying \( B \) first, then \( A \). Now notice that \( C \) is a mapping from \( V \) to \( V \) and that conditions (1.23)

\[
C(|a\rangle + |b\rangle) = A(B(|a\rangle + |b\rangle)) = A(B|a\rangle + B|b\rangle) = C|a\rangle + C|b\rangle,
C(|a\rangle B|a\rangle) = A(a B|a\rangle) = a(A B|a\rangle) = aC|a\rangle
\]

hold for any \( |a\rangle, |b\rangle \in V \) and \( a \in \mathcal{F} \).

Exercises: Show that the matrix representing \( C \) is identical to the matrix obtained by multiplying the matrix representations of the operators \( A \) and \( B \).

Exercises: Show that sum of two linear operators is another linear operator.

In general \( AB \neq BA \): the order of composition matters. The difference

\[
AB - BA \equiv [A,B]
\]

is another linear operator, known as the commutator of \( A \) and \( B \).

Functions of operators

We can construct new linear operators by composition and addition. For example, given a linear operator \( A \), let us define exp \( A \) in the obvious way as

\[
\exp A = \lim_{N \to \infty} \left( I + \frac{1}{N} A \right)^N = I + A + \frac{1}{2!} A^2 + \frac{1}{3!} A^3 + \cdots + \frac{1}{m!} A^m + \cdots,
\]

in which \( I \) is the identity operator. \( \exp A \) is another linear operator, but notice that it is a mapping from \( V \) to itself and that for any \( |a\rangle, |b\rangle \in V \) and \( a \in \mathcal{F} \) we have that

\[
\langle a | \exp A |b\rangle = \langle a | A |b\rangle + (\exp A) |a\rangle \langle a | b\rangle,
\]

\[
\langle a | (\exp A) |a\rangle = a \langle a | (\exp A) |a\rangle.
\]

\[
\exp A |a\rangle + |b\rangle
\]

\[
(\exp A) |a\rangle + (\exp A) |b\rangle,
\]

\[
\exp A (\alpha |a\rangle) = \alpha (\exp A) |a\rangle.
\]

\[
(\exp A) (\alpha |a\rangle) = \alpha (\exp A) |a\rangle.
\]
Example: what is $\exp(\alpha G)$, where $G = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$?

First note that $G^2 = -I$. Therefore $G^{2n} = (-1)^n I$ and $G^{2n+1} = (-1)^n G$. We can use this to split the expansion of the exponential into a sum of even and odd terms:

$$\exp(\alpha G) = \sum_{n=0}^{\infty} \frac{(\alpha G)^n}{n!}$$

$$= \sum_{n=0}^{\infty} \frac{1}{(2n)!} (-1)^n \alpha^{2n} I + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} (-1)^n \alpha^{2n+1} G$$

$$= \sum_{n=0}^{\infty} \frac{1}{(2n)!} \alpha^{2n} I + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \alpha^{2n+1} G$$

$$= \cos \alpha I + \sin \alpha G$$

For this reason $G$ is known as the generator of two-dimensional rotations: for $\epsilon \ll 1$ the operator $I + \epsilon G$ is a rotation by an angle $\epsilon$; from the definition (3.9) the operator $\exp(\alpha G)$ is obtained by chaining together many such infinitesimal rotations.

### 3.3 Special types of matrix

Let $A$ be an $n \times n$ matrix with elements $A_{ij}$. The transpose $A^T$ is obtained from $A$ by swapping rows and columns. It has elements

$$(A^T)_{ij} = A_{ji}. \quad (3.12)$$

The Hermitian conjugate $A^\dagger$ is obtained by swapping rows and columns and taking the complex conjugate. It has elements

$$(A^\dagger)_{ij} = A^{\ast}_{ji}. \quad (3.13)$$

**Exercises:** Show that $(AB)^T = B^T A^T$ and $(AB)^\dagger = B^\dagger A^\dagger$.

The following table lists the conditions that are satisfied by members of some important families of matrices:

<table>
<thead>
<tr>
<th>Symmetric: $A = A^T$</th>
<th>Hermitian: $A = A^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orthogonal: $A A^T = I$</td>
<td>Unitary: $A A^\dagger = I$</td>
</tr>
<tr>
<td>[Normal: $A^\dagger A = A A^\dagger$]</td>
<td></td>
</tr>
</tbody>
</table>

Hermitian matrices are sometimes also known as self-adjoint matrices. Recall that orthogonal matrices represent reflections and rotations.

It turns out (§3.5 below) that the conditions for a matrix to be Hermitian or unitary are independent of (orthonormal) basis: if the matrix representing a particular linear operator is Hermitian (unitary) in one basis then it is Hermitian (unitary) in all bases and one can sensibly say that the underlying operator itself is Hermitian (unitary).

For the special case of real vector spaces this implies that if a matrix is symmetric (orthogonal) in one basis then it is symmetric (orthogonal) in all bases.

**Exercises:** Show that the product of two unitary matrices is another unitary matrix. Does the same result hold for Hermitian matrices? If not, find a condition under which it does hold.

### 3.4 Dual of an operator

Given an operator $A$, let $|b\rangle = A(a)$ be the result of applying $A$ to an arbitrary ket $|a\rangle$. How are the corresponding dual vectors $|a\rangle$ and $|b\rangle$ related?

Expanding $|b\rangle = A(a)$ in terms of components we have that

$$\sum_{k=1}^{n} b_k |e_k\rangle = \sum_{j=1}^{n} a_j \epsilon_{jk} |e_j\rangle.$$

(3.14)

Using (2.13) to obtain the dual of each side of this equation gives

$$\sum_{k=1}^{n} \epsilon_{kj} b^*_k = \sum_{j=1}^{n} a^*_j \epsilon^*_{jk} |e_j\rangle = \sum_{j=1}^{n} a^*_j (A^\dagger)^* |e_j\rangle.$$

(3.15)

The LHS is clearly just $\langle a | A^\dagger$. The RHS is $\langle a | A^\dagger| e\rangle$. To see this, apply the RHS to an arbitrary basis ket $|e_k\rangle$. The result is $\sum_{j=1}^{n} a^*_j (A^\dagger)_{kj}$. The result of operating on the same ket with $|a\rangle$ is

$$\langle a | A^\dagger | e\rangle = \sum_{j=1}^{n} a^*_j \epsilon_{jk} (A^\dagger)^* |e_j\rangle.$$

(3.16)

Therefore, if $|b\rangle = A(a)$ then the corresponding dual vectors are related via $|b\rangle = \langle a | A^\dagger$: the dual to the operator $A$ is its Hermitian conjugate $A^\dagger$.

**Exercises:** Use property (2.3) of the inner product to show that $\langle a | A^\dagger (a^\dagger | b\rangle$.

### 3.5 Change of basis

Bases are not unique. Sometimes it is convenient to do parts of a calculation in one basis and the rest in another. Therefore it is important to know how to transform vector coordinates and matrix elements from one basis to another.

Suppose we have two different orthonormal basis sets, $\{|e_1\rangle, \ldots, |e_n\rangle\}$ and $\{|e'_1\rangle, \ldots, |e'_n\rangle\}$, with

$$|e'_k\rangle = \sum_{j=1}^{n} U_{kj} |e_j\rangle. \quad (3.17)$$

That is, the $i^{th}$ column of the matrix $U$ gives the representation of $|e'_i\rangle$ in the unprimed basis. The corresponding relationship for the basis bra is

$$\langle e'_i | = \sum_{j=1}^{n} \epsilon_{ij} U^\dagger_{ji} |e_j\rangle. \quad (3.18)$$

The orthonormality of the bases places constraints on the transformation matrix $U$.

$$\delta_{ij} = \langle e'_i | e'_j \rangle = \sum_{k=1}^{n} \langle e'_i | U^\dagger_{ik} |e_k\rangle = \sum_{k=1}^{n} U^\dagger_{ik} U_{kj} = \sum_{k=1}^{n} (U^\dagger U)_{ik} U_{kj}. \quad (3.19)$$
or, in matrix form, \( I = UU^\dagger \): the matrix \( U \) describing the coordinate transformation must be unitary: unitary matrices are the generalization of real orthogonal transformations (i.e., rotations and reflections) to complex vector spaces.

**Transformation of vector components** Taking an arbitrary vector \( |a\rangle = \sum_{j=1}^n a_j |e_j\rangle = \sum_{j=1}^n U_{ij} |e_j\rangle \) and using \( |e_j\rangle \) to project \(|a\rangle\) along the \( i^{th} \) primed basis vector gives

\[
a_i' = \langle e_i|a\rangle = \sum_{j=1}^n \langle e_i|e_j\rangle U_{ij} = \sum_{j=1}^n U_{ij} a_j = \sum_{j=1}^n (U^\dagger)_{ij} a_j, \tag{3.20}
\]

In matrix form, the components in the primed basis are given by \( a' = U^\dagger a \).

**Transformation of matrix elements** In the primed basis, the operator \( A \) has matrix elements

\[
A'_{ij} = \langle e_i'|A|e_j\rangle = \sum_{k=1}^n \langle e_k'|A|e_k\rangle = \sum_{k=1}^n U_{ik} A_{kj} U_{jk} = \sum_{k=1}^n (U^\dagger A U)_{ik} (U^\dagger A U)_{jk}, \tag{3.21}
\]

so that \( A' = U^\dagger A U \). What does this mean? When applying the matrix \( A' \) to a vector whose coordinates \( a' \) are given with respect to the \( |e_j\rangle \) basis, think of \( U \) as transforming from \( a' \) to \( a \). Then we apply the matrix \( A \) to the result before transforming back to the primed basis with \( U^\dagger \).

**Exercises:** Derive the change-of-basis formulae (3.20) and (3.21) by resolving the identity (3.5). That is, write

\[
a_i' = \langle e_i'|a\rangle = \langle e_i'|I|a\rangle = \sum_{j=1}^n \langle e_i'|e_j\rangle A_{ji} a_j = \sum_{j=1}^n \langle e_i'|e_j\rangle U_{ij} a_j = \sum_{j=1}^n \langle e_i'|e_j\rangle U_{ij} a_j = \sum_{j=1}^n (U^\dagger A U)_{ij} a_j, \tag{3.22}
\]

and use the fact that \( I \) can be expressed as \( I = \sum |e_k\rangle \langle e_k| = \sum |e_k\rangle \langle e_k|e_k\rangle \). You should find that

\[
a_{i}' = \sum_{j} \langle e_j'|e_i\rangle a_j = \sum_{j} \langle e_j'|e_i\rangle U_{ij} a_j, \tag{3.23}
\]

where \( \langle e_j'|e_i\rangle \) is the projection of \( |e_i\rangle \) onto the \( |e_j\rangle \) basis — that is, a matrix whose \( j^{th} \) column expresses \( |e_i\rangle \) in the primed basis. How is this matrix related to \( U \) introduced above?

**Exercises:** Justify the statement at the end of §3.3: show that if a matrix is Hermitian (unitary) in one orthonormal basis then it is Hermitian (unitary) in any other orthonormal basis.

**Example:** Two-dimensional rotations Suppose that the \( \{|e'_1\rangle, |e'_2\rangle\} \) basis is related to the \( \{|e_1\rangle, |e_2\rangle\} \) basis by a rotation through an angle \( \alpha \). From the diagram, the basis vectors are related through

\[
\begin{align*}
|e'_1\rangle &= \cos \alpha |e_1\rangle + \sin \alpha |e_2\rangle \\
|e'_2\rangle &= -\sin \alpha |e_1\rangle + \cos \alpha |e_2\rangle
\end{align*}
\]

\[
\Rightarrow \quad \left( \begin{array}{c} |e'_1\rangle \\ |e'_2\rangle \end{array} \right) = U \left( \begin{array}{c} |e_1\rangle \\ |e_2\rangle \end{array} \right), \tag{3.24}
\]

with

\[
U = \begin{pmatrix}
\cos \alpha & \sin \alpha \\
-\sin \alpha & \cos \alpha
\end{pmatrix}, \quad \text{so that} \quad U^\dagger = \begin{pmatrix}
\cos \alpha & -\sin \alpha \\
\sin \alpha & \cos \alpha
\end{pmatrix}. \tag{3.25}
\]

The \( i^{th} \) column of \( U \) expresses \( |e_i\rangle \) in the \( |e_j\rangle \) basis: \( U_{ij} = \langle e_j|e_i\rangle \). Clearly the coordinates \( a, a' \) of a vector \( |a\rangle \) in the two bases are related through

\[
a' = U^\dagger a. \]

### 3.6 Rank

The range of a linear operator is the set of all possible output vectors:

\[
\text{range } A = \{ A|v\rangle : |v\rangle \in V \}. \tag{3.26}
\]

The rank is clearly the subspace of \( V \) spanned by the vectors \( A|e_1\rangle, A|e_2\rangle, \ldots, A|e_n\rangle \). The rank of an operator is the dimension of its range.

**Examples:**

\[
\begin{align*}
\text{rank } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} &= 3, \quad \text{rank } \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} &= 2, \quad \text{rank } \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 2 & 1 \end{pmatrix} &= 1. \tag{3.27}
\end{align*}
\]

The rank of a matrix is unchanged under any of the following *elementary row operations*:

- swap two rows;
- multiply one row by a non-zero constant;
- add a multiple of one row to another.

Each such operation produces a new matrix whose columns are linear combinations of the columns of the original matrix.

A simple way of calculating the rank of a matrix is to use elementary row operations to reduce it to an upper triangular matrix and counting the number of LI columns in the result.

Every elementary row operation can be expressed as an *elementary matrix*. For example, starting with a \( 3 \times 3 \) matrix \( A \) and adding \( \alpha \) times row 3 to row 2, then swapping the first two rows gives a new matrix \( E_2 E_1 A \), where the elementary matrices \( E_1 \) and \( E_2 \) are

\[
E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.28}
\]

**Aside: linear equations** To solve the linear equation \( Ax = b \) apply elementary row operations \( E_1, E_2, \ldots \) to \( A \) to reduce it to upper triangular form. That is,

\[
\begin{align*}
Ax &= b \\
\Rightarrow (E_m \cdots E_2 E_1) Ax &= (E_m \cdots E_2 E_1) b \\
\Rightarrow \begin{pmatrix}
A_{11} & A_{12} & A_{13} & \cdots & A_{1n-1} & A_{1n} \\
0 & A_{22} & A_{23} & \cdots & A_{2n-1} & A_{2n} \\
0 & 0 & A_{33} & \cdots & A_{3n-1} & A_{3n} \\
& \vdots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \cdots & A_{n-1,n} & A_{n,n}
\end{pmatrix} \begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
\vdots \\
x_n
\end{pmatrix} &= \begin{pmatrix}
b_1' \\
b_2' \\
b_3' \\
\vdots \\
b_n'
\end{pmatrix}. \tag{3.29}
\end{align*}
\]

where \( A_{ij}' = (E_m \cdots E_2 E_1) A_{ij} \) and \( b_i' = (E_m \cdots E_2 E_1) b_i \). Then the \( x_i \) can be found by backsubstitution.
### 3.7 Determinant

**Permutations** A permutation of the list $(1, 2, \ldots, m)$ is another list that contains each of the numbers $1, 2, \ldots, m$ exactly once. In other words, it is a straightforward shuffling of the order of the elements. There are $m!$ permutations of an $m$-element list.

Given a permutation $P$ we write $P(i)$ for the first element in the shuffled list, $P(2)$ for the second, etc. Then $P$ can be written as $(P(1), P(2), \ldots, P(m))$. An alternative notation is

$$P = \begin{pmatrix}
1 & 2 & \cdots & m \\
P(1) & P(2) & \cdots & P(m)
\end{pmatrix},$$

(3.30)

which emphasizes that $P$ is a mapping from the set $(1, 2, \ldots, m)$ (top row) to itself (values given on bottom row). From any two permutation mappings $P$ and $Q$ we can compose a new one $PQ$ defined through $(PQ)(i) = P(Q(i))$. There is an identity mapping for which $P(i) = i$ and every $P$ has an inverse

$$P^{-1} = \begin{pmatrix}
P(1) & P(2) & \cdots & P(m) \\
1 & 2 & \cdots & m
\end{pmatrix},$$

(3.31)

which is well defined because each number $1, 2, \ldots, m$ appears exactly once in the top row of (3.31).

Any permutation $(P(1), P(2), \ldots, P(m))$ can be constructed from $(1, 2, \ldots, m)$ by a sequence of pairwise element exchanges. **Even (odd) permutations** require an even (odd) number of exchanges. The **sign** of a permutation is defined as

$$\text{sgn}(P) = \begin{cases} +1, & \text{if } P \text{ is an even permutation of } (1, 2, \ldots, m), \\ -1, & \text{if } P \text{ is an odd permutation of } (1, 2, \ldots, m). \end{cases}$$

(3.32)

Given two permutations $P$, $Q$, we have $\text{sgn}(PQ) = \text{sgn}(P)\text{sgn}(Q)$. The identity permutation is even. Therefore $+1 = \text{sgn}(P^{-1}) = \text{sgn}(P^{-1})\text{sgn}(P)$, showing that $\text{sgn}(P^{-1}) = \text{sgn}(P)$.

The following table shows all 6 permutations of the 3-elements list $(1, 2, 3)$:

<table>
<thead>
<tr>
<th>$P(1)$</th>
<th>$P(2)$</th>
<th>$P(3)$</th>
<th>$\text{sgn}(P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$P_2$</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$P_3$</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>$P_4$</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$P_5$</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$P_6$</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

**Determinant** The determinant of an $n \times n$ matrix $A$ is defined through

$$\det A = \sum_{P} \text{sgn}(P)A_{P(1)}A_{P(2)}\cdots A_{P(n)},$$

(3.33)

where the sum is over all permutations $P$ of $(1, 2, 3, \ldots, n)$. For example, for the case $n = 3$

$$\det A = A_{11}A_{22}A_{33} - A_{12}A_{21}A_{33} + A_{12}A_{23}A_{31} - A_{13}A_{22}A_{31} + A_{13}A_{21}A_{32} - A_{11}A_{23}A_{32},$$

(3.34)

where in the sum (3.33) I have taken the permutations $P$ in the order given in the table above.

**Exercise:** Confirm that this expression agrees with the result obtained by the more familiar minor-expansion method.

### Properties of the determinant

In the following $A$ and $B$ are any two square matrices.

(i) If $A$ is $n$ by $n$, then $\det A = 0$.

(ii) If $B$ can be obtained from $A$ by multiplying one row by a factor $\alpha$ then $\det B = \alpha \det A$.

(iii) If $B$ can be obtained from $A$ by adding a multiple of one row to another then $\det B = \det A$.

(iv) If $B$ can be obtained from $A$ by swapping two rows then $\det B = -\det A$.

(v) $\det(A^T) = \det A$.

(vi) $\det(AB) = \det A \det B$.

**Proof of (i):** Suppose that the first two rows of $A$ are equal. $A_{1j} = A_{2j}$, and consider the factor $F(P)$ below in each term of the sum

$$\det A = \sum_{P} \text{sgn}(P)A_{1P(1)}A_{2P(2)}\cdots A_{nP(n)}.$$

(3.35)

Since $A_{1j} = A_{2j}$, we have that $F(P) = \text{sgn}(P)A_{1P(1)}A_{2P(2)}\cdots A_{nP(n)}$. For each permutation $P$ there is another, $P'$, that swaps only $P(1)$ and $P(2)$, leaving the other elements unchanged. This new permutation has $F(P') = \text{sgn}(P')A_{1P(1)}A_{2P(2)}\cdots A_{nP(n)} = -F(P)$. So, the term involving $P'$ cancels out the contribution of the term involving $P$ to the sum (3.35). Summing over all $P$ the result is $\det A = 0$.

**Proof of (ii):** Without loss of generality, suppose that

$$B = \begin{pmatrix}
\alpha A_{11} & \alpha A_{12} & \cdots & \alpha A_{1n} \\
A_{21} & A_{22} & \cdots & A_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
A_{n1} & A_{n2} & \cdots & A_{nn}
\end{pmatrix}$$

(3.36)

Then

$$\det B = \sum_{P} \text{sgn}(P)B_{1P(1)}B_{2P(2)}\cdots B_{nP(n)}$$

$$= \sum_{P} \text{sgn}(P)(\alpha A_{1P(1)})A_{2P(2)}\cdots A_{nP(n)}$$

$$= \alpha \sum_{P} \text{sgn}(P)A_{1P(1)}A_{2P(2)}\cdots A_{nP(n)} = \alpha \det A.$$  

(3.37)

**Proof of (iii):** Suppose that

$$B = \begin{pmatrix}
A_{11} + \alpha A_{21} & A_{12} + \alpha A_{22} & \cdots & A_{1n} + \alpha A_{2n} \\
A_{21} & A_{22} & \cdots & A_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
A_{n1} & A_{n2} & \cdots & A_{nn}
\end{pmatrix}$$

(3.38)

Then

$$\det B = \sum_{P} \text{sgn}(P)\left[A_{1P(1)} + \alpha A_{2P(1)}\right]A_{2P(2)}\cdots A_{nP(n)}$$

$$= \sum_{P} \text{sgn}(P)A_{1P(1)}A_{2P(2)}\cdots A_{nP(n)} + \alpha \sum_{P} \text{sgn}(P)A_{2P(1)}A_{2P(2)}\cdots A_{nP(n)}.$$

(3.39)

The second term is zero by property (i) above.
Proof of (iv): Suppose, for example, that $B$ is obtained from $A$ by swapping the first two rows. Then

$$
det B = \sum_{P} \text{sgn}(P) A_{P(1)} A_{P(2)} \cdots A_{P(n)}
$$

$$= \sum_{P} \text{sgn}(P) A_{P(2)} A_{P(3)} \cdots A_{P(Q)} A_{P(Q)}
$$

where $Q$ is a permutation that (un)wraps the first two rows and leaves the rest unchanged: $Q(1) = 2$, $Q(2) = 1$ and $Q(i) = i$ for $i > 2$. The sum over $P$ can be replaced by another sum over $P'$ with $\text{sgn}(P') = -\text{sgn}(P)$. Therefore

$$
det B = - \sum_{P'} \text{sgn}(P') A_{P'(1)} A_{P'(2)} \cdots A_{P'(n)}
$$

$$= - \det A.
$$

Proof of (v):

$$
\det(A^T) = \sum_{P} \text{sgn}(P) (A^T)_{P(1)} (A^T)_{P(2)} \cdots (A^T)_{P(n)}
$$

$$= \sum_{P} \text{sgn}(P) A_{P(1)} A_{P(2)} \cdots A_{P(n)}
$$

$$= \sum_{P} \text{sgn}(P) \prod_{i=1}^{n} A_{P(i,i)}
$$

Let $Q$ be the inverse permutation to $P$: if $P(i) = j$, then $Q(j) = i$. Then $A_{P(i,j)} = A_{Q(j,i)}$ and we have that $\prod_{i=1}^{n} A_{P(i,i)} = \prod_{i=1}^{n} A_{Q(i,i)}$. Noting also that $\text{sgn}(P) = \text{sgn}(Q)$ and replacing the sum over $P$ by an equivalent sum over $Q$, obtain

$$
\sum_{P} \text{sgn}(P) \prod_{i=1}^{n} A_{P(i,i)} = \sum_{Q} \text{sgn}(Q) \prod_{i=1}^{n} A_{Q(i,i)} = \det A.
$$

Proof of (vi): homework exercise!

Exercises: Use the properties above together with results derived in §3.5 to show that the determinant is independent of basis.

3.8 Trace

The trace $\text{tr} A$ of an $n \times n$ matrix $A$ is the sum of its diagonal elements:

$$
\text{tr} A = \sum_{i=1}^{n} A_{ii}.
$$

The trace satisfies

$$
\text{tr}(AB) = \text{tr}(BA)
$$

because

$$
\text{tr}(AB) = \sum_{i=1}^{n} \left( \sum_{j=1}^{n} A_{ij} B_{ji} \right) = \sum_{j=1}^{n} \left( \sum_{i=1}^{n} B_{ji} A_{ij} \right) = \text{tr}(BA).
$$

Taking $A = A_1$, $B = A_2 A_3 \cdots A_n$ we see that $\text{tr}(A_1 A_2 \cdots A_n) = \text{tr}(A_2 \cdots A_n A_1)$: the trace is invariant under cyclic permutations.

Exercises: Show that the trace is independent of basis. Explain then how, given a $3 \times 3$ rotation matrix, one can find the rotation angle directly from the trace.

3.9 Eigenvectors and diagonalisation

Recall that $|v\rangle \neq 0$ is an eigenvector of an operator $A$ with eigenvalue $\lambda$ if it satisfies the eigenvalue equation

$$
A|v\rangle = \lambda|v\rangle.
$$

To find $|\varphi\rangle$ we first find $\lambda$ by rewriting eq above as

$$
A|\varphi\rangle - \lambda|\varphi\rangle = 0.
$$

Clearly $A - \lambda I$ can't be invertible: if it were, then we could operate on 0 with $(A - \lambda I)^{-1}$ to get $|\varphi\rangle$. So, we must have that

$$
det(A - \lambda I) = 0.
$$

which is known as the characteristic equation for $A$. The characteristic equation is an $n$th-order polynomial in $\lambda$ which can always be written in the form $(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) = 0$, where the $n$ roots (i.e., eigenvalues) $\lambda_1, \ldots, \lambda_n$ are in general complex and not necessarily distinct. If a particular value of $\lambda$ appears $k > 1$ times then that eigenvalue is said to be $k$-fold degenerate. (The integer $k$ is also sometimes called the multiplicity of the eigenvalue.)

Exercises: Let $A$ be a linear operator with eigenvector $|v\rangle$ and corresponding eigenvalue $\lambda$. Show that $|v\rangle$ is also an eigenvector of the linear operator $\exp(A)$, but with eigenvalue $\exp(\lambda)$.

Example: Example: find the eigenvalues and eigenvectors of

$$
A = \begin{pmatrix}
1 & 0 & 1 \\
0 & 2 & 0 \\
1 & 0 & 1
\end{pmatrix}
$$

The characteristic equation is $(1 - \lambda)(2 - \lambda)(1 - \lambda) - (2 - \lambda) = 0$, which, when factorised, is $(\lambda - 2)^2 \lambda = 0$. Therefore the eigenvalues are $\lambda = 0, 2, 2$: the eigenvalue $\lambda = 2$ is doubly degenerate.

To find the eigenvector $|v_1\rangle$ corresponding to the first eigenvalue with $\lambda = \lambda_1 = 0$, take $|v_1\rangle = |x_1, x_2, x_3\rangle^T$ and substitute into eigenvalue equation (3.47) to find $x_1 = -x_2$ and $x_2 = 0$. Any vector $|v\rangle$ that satisfies these constraints is an eigenvector of $A$ with eigenvalue 0. For example, we could take $|v_2\rangle = (1, 0, -1)^T$ or $(-1, 0, 0)^T$ or even $i |v_0\rangle = (0, 0, -i)^T$ (assuming we have a complex vector space). It is usually most convenient, however, to choose to make $|v\rangle$ normalized. Therefore we choose $|v_3\rangle = |1, 0, -1\rangle^T$.

Taking $\lambda = \lambda_2 = \lambda_3 = 2$ and substituting $v = |x_1, x_2, x_3\rangle^T$ into the eigenvalue equation yields the constraints $-x_1 + x_3 = 0$ and $x_1 - x_3 = 0$. So, we must set $x_1 = x_3$, but we are free to choose anything for $x_2$. For $|v_2\rangle$ let us take $x_1 = x_2 = \sqrt{2}$ and $x_2 = 0$, while for $|v_3\rangle$ we choose $x_1 = x_2 = 0$ and $x_3 = 1$. To summarise, we have found the following eigenvalues and eigenvectors for the matrix $A$:

$$
\lambda_1 = 0, \quad |v_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \lambda_2 = 2, \quad |v_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_3 = 2, \quad |v_3\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
$$

Diagonalisation Notice that for the $3 \times 3$ matrix $A$ above we were able to find three eigenvectors $|v_1\rangle, |v_2\rangle, |v_3\rangle$. These three eigenvectors turn out to be orthogonal. Therefore they can be used as a basis for the 3-dimensional vector space on which $A$ operates. In this eigenbasis the matrix representing the operator $A$ takes on a particularly simple form: it is diagonal, with matrix elements

$$
\langle n| A |m\rangle = \lambda_n \delta_{m,n}.
$$
3.10 The complex spectral theorem

A fundamental problem of linear algebra is to find the conditions under which a given matrix $A$ can be diagonalised, i.e., whether there exists an invertible coordinate transformation $P$ for which $P^{-1}AP$ becomes diagonal. For our purposes, we need only accept that a sufficient condition is that $A$ be Hermitian (see next section).

If you don’t like to accept such things, let $A$ be an operator on a complex inner-product space $V$. It is relatively easy to show that $V$ has an orthonormal basis consisting of eigenvectors of $A$ if and only if $A$ is normal. Recall that normal matrices satisfy $AA^\dagger = A^\dagger A$. Hermitian and unitary matrices are special cases of normal matrices.

Here is an outline sketch of the proof. Let $|v_1\rangle$, $|v_2\rangle$, $|v_3\rangle$ be the eigenvectors of $A$ and $\lambda_1$, $\lambda_2$, $\lambda_3$ the corresponding eigenvalues.

First suppose that $V$ has an orthonormal basis consisting of eigenvectors of $A$. In this case the matrix representing $A$ is diagonal with matrix elements $\lambda_i = \langle v_i | A | v_i \rangle$. The matrix representing $A$ has elements $A_{ij} = \langle v_i | A | v_j \rangle = \langle v_j | A^\dagger | v_i \rangle$, which is also diagonal. Diagonal matrices commute. Therefore $AA^\dagger = A^\dagger A$. $A$ is normal.

Now the converse: we need to show that if $A$ is normal then $V$ has an orthonormal basis consisting of eigenvectors of $A$. Any matrix can be reduced to upper triangular form by applying an appropriate set of elementary row operations (see §3.6).

In particular, there is an orthonormal basis $|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle$ in which the matrix representing $A$ becomes

$$
\langle \psi_i | A | \psi_j \rangle = \begin{pmatrix}
0 & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{n1} & \cdots & 0
\end{pmatrix}.
$$

From (3.53) we see that

$$
\| A | \psi_i \rangle \|^2 = |a_{1i}|^2
$$

and

$$
\| A^\dagger | \psi_i \rangle \|^2 = |a_{i1}|^2 + |a_{i2}|^2 + \cdots + |a_{in}|^2.
$$

But $A$ is normal, so $\| A | \psi_i \rangle \|^2 = \| A^\dagger | \psi_i \rangle \|^2$. This means that all entries in the first row of the matrix in the RHS of (3.53) vanish, except possibly for the first.

Similarly, from equation (3.53) we have that

$$
\| A | \psi_i \rangle \|^2 = |a_{1i}|^2 + |a_{2i}|^2 + \cdots + |a_{ni}|^2.
$$

using the result that $a_{ij} = 0$ from the previous paragraph. Because $A$ is normal we have that $\| A | \psi_i \rangle \|^2 = \| A^\dagger | \psi_i \rangle \|^2$; the whole second row must equal zero, except possibly for the diagonal element $a_{ii}$.

Repeating this procedure for $A | \psi_1 \rangle$ to $A | \psi_n \rangle$ shows that all of the off-diagonal elements in (3.53) must be zero. Therefore the orthonormal basis vectors $|\psi_1\rangle, \ldots, |\psi_n\rangle$ are eigenvectors of $A$.

The corresponding result for real vector spaces is harder to prove. It states that a real vector space $V$ has an orthonormal basis consisting of eigenvectors of $A$ if and only if $A$ is symmetric (which is equivalent to Hermitian for the special case of a real vector space).

3.11 Hermitian operators

Hermitian operators are particularly important. Here are two central results. If $A$ is Hermitian then:

1. its eigenvalues are real
2. its eigenvectors are orthogonal (and therefore form a basis of $V$)

Proof of (1): the eigenvalues of a Hermitian operator are real Let $|v\rangle$ be an eigenvector of $A$ with eigenvalue $\lambda$. Then the eigenvalue equation (3.47) and its dual are

$$
A |v\rangle = \lambda |v\rangle,
$$

$$
|v\rangle A^\dagger = \lambda^* |v\rangle.
$$

As $A$ is Hermitian we have $A^\dagger = A$. Operate on the first of (3.56) with $|v\rangle$ and use the second to operate on $|v\rangle$. Subtracting, the result is

$$
0 = (\lambda - \lambda^*) |v\rangle.
$$

But $|v\rangle |v\rangle = 0$. Therefore $\lambda = \lambda^*$; the eigenvalues are real.

Proof of (2): the eigenvectors of a Hermitian operator are orthogonal Let $|v_1\rangle$ and $|v_2\rangle$ be two eigenvectors with corresponding eigenvalues $\lambda_1$, $\lambda_2$. The eigenvalue equations are

$$
A |v_1\rangle = \lambda_1 |v_1\rangle,
$$

$$
A |v_2\rangle = \lambda_2 |v_2\rangle.
$$

For simplicity, let us first consider the case $\lambda_1 \neq \lambda_2$. Operating on the first of (3.58) with $|v_2\rangle$ and on the second with $|v_1\rangle$ results in

$$
\langle v_2 | A | v_1 \rangle = \lambda_1 \langle v_2 | v_1 \rangle,
$$

$$
\langle v_1 | A | v_2 \rangle = \lambda_2 \langle v_1 | v_2 \rangle.
$$

Taking the complex conjugate of the second of these gives

$$
\langle v_1 | A | v_2 \rangle^* = \lambda_2^* \langle v_1 | v_2 \rangle^*.
$$

so that

$$
0 = (\lambda_1 - \lambda_2^*) \langle v_1 | v_2 \rangle.
$$

Under our assumption that $\lambda_1 \neq \lambda_2$ we must have $\langle v_1 | v_2 \rangle = 0$: the eigenvectors are orthogonal. If all $n$ eigenvalues are distinct, then it is clear that the eigenvectors span $V$ and therefore form a basis.

If $\lambda_1 = \lambda_2$ then we can use the Gram-Schmidt procedure to construct an orthonormal pair from $|v_1\rangle$ and $|v_2\rangle$ the complex spectral theorem (§3.10) guarantees that a Hermitian operator on an $n$-dimensional space always has $n$ orthogonal eigenvectors and so there must exist appropriate linear combinations of $|v_1\rangle$ and $|v_2\rangle$ that are orthogonal.

Exercise: Show that (i) the eigenvalues of a unitary operator are complex numbers of unit modulus and (ii) the eigenvectors of a unitary operator are mutually orthogonal.

How to diagonalise a Hermitian operator Let $A$ be a Hermitian operator with normalised eigenvectors $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$. When the matrix representing $A$ is expressed in terms of its eigenbasis then the matrix elements are

$$
\langle v_l | A | v_k \rangle = \lambda_l \delta_{lk},
$$

where $\lambda_l$ is the eigenvalue corresponding to $|v_l\rangle$. In our standard $|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle$ basis, we have that the matrix elements of $A$ are given by

$$
\langle v_l | A | v_k \rangle = \langle v_l | I A I^\dagger | v_k \rangle
$$

resolving the identity (3.5) through $I = \sum_k |v_k\rangle \langle v_k|$. Written as a matrix equation, this is

$$
A = U \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) U^\dagger,
$$

where $U_{lj} = \langle v_l | v_j \rangle$ so that the $l$th column of $U$ expresses $|v_l\rangle$ in terms of the $|v_1\rangle, \ldots, |v_n\rangle$ basis.
II–13 Mathematical Methods wk 2: Linear Operators

Exercise: Using $UU^\dagger = U^\dagger U = I$ show the following:

$U^\dagger AU = \text{diag}(\lambda_1, \ldots, \lambda_n)$,

$A^m = U \text{diag}(\lambda_1^m, \ldots, \lambda_n^m) U^\dagger$.

$\text{tr} A = \sum_{i=1}^n \lambda_i$,

$\text{det} A = \prod_{i=1}^n \lambda_i$.

(3.65)

Simultaneous diagonalisation of two Hermitian matrices

Let $A$ and $B$ be two Hermitian operators. There exists a basis in which $A$ and $B$ are both diagonal if and only if $[A, B] = 0$. Some comments before proving this:

1. Because the eigenvectors of Hermitian operators are orthogonal, any basis in which such an operator is diagonal must be an eigenbasis.

2. An equivalent statement is that “$A$ and $B$ both have the same eigenvectors if and only if $[A, B] = 0$.”

3. In this eigenbasis the only difference between $A$ and $B$ is the values of their diagonal elements (i.e., their eigenvalues).

Proof: We first show that if there is a basis in which $A$ and $B$ are both diagonal then $[A, B] = 0$. This is obvious: diagonal matrices commute.

The converse is that if $[A, B] = 0$ then there is a basis in which both $A$ and $B$ are diagonal. To prove this, note that because $A$ is Hermitian we can find a basis in which $A = \text{diag}(a_1, \ldots, a_n)$, where the $a_i$ are the eigenvalues of $A$. In this basis $B$ will be represented by some matrix

$B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots \\ B_{n1} & \cdots & \cdots & B_{nn} \end{pmatrix}$.

(3.66)

The commutator

$AB - BA = \begin{pmatrix} 0 & (a_1 - a_2)B_{12} & \cdots & (a_1 - a_n)B_{1n} \\ (a_2 - a_1)B_{21} & 0 & \cdots & (a_2 - a_n)B_{2n} \\ \vdots & \ddots & \ddots & \ddots \\ (a_n - a_1)B_{n1} & \cdots & \cdots & 0 \end{pmatrix}$.

(3.67)

By assumption $[A, B] = 0$. From the matrix (3.67) we see that this means that $B_{ij} = B_{ji} = 0$ for all indices $(i, j)$ for which $a_i \neq a_j$: if all of A's eigenvalues are distinct then $B$ must be diagonal.

If some of the $a_i$ aren’t distinct we have just a little more work to do. Take, for example, the case $a_1 = a_2$.

Then we have that

$B = B' = \begin{pmatrix} B_{11} & B_{12} & 0 & 0 & \cdots \\ B_{21} & B_{22} & 0 & 0 & \cdots \\ 0 & 0 & B_{33} & \cdots & \cdots \\ 0 & 0 & 0 & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & B_{nn} \end{pmatrix}$.

(3.68)

using $B_2 = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ and $B_{n-2} = \text{diag}(B_{33}, \ldots, B_{nn})$ to denote block submatrices of $B$. Now introduce a unitary change-of-basis matrix

$U = \begin{pmatrix} U_{11} & U_{12} & 0 & \cdots \\ U_{21} & U_{22} & 0 & \cdots \\ 0 & 0 & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}$.

(3.69)

Then in this new basis $B$ will be represented by the matrix

$U^\dagger BU = \begin{pmatrix} B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & \ddots & \ddots & \ddots \\ B_{n1} & \cdots & \cdots & B_{nn} \end{pmatrix}$.

(3.70)

Notice that $U$ changes only the upper-left corner of $B$. We can always choose $U_2$ to make $U_2^\dagger B_2 U_2 = \text{diag}(b_1, b_2)$, so that the matrix $U^\dagger BU$ becomes diagonal. The basis change $U$ has no effect on the matrix representing $A$ because $U^\dagger AU = \text{diag}(a_1, a_2, \ldots) = a_1 \text{diag}(1, 1)$ is proportional to the identity, we have that

$U^\dagger AU = \begin{pmatrix} A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \ddots & \ddots & \ddots \\ A_{n1} & \cdots & \cdots & A_{nn} \end{pmatrix}$.

(3.71)

3.12 Odds and ends

Quadratic forms

Expressions such as $ax^2 + 2bxy + cy^2$, or, more generally,

$\sum_{i=1}^n A_{ii}x_i^2 + 2 \sum_{i<j} A_{ij}x_i x_j = \sum_{i=1}^n \sum_{j=1}^n A_{ij}x_i x_j$

are known as (homogeneous) quadratic forms. They can be written in matrix form as $x^T A x$, where $A$ is a symmetric matrix with elements $A_{ij} = A_{ji}$. For example,

$4x^2 + 6xy + 7y^2 = (x y) \begin{pmatrix} 4 & 3 \\ 3 & 7 \end{pmatrix} (x y)$.

(3.73)

Exercise: Explain why under a certain change of basis $x \rightarrow x'$ the quadratic form (3.72) can be expressed as $\lambda_{x'}x'^2 + \cdots + \lambda_n x'^n$, where the $\lambda_i$ are the eigenvalues of $A$. What is the relationship between $x$ and $x'$?

Do you need to make any assumptions about the elements $A_{ij}$ or $x_i$?

Lorentz transformations

Our definition of scalar product includes a condition (2.8) that scalar products are positive unless one of the vectors involved is zero. This condition is relaxed in the special theory of relativity, where the scalar product of two four-vectors $x = (ct, x, y, z)$ and $x' = (ct', x', y', z')$ is defined to be

$x \cdot x' = (ct x' x y' z' y z) = \gamma (ct')^2 - (x')^2 - (y')^2 - (z')^2$.

(3.74)

where the metric $\eta = \text{diag}(1, -1, -1, -1)$. The (square of the) “length” of a spacetime interval $dx = (dt, dx, dy, dz)$ is then

$(dx)^2 = (c dt)^2 - (dx)^2 - (dy)^2 - (dz)^2$.

(3.75)

which can be positive, negative or zero, depending on whether the interval is time-like, space-like or light-like.

A Lorentz transformation is a change of basis $x \rightarrow x'$: $x \rightarrow x'$ that preserves the scalar product (3.74). An example familiar from the first-year course is $x' = Ax$, where

$A = \begin{pmatrix} \gamma & -\beta \gamma \\ -\beta \gamma & \gamma \end{pmatrix}$.

(3.76)

with $\beta = c/e$ and $\gamma = 1/\sqrt{1 - \beta^2}$. It is easy to confirm that $\Lambda^T \eta \Lambda = \eta$. 


Tensor product  Let $V$ and $W$ be $n$- and $m$-dimensional vector spaces respectively. The tensor product $V \otimes W$ is a vector space of dimension $n \times m$. A basis for $V \otimes W$ is $\{|e^V_i\rangle \otimes |e^W_j\rangle\}$, where $\{|e^V_1\rangle, \ldots, |e^V_n\rangle\}$ is a basis of $V$ and $\{|e^W_1\rangle, \ldots, |e^W_m\rangle\}$ is a basis of $W$. For any $|v_i\rangle \in V$, $|w_i\rangle \in W$ and scalar $\alpha$, the tensor product satisfies

\[
(v_1 + v_2) \otimes |w_i\rangle = (v_1 \otimes |w_i\rangle) + (v_2 \otimes |w_i\rangle),
\]

\[
\alpha |v_i\rangle \otimes |w_i\rangle = \alpha (|v_i\rangle \otimes |w_i\rangle).
\]

(3.77)

Given linear operators $A : V \to V$ and $B : W \to W$ we can define a new linear operator $(A \otimes B) : V \otimes W \to V \otimes W$ through

\[
(A \otimes B)(|v_i\rangle \otimes |w_i\rangle) = A|v_i\rangle \otimes B|w_i\rangle.
\]

(3.78)

If $V$ and $W$ are inner-product spaces then there is a natural inner product between two vectors $|v_1\rangle \otimes |v_1\rangle$ and $|v_2\rangle \otimes |v_2\rangle \in V \otimes W$:

\[
(|v_1\rangle \otimes |v_1\rangle) (|v_2\rangle \otimes |v_2\rangle) = (v_1|v_1\rangle)(v_2|v_2\rangle).
\]

(3.79)

Further reading

See RHB§8, DR§II.

Jordan normal form  Not all matrices can be diagonalised, but it turns out that for every matrix $A$ there is an invertible transformation $P$ for which $P^{-1}AP$ takes on the block-diagonal form

\[
P^{-1}AP = \begin{pmatrix}
A_1 & & \\
& \ddots & \\
& & A_n
\end{pmatrix},
\]

(3.80)

where the blocks are

\[
A_i = \begin{pmatrix}
\lambda_i & 1 & & \\
& \lambda_i & 1 & \\
& & \ddots & 1 \\
& & & \lambda_i
\end{pmatrix},
\]

(3.81)

having ones immediately above the main diagonal. Each $\lambda_i$ here is an eigenvalue of $A$ with multiplicity given by the number of diagonal elements in the corresponding $A_i$.

Singular value decomposition  A generalisation of the eigenvector decomposition we have been applying to square matrices holds to any $m \times n$ matrix $A$: any such $A$ can be factorized as

\[
A = UDV^\dagger,
\]

(3.82)

where $U$ is an $m \times m$ unitary matrix, $D$ is an $m \times n$ diagonal matrix consisting of the so-called singular values of $A$, and $V$ is an $n \times n$ unitary matrix.