# Contents

# Books

In developing these lectures, I have not followed a particular text book. I recommend *Quantum Physics* by Stephen Gasiorowicz as a good 'second level' book for quantum mechanics. *Quantum Mechanics* by Nouredine Zettili is also useful, because it contains many worked examples and exercises. However, these books do not cover some of the more advanced material in the course. *Quantum Processes, Systems, and Information*, by Schumacher and Westmoreland, is a good source for this.

# 1 Linear Algebra

## 1.1 Matrices

### 1.1.1 A Few Definitions

A matrix is simply an ordered array of numbers, or *elements*:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \dots & \dots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \tag{1.1}$$

This matrix has $m$ *rows* and $n$ *columns*. We shall normally deal with *square* matrices with $m = n$. Then $n$ is sometimes known as the *order* of the matrix.

To do anything with matrices, we need to define basic arithmetic operations:

1. *Equality*: $A = B$ if $A_{ij} = B_{ij}$ for all $i, j$.
2. *Addition*: $C = A + B$ if $C_{ij} = A_{ij} + B_{ij}$ for all $i, j$. So addition is *commutative*, A+B=B+A, and *associative* (A+B)+C=A+(B+C).
3. *Multiplication by a scalar*: The elements of $\alpha A$ are $\alpha A_{ij}$.
4. *Matrix multiplication*: $C = AB$ if $C_{ij} = \sum_k A_{ik} B_{kj}$. This is only meaningful if the number of columns in $A$ is the same as the number of rows in $B$. Multiplication is associative $(AB)C = A(BC)$ and *distributive* $A(B + C) = AB + AC$, but not. in general, commutative $AB \neq BA$.
5. *Matrix inverse*: The inverse of a square matrix $A$, written $A^{-1}$, is defined by $AA^{-1} = \mathbb{1}$, where $\mathbb{1}$ is the *unit* matrix, $\mathbb{1}_{ij} = \delta_{ij}$, the Kronecker delta. This inverse does not always exist.
6. *Transpose*: The transpose of $A$, written $A^T$, is formed by swapping the rows and columns. Thus $(A^T)_{ij} = A_{ji}$.
7. *Adjoint*: The adjoint, or *Hermitian conjugate*, of $A$, written $A^\dagger$, is the complex conjugate of the transpose. Thus $(A^\dagger)_{ij} = A^*_{ji}$. (This is the physicists use of the term adjoint: there is another definition which is the transpose of the matrix of cofactors)

In these notes, matrices will generally be denoted by upper case Roman letters ($A$), vectors by lower case Roman ($x$) and scalars by Greek letters ($\lambda$). Later, hatted capitals ($\hat{A}$) will be used for operators. As ever, this convention will be abandoned where it conflicts with standard notation in a particular topic.

## 1.2 Determinants

The *determinant* of a matrix, $|A|$, is defined to be the determinant of the array in Eq.(1.1). This is a useful number for characterising a matrix. We shall see later that the determinant is equal to the product of the eigenvectors.

Determinant are conventionally evaluated using an expansion over *minors*:

$$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{vmatrix} = a\begin{vmatrix} f & g & h \\ j & k & l \\ n & o & p \end{vmatrix} - b\begin{vmatrix} e & g & h \\ i & k & l \\ m & o & p \end{vmatrix} + c\begin{vmatrix} e & f & h \\ i & j & l \\ m & n & p \end{vmatrix} - d\begin{vmatrix} e & f & g \\ i & j & k \\ m & n & o \end{vmatrix} . \tag{1.2}$$

Each of these $3 \times 3$ determinants is then evaluated in a similar way, until we reduce to $1 \times 1$ determinants, which are equal to their (only) element. Note that it is not necessary to work along the top row: we can, in fact, chose any row or column, and get the same result.

The determinant of a product is equal to the product of the determinants: $|AB| = |A||B|$ (*product theorem*).

## 1.3 Trace

Another useful way of characterising a matrix by a single number is to find the *trace*, which is the sum of the diagonal elements: $\mathrm{Tr}\{A\} = \sum_i A_{ii}$. We shall see that the trace is equal to the sum of the eigenvalues of the matrix.

The trace of a product is independent of the order of multiplication: $\mathrm{Tr}\{AB\} = \mathrm{Tr}\{BA\}$, even if $AB \neq BA$.

**Proof**

$$\begin{aligned} \mathrm{Tr}\{AB\} &= \sum_i (AB)_{ii} = \sum_{ij} A_{ij}B_{ji} \\ &= \sum_{ij} B_{ji}A_{ij} = \sum_j (BA)_{jj} = \mathrm{Tr}\{BA\} \end{aligned}$$

More generally, any cyclical permutation of a product of operators has the same trace:

$$\mathrm{Tr}\{ABC\} = \mathrm{Tr}\{BCA\} = \mathrm{Tr}\{CAB\} \ . \tag{1.3}$$

## 1.4 Finding the Inverse

The conventional analytic way of finding the inverse of a matrix is to use the formula

$$A^{-1} = \frac{1}{|A|} C^T \tag{1.4}$$

where C is the matrix of the *cofactors* of the elements of $A$. The cofactor is the determinant of the matrix with the row and column containing $A_{ij}$ struck out (the *minor* from above), multiplied by the same sign pattern as in a determinant. For example, if

$$A = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{pmatrix} \text{ then } C_{23} = -\begin{vmatrix} a & b & d \\ i & j & l \\ m & n & p \end{vmatrix} \tag{1.5}$$

$A^{-1}$ does not exist if $|A| = 0$ – the matrix is *singular*.

## 1.5 Classification of Matrices

There are many different types of matrix, with special properties. Some which occur frequently are:

1. Diagonal: Only the diagonal elements are non-zero, $A_{ij} = 0$ for $i \neq j$.
2. Symmetric: The matrix is equal to its transpose, $A^T$.
3. Hermitian: The matrix is equal to its adjoint, $A^\dagger$.
4. Orthogonal: The inverse of the matrix is equal to its transpose: $O^{-1} = O^T$.
5. Unitary: The inverse of the matrix is equal to its adjoint: $U^{-1} = U^\dagger$. Real unitary matrices are also orthogonal.
6. Normal: The matrix commutes with its adjoint, $AA^\dagger = A^\dagger A$.

**Exercise:** Show that $(AB)^\dagger = B^\dagger A^\dagger$

**Exercise:** Show that if $A$ is Hermitian, $\langle\alpha|A|\beta\rangle = \langle\beta|A|\alpha\rangle^*$, for any vectors $|\alpha\rangle$, $|\beta\rangle$.

**Exercise:** Show that the modulus of the determinant of a unitary matrix is always 1.

## 1.6 Vectors

A vector can be defined as a matrix with a single column:

$$|x\rangle = \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{1.6}$$

In these notes, we use *Dirac notation*. A column vector, also known as a *ket* is written $|x\rangle$.

The adjoint of our ket is a row vector, also known as a *bra*, with symbol $\langle x|$.

$$\langle x| = (|x\rangle)^\dagger = (x_1^* \; x_2^* \; \ldots \; x_n^*)\,. \tag{1.7}$$

With this definition, the same arithmetic rules apply as for matrices. There is no well-formed product of two vectors $x$ and $y$, because they do not satisfy the rule that the number of rows in one is the same as the number of columns in the other.

However, there are two well formed products of a ket and a bra. The *inner product* of $|y\rangle$ and $\langle x|$ is

$$\langle x|y\rangle = \sum_i x_i^* y_i\,, \tag{1.8}$$

which is a scalar. It is the equivalent of the dot product $\boldsymbol{x} \cdot \boldsymbol{y}$ for Cartesian vectors. In the language of Dirac notation, this combination of a bra and a ket is known as a *braket*. Note that, reversing the order of the vectors,

$$\langle y|x\rangle = \sum_i y_i^* x_i = \langle x|y\rangle^*\,. \tag{1.9}$$

The other product is the *outer product*. This can be thought of as the $xy^\dagger$ matrix product (summing over the index which only goes to '1'), but it is never written like this. We will use the notation $|x\rangle\langle y|$ (in other applications, you may see $x \otimes y$). It is a matrix with elements

$$(|x\rangle\langle y|)_{ij} = x_i y_j^* \,. \tag{1.10}$$

Note, this does not correspond to the Cartesian vector product!

**Exercise:** Show that

$$(|x\rangle\langle y|)\,|z\rangle = \langle y|z\rangle|x\rangle \tag{1.11}$$

## 1.7 Commutators and Anti-Commutators

The commutator of two matrices $A$ and $B$, is defined as

$$[A,B] = AB - BA \,. \tag{1.12}$$

Note that $[B,A] = -[A,B]$ and $[A,A] = 0$. The commutator is an important quantity in quantum mechanics. The fact that operators representing quantities such as position and momentum do not commute is what makes it different from classical mechanics.

We will also meet the anti-commutator

$$\{A,B\} = AB + BA \,. \tag{1.13}$$

## 1.8 Eigenvalues and Eigenvectors

The product of a vector with a matrix is another vector: $(Ax)_i = \sum_j A_{ij}x_j$, Thus the equation $A|x\rangle = |b\rangle$ represents a set of simultaneous linear equations. If the inverse $A^{-1}$ exists, its formal solution is $|x\rangle = A^{-1}|b\rangle$.

The matrix eigenvalue problem is defined by the equation

$$A|x\rangle = \lambda|x\rangle \,, \tag{1.14}$$

where the eigenvalues, $\lambda$, and eigenvectors, $|x\rangle$, are to be determined. Clearly if $|x\rangle$ is a solution, then so is $\mu|x\rangle$, where $\mu$ is any scalar, so we conventionally *normalise* the eigenvectors by insisting $\langle x|x\rangle = \sum_i |x_i|^2 = 1$

$(A-\lambda)|x\rangle = 0$ implies that $|A - \lambda\mathbb{1}| = 0$, the *secular equation*. One way of finding the eigenvalues is to solve this equation, which, for an $N \times N$ matrix, is an order-$N$ polynomial, with $N$ solutions. These solutions are not always distinct; identical eigenvalues are said to be *degenerate*.

The eigenvalues of an Hermitian (also real symmetric) matrix are all real. The eigenvectors are orthogonal, that is $\langle x_\alpha|x_\beta\rangle = 0$, if the corresponding eigenvalues are non-degenerate ($\lambda_\alpha \neq \lambda_\beta$).

**Proof**
Consider two eigenvectors satisfying the equations

$$\begin{aligned} A|x_\alpha\rangle &= \lambda_\alpha|x_\alpha\rangle \\ A|x_\beta\rangle &= \lambda_\beta|x_\beta\rangle \end{aligned}$$

Multiplying the first of these equations by $\langle x_\beta|$, the second by $\langle x_\alpha|$

$$\begin{aligned} \langle x_\beta|A|x_\alpha\rangle &= \lambda_\alpha\langle x_\beta|x_\alpha\rangle \\ \langle x_\alpha|A|x_\beta\rangle &= \lambda_\beta\langle x_\alpha|x_\beta\rangle \end{aligned}$$

Taking the Hermitian conjugate of the first equation, and using the fact that $A$ is Hermitian:

$$\langle x_\alpha|A|x_\beta\rangle = \lambda_\alpha^*\langle x_\alpha|x_\beta\rangle$$

Hence $\lambda_\beta\langle x_\alpha|x_\beta\rangle = \lambda_\alpha^*\langle x_\alpha|x_\beta\rangle$. If $\alpha = \beta$, this implies $\lambda_\alpha = \lambda_\alpha^*$, so $\lambda_\alpha$ must be real.
Thus $\lambda_\beta\langle x_\alpha|x_\beta\rangle = \lambda_\alpha\langle x_\alpha|x_\beta\rangle$, which means that $\langle x_\alpha|x_\beta\rangle = 0$ if $\lambda_\alpha \neq \lambda_\beta$.

The situation is more complicated for degenerate eigenvalues, when this proof does not hold. Suppose we have a set of $k$ degenerate eigenvalues. Then, any linear combination of the corresponding eigenvectors is also an eigenvector. For normal matrices, it is always possible to find a set of $k$ such linear combinations which are mutually orthogonal. As a simple example, for a degenerate pair of eigenvectors $|x_1\rangle$, $|x_2\rangle$, it is easy to show that the combinations $|z_1\rangle = |x_1\rangle$, $|z_2\rangle = |x_2\rangle - \langle z_1|x_2\rangle/\langle z_1|z_1\rangle|z_1\rangle$ are orthogonal (but not normalised). Numerical eigenvector solvers usually generate such orthogonalised combinations.

**Exercise:** Show that for three degenerate states, the combinations

$$\begin{aligned} |z_1\rangle &= |x_1\rangle \\ |z_2\rangle &= |x_2\rangle - \frac{\langle z_1|x_2\rangle}{\langle z_1|z_1\rangle}|z_1\rangle \\ |z_3\rangle &= |x_3\rangle - \frac{\langle z_1|x_3\rangle}{\langle z_1|z_1\rangle}|z_1\rangle - \frac{\langle z_2|x_3\rangle}{\langle z_2|z_2\rangle}|z_2\rangle \end{aligned}$$

are orthogonal. This is the Gram-Schmidt orthogonalisation method.

In quantum mechanics, we will frequently find that the matrices we are dealing with have a *block diagonal* from. This means that the only non-zero matrix elements are arranged in square blocks along the diagonal. For example, the matrix

$$\begin{pmatrix} A_{11} & A_{12} & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 \\ 0 & 0 & A_{33} & A_{34} \\ 0 & 0 & A_{43} & A_{44} \end{pmatrix} \tag{1.15}$$

consist of a two $2 \times 2$ block. Solving the eigenvalue problem for a block diagonal matrix is particularly simple, because, when written out as simultaneous equations, we find that each block behaves independently. For the above matrix, instead of a $4 \times 4$ eigenvalue problem, we can solve separately the two $2 \times 2$ problems.

## 1.9 Similarity Transformations

A *similarity transformation* is a transformation of a matrix according to

$$A' = S^{-1}AS \tag{1.16}$$

where $S$ is any matrix whose inverse exists. An important class of similarity transformations are those for which $S$ is unitary, so $A' = U^\dagger A U$. These are known as unitary transformations.

**Exercise:** Show that (i) the determinant, (ii) the trace, and (iii) the eigenvalues of a matrix are unchanged by a similarity transformation. When the transformation is unitary, show additionally that (iv) if $A$ is Hermitian, so is $A'$.

A matrix is diagonalised by finding a similarity transformation which produces a diagonal matrix. The required matrix is formed from columns which are the normalised eigenvectors of the original matrix, that is $S_{ij} = (|x^j\rangle)_i = x_i^j$. Here the superscript $j = 1 \ldots n$ labels the $n$ eigenvectors, $|x^j\rangle$, of $A$.

**Proof**

The $(i,j)$ element of the transformed matrix is

$$\begin{aligned}(S^{-1}AS)_{ij} &= \sum_{kl}(S^{-1})_{ik}A_{kl}S_{lj} = \sum_{kl}(S^{-1})_{ik}A_{kl}x_l^j \\ &= \sum_{k}(S^{-1})_{ik}\lambda_j x_k^j = \lambda_j \sum_k (S^{-1})_{ik}S_{kj} = \lambda_j \delta_{ij}.\end{aligned}$$

Hence the transformed matrix is diagonal, with elements equal to the eigenvalues.

The most general type of matrix which has orthogonal eigenvectors is the normal matrix. $S$ will then be a unitary matrix – only normal matrices can be diagonalised by a unitary transformation. A non-normal matrix may be diagonisable by a non-unitary similarity transformation. However, it is also possible that it does not have $n$ linearly independent eigenvectors, in which case it is *defective*. A defective matrix cannot be diagonalised by a similarity transformation.

**Exercise:** Show that, if the eigenvectors are orthonormal, $S$ is indeed unitary.

Since the determinant and trace of a matrix are unchanged by a similarity transformation, it follows immediately that if we can diagonalise a matrix its determinant is equal to the product of its eigenvalues, and its trace is equal to their sum. In fact this is true for any matrix, even one which is not diagonalisable.

## 1.10 Kronecker Delta and Diagonal Matrices

The Kronecker delta, written $\delta_{ij}$ has two indicies which take integer values:

$$\begin{aligned}\delta_{ij} &= 1 \quad (i = j) \\ &= 0 \quad (i \neq j)\end{aligned} \tag{1.17}$$

It can be thought of as a matrix with '1's along the diagonal '0's elsewhere. This is the unit matrix

$$\mathbb{1} = \begin{pmatrix} 1 & 0 & 0 & \ldots \\ 0 & 1 & 0 & \ldots \\ 0 & 0 & 1 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{1.18}$$

Sums involving the Kronecker delta are easily evaluated:

$$\sum_j \delta_{ij} x_j = x_i \,. \tag{1.19}$$

This is equivalent to the identity $\mathbb{1}|x\rangle = |x\rangle$.

It is convenient to write diagonal matrices in terms of the Kronecker delta. The matrix

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots \\ 0 & \lambda_2 & 0 & \dots \\ 0 & 0 & \lambda_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{1.20}$$

has elements $D_{ij} = \lambda_i \delta_{ij}$.

## 1.11 Singular Matrices

A matrix with zero determinant is said to be *singular*. The determinant of the inverse of $A$, $|A^{-1}|$ is $|A|^{-1}$, so the inverse of a singular matrix does not exist. If we have a set of $N$ simultaneous equations with $N$ unknowns, $Ax = b$, and $A$ is singular, there is no solution.

If the matrix can be diagonalised by a similarity transformation, its determinant is the product of the eigenvalues. Hence a singular matrix must have at least one zero eigenvalue. The number of non-zero eigenvalues of a matrix is known as its *rank*.

In fact, these are not the best definitions of rank and singularity; they only apply for square matrices, while the concepts can be applied to rectangular matrices. However, we will not consider the more general definitions here.

## 1.12 Pauli Matrices

The Pauli matrices, $\sigma_x$, $\sigma_y$, $\sigma_z$ (or sometimes $X$, $Y$, $Z$) play an important role in quantum mechanices, particularly for treatments two level systems such as the spin-$\frac{1}{2}$. Here, I will introduce them and mention a few of their properties.

The matrices are

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.21}$$

The Pauli matrices are Hermitian and unitary. They all have zero trace and eigenvalues $\pm 1$. The square of each is the unit matrix:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbb{1}. \tag{1.22}$$

The commutators of the Pauli matrices are

$$[\sigma_x, \sigma_y] = 2i\sigma_z \tag{1.23}$$

along with cyclic permutations ($[\sigma_z, \sigma_x] = 2i\sigma_y$ and $[\sigma_y, \sigma_z] = 2i\sigma_x$). This is sometimes written in terms of the *Levi-Civita symbol* as

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k. \tag{1.24}$$

The anti-commutators are

$$\{\sigma_i, \sigma_j\} \equiv \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}\mathbb{1}. \tag{1.25}$$

## 1.13 Problems

1.1. Given the matrices

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

(a) Calculate $AB$ and $AC$.

(b) Find the eigenvalues and eigenvectors of the matrix $A$. Show that the eigenvectors are orthogonal.

(c) Find the inverse of the matrix $B$.

1.2. Given the matrices

$$A = \begin{pmatrix} 2 & 1 & 2 \\ 3 & 5 & 7 \\ 1 & 0 & 1 \end{pmatrix}, \; B = \begin{pmatrix} -3 & 1 & 0 \\ 6 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix}$$

Evaluate the products $AB$ and $BA$. Show that the transpose $(AB)^T = B^T A^T$. Calculate the determinant $|A|$ and trace, $\mathrm{Tr}\{A\}$.

Given

$$C = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \end{pmatrix}, \; D = \begin{pmatrix} 5 & -2 \\ 3 & 1 \\ -1 & 0 \end{pmatrix}$$

Evaluate the products $CD$ and $DC$.

1.3. Consider the following matrix $A$ and vectors $x$ and $y$:

$$A = \begin{pmatrix} 5 & 3+2i & 3i \\ -i & 3i & 8 \\ 1-i & 1 & 4 \end{pmatrix}, \qquad x = \begin{pmatrix} -3i \\ 2+i \\ 4 \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} 2 \\ -i \\ 2-3i \end{pmatrix}$$

(a) Write down the complex conjugate, the transpose and the Hermitian conjugates of $A$ and $x$.

(b) Evaluate the inner products $\langle x|y\rangle$ and $\langle y|x\rangle$.

(c) Evaluate $Ax$, $Ay$.

(d) Calculate $\langle y|Ax\rangle$, $\langle Ay|x\rangle$ and $\langle A^\dagger y|x\rangle$.

1.4. Given the matrices

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

Find the eigenvalues and eigenvectors of both matrices. What is the matrix $U$ which diagonalises $A$? Verify this by direct multiplication.

1.5. Find the eigenvalues and eigenvectors of

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

(Note: this is not a *normal* matrix.)

1.6. Find the eigenvalues and an orthonormal set of eigenvectors for the matrix

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

Note: two of the eigenvectors are degenerate in this problem; take care that the corresponding eigenvectors you find are orthogonal.

1.7. Find the eigenvalues and normalised eigenvectors of the matrix

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}.$$

Without further calculation, write down the eigenvalues and eigenvectors of

$$B = \begin{pmatrix} 1 & 3 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 5 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

1.8. We saw that the eigenvalues and eigenvectors of

$$A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$$

are $\lambda_1 = 4$, $|x_1\rangle = (1/\sqrt{2})(1,1)^T$ and $\lambda_2 = 6$, $|x_2\rangle = (1/\sqrt{2})(1,-1)^T$.

By calculating the outer products $|x_1\rangle\langle x_1|$ and $|x_2\rangle\langle x_2|$, show that

$$\sum_{\alpha=1}^{2} |x_\alpha\rangle\langle x_\alpha| = \mathbb{1} \text{ and } \sum_{\alpha=1}^{2} \lambda_\alpha |x_\alpha\rangle\langle x_\alpha| = A$$

1.9. Verify some of the properties of the Pauli matrices, particularly Eq.(1.23) and Eq.(1.25). You do not have to do every combination, just enough to practice doing the calculations.

1.10. Find the eigenvalues and normalised eigenvectors of each of the Pauli matrices.

1.11. Use the anti-commutator relationship, Eq.(1.25), to show that the Pauli matrices satisfy the trace identity

$$\mathrm{Tr}\{\sigma_i \sigma_j\} = 2\delta_{ij}\,.$$

# 2 Linear Vector Spaces

So far, we have been treating vectors as just a list of numbers. However, we are very used to the idea that a three component vector represents the Cartesian coordinates of a point in three dimensional space, $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$. We call this $R_3$ ($R$ for real, as the components of the vector must be real). Similarly, the points on a plane $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}$ make up the two dimensional space $R_2$.

This idea is generalised to produce the concept of a linear vector space. An $n$ dimensional space is defined by $n$ linearly independent basis vectors. The space is made up of all possible linear combinations of these basis vectors. If the coefficients are real, this gives $R_n$, while if they are complex, it is $C_n$. For example, the four-vectors of special relativity exist in an $R_4$ space-time.

Formally, a set of vectors $|a\rangle, |b\rangle, |c\rangle, \ldots$ form a linear vector space if:

1. Addition is commutative and associative, so

$$|a\rangle + |b\rangle = |b\rangle + |a\rangle \tag{2.1}$$
$$(|a\rangle + |b\rangle) + |c\rangle = |a\rangle + (|b\rangle + |c\rangle) \tag{2.2}$$

2. Multiplication by a scalar is both distributive and associative, so

$$\lambda(|a\rangle + |b\rangle) = \lambda|a\rangle + \lambda|b\rangle \tag{2.3}$$
$$(\lambda + \mu)|a\rangle = \lambda|a\rangle + \mu|a\rangle \tag{2.4}$$
$$\lambda(\mu|a\rangle) = (\lambda\mu)|a\rangle \tag{2.5}$$

   where $\lambda$ and $\mu$ are any scalars (may be complex in $C_n$).

3. There exists a *null vector*, $|0\rangle$, such that $|a\rangle + |0\rangle = |a\rangle$ for any vector $|a\rangle$.

4. Multiplying by the unit scalar leaves any vector unchanged: $1|a\rangle = |a\rangle$.

5. Multiplying by the zero scalar gives the null vector: $0|a\rangle = |0\rangle$.

6. All vectors have a corresponding negative vector $-|a\rangle$, such that $|a\rangle + (-|a\rangle) = 0$.

In these definitions, the '+' symbol means some rule for combining vectors, such that the combination is also a vector in the space.

Many other mathematical objects form vector spaces, as they obey the above rules. Examples include the spaces of functions, and the spaces of matrices with a given number of rows and columns.

## 2.1 Basis Vectors

A basis is a set of vectors which spans the space, so in the basis $(|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle)$, any vector $|x\rangle$ can be written

$$|x\rangle = \sum_i x_i |e_i\rangle = x_1|e_1\rangle + x_2|e_2\rangle + \ldots + x_n|e_n\rangle \,. \tag{2.6}$$

We now have a distinction between a physical vector $|x\rangle$, independent of a basis, and its representation in a particular basis, which is just a list of numbers, $x_i$.

The corresponding bra is obtained by taking the adjoint of both sides

$$\langle x| = \sum_i x_i^* \langle e_i| = x_1^*\langle e_1| + x_2^*\langle e_2| + \ldots + x_n^*\langle e_n| \,. \tag{2.7}$$

Do not forget to take the complex conjugate of the coefficients – this is a common student mistake!

## 2.2 The Inner Product

An important concept in a vector space is the inner product of two vectors $\langle a|b\rangle$. The inner product is something which has to be defined for a given vector space; for example, in Cartesian $R_3$ it is $\boldsymbol{a}.\boldsymbol{b} = ab\cos\theta$, where $\theta$ is the angle between the vectors. The inner product of the vector with itself is the square of the norm of the vector; $\langle a|a\rangle = ||a||^2$. In $R_3$ this is the length of the vector, and it is useful to generalise this idea to other spaces.

To be described as an inner product $\langle a|b\rangle$ must have the following properties:

$$\langle a|b\rangle = \langle b|a\rangle^* \text{ and } \langle a|(\lambda|b\rangle + \mu|c\rangle) = \lambda\langle a|b\rangle + \mu\langle a|c\rangle. \tag{2.8}$$

It follows that, if $|d\rangle = \lambda|a\rangle + \mu|b\rangle$

$$\langle d|c\rangle = (\lambda^*\langle a| + \mu^*\langle b|)|c\rangle = \lambda^*\langle a|c\rangle + \mu^*\langle b|c\rangle. \tag{2.9}$$

Using these results, with $|a\rangle$ and $|b\rangle$ expanded in terms of the basis states $|e_i\rangle$,

$$|a\rangle = \sum_i a_i|e_i\rangle \qquad \text{and} \qquad |b\rangle = \sum_i b_i|e_i\rangle\,, \tag{2.10}$$

we get

$$\langle a|b\rangle = \sum_i\sum_j a_i^* b_j\langle e_i|e_j\rangle\,. \tag{2.11}$$

We nearly always want to work in a basis which is *orthonormal*, $\langle e_i|e_j\rangle = \delta_{ij}$: the basis vectors are mutually orthogonal and of unit length. Then,

$$\langle a|b\rangle = \sum_i a_i^* b_i \tag{2.12}$$

Defining an orthonormal basis in a space completely defines the inner product. Note that if we think of the $a_i$ and $b_i$ as forming the components of column vectors, this inner product is the same as the one we defined in the previous section. The representation of a vector in a particular basis is a column vector.

Given an orthonormal basis $|e_i\rangle$, it is easy to evaluate the coefficients in the expansion Eq.(2.6). If

$$|x\rangle = \sum_j x_j|e_j\rangle \tag{2.13}$$

then

$$\langle e_i|x\rangle = \sum_j x_j\langle e_i|e_j\rangle = \sum_j x_j\delta_{ij} = x_i\,. \tag{2.14}$$

Thus

$$x_i = \langle e_i|x\rangle\,. \tag{2.15}$$

## 2.3 Linear Operators in Vector Spaces

An *operator* transforms every vector in a space into another vector

$$|y\rangle = \hat{A}|x\rangle. \tag{2.16}$$

$\hat{A}$ is a *linear* operator if

$$\hat{A}(\lambda|a\rangle + \mu|b\rangle) = \lambda\hat{A}|a\rangle + \mu\hat{A}|b\rangle \tag{2.17}$$

where $\lambda, \mu$ are scalars. This definition is independent of any basis. We now introduce a basis $|e_i\rangle$. $\hat{A}$ acting on a particular basis vector produces another vector within the space, which can be written as a linear combination of basis vectors:

$$\hat{A}|e_j\rangle = \sum_i A_{ij}|e_i\rangle \tag{2.18}$$

Hence in this particular basis, using the linearity property, Eq.2.17, we get

$$|y\rangle = \sum_i y_i|e_i\rangle = \hat{A}\sum_j x_j|e_j\rangle = \sum_j x_j \sum_i A_{ij}|e_i\rangle = \sum_i \left(\sum_j A_{ij}x_j\right)|e_i\rangle \,, \tag{2.19}$$

so equating the coefficients of $|e_i\rangle$

$$y_i = \sum_j A_{ij}x_j \,. \tag{2.20}$$

This looks like a matrix-vector multiplication: the representation of linear operator in a particular basis is a matrix.

If the basis $|e_i\rangle$ is orthonormal, taking the inner product of Eq.(2.18) with $|e_i\rangle$ gives

$$A_{ij} = \langle e_i|\hat{A}|e_j\rangle \,. \tag{2.21}$$

## 2.4 The Outer Product

We define the outer product of two vectors $|a\rangle\langle b|$ by using Eq.(1.11):

$$(|a\rangle\langle b|)|c\rangle = \langle b|c\rangle\,|a\rangle \tag{2.22}$$

From this, it follows that the outer product is an operator. It is straight forward to show that

$$\begin{aligned}(\lambda|a\rangle + \mu|b\rangle)\langle c| &= \lambda|a\rangle\langle c| + \mu|b\rangle\langle c| \\ |a\rangle(\lambda^*\langle b| + \mu^*\langle c|) &= \lambda^*|a\rangle\langle b| + \mu^*|a\rangle\langle c| \,.\end{aligned} \tag{2.23}$$

So writing $|a\rangle$ and $|b\rangle$ in terms of basis vectors $|e_i\rangle$:

$$|a\rangle\langle b| = \sum_{kl} a_k b_l^*\, |e_k\rangle\langle e_l| \,. \tag{2.24}$$

For an orthonormal basis, using Eq(2.21), we get

$$(|a\rangle\langle b|)_{ij} = \sum_{kl} a_k b_l^*\, \langle e_i|(|e_k\rangle\langle e_l|)|e_j\rangle = \sum_{kl} a_k b_l^*\, \langle e_l|e_j\rangle\langle e_i|e_k\rangle = a_i b_j^* \,. \tag{2.25}$$

Hence the matrix representation of the outer product operator is just the outer product matrix we defined before.

Using the expansion Eq.(2.15) and our definition of the outer product,

$$|x\rangle = \sum_i \langle e_i|x\rangle|e_i\rangle = \sum_i (|e_i\rangle\langle e_i|)\,|x\rangle \,. \tag{2.26}$$

From this, it follows that

$$\sum_i |e_i\rangle\langle e_i| = \mathbb{1} \,. \tag{2.27}$$

This is known as the *completeness relation* for the basis.

We will often write matrices using outer product notation. This is particularly useful when they are large but sparse, with few non-zero matrix elements. As a simple example

$$\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} = a\,|e_1\rangle\langle e_2| = a\,|1\rangle\langle 2|\,. \tag{2.28}$$

In the second form, we assume that if the ket is just labelled by a number it is a basis vector. It is important that you are comfortable moving between these notations.

If we know the eigenvalues, $\lambda_\alpha$ and eigenvectors $|x_\alpha\rangle$ of a matrix, $A$, we can express it in terms of the sum of outer products of eigenvectors:

$$A = \sum_\alpha \lambda_\alpha\,|x_\alpha\rangle\langle x_\alpha|\,. \tag{2.29}$$

This is known as the *spectral decomposition* of the matrix.

**Proof**

We act with the sum on an arbitrary $|y\rangle$,

$$\sum_\alpha \lambda_\alpha(|x_\alpha\rangle\langle x_\alpha|)|y\rangle = \sum_\alpha \lambda_\alpha\langle x_\alpha|y\rangle|x_\alpha\rangle.$$

If we expand $y$ in a basis of the eigenstates $x_\alpha$

$$|y\rangle = \sum_\alpha \langle x_\alpha|y\rangle|x_\alpha\rangle$$

So

$$A|y\rangle = \sum_\alpha \langle x_\alpha|y\rangle A|x_\alpha\rangle = \sum_\alpha \langle x_\alpha|y\rangle\lambda_\alpha|x_\alpha\rangle\,.$$

## 2.5 Changes of Basis

The vector $|x\rangle$ is expanded in the basis $|e_i\rangle$ as

$$|x\rangle = \sum_i x_i|e_i\rangle. \tag{2.30}$$

We now introduce a new basis $|e'_j\rangle$ defined by

$$|e'_j\rangle = \sum_i S_{ij}|e_i\rangle \tag{2.31}$$

If the original basis is orthonormal, this means $S_{ij} = \langle e_i|e'_j\rangle$. Writing the same vector in both bases

$$|x\rangle = \sum_i x_i|e_i\rangle = \sum_j x'_j|e'_j\rangle = \sum_{ij} x'_j S_{ij}|e_i\rangle \tag{2.32}$$

So we can identify how $x$ transforms:

$$x_i = \sum_j S_{ij}x'_j \quad \text{or} \quad |x\rangle = S|x\rangle'\,. \tag{2.33}$$

The notation is a bit limiting here, because we want to talk about the same physical vector $|x\rangle$ in two different bases. For this reason, I have used $|x\rangle'$ rather than $|x'\rangle$.

We now consider how to transform a matrix $A$, by requiring that the equation $|y\rangle = \hat{A}|x\rangle$ is satisfied in both bases. In the two bases, this corresponds to the matrix equations $|y\rangle = A|x\rangle$ transforming into $|y\rangle' = A'|x\rangle'$. Rewriting the first of these equations using (2.33) to transform the vectors, $S|y\rangle' = AS|x\rangle'$. Hence we can identify $A' = S^{-1}AS$. Thus a change of basis corresponds to a similarity transformation of a matrix.

The inner product of the basis vectors in the transformed space is

$$\langle e'_k|e'_l\rangle = \sum_{ij} S^*_{ik} S_{jl} \langle e_i|e_j\rangle \,. \tag{2.34}$$

If the original basis was orthonormal, $\langle e_i|e_j\rangle = \delta_{ij}$, so

$$\langle e'_k|e'_l\rangle = \sum_{ij} S^*_{ik} S_{jl} \delta_{ij} = \sum_i S^*_{ik} S_{il} = \sum_i S^\dagger_{ki} S_{il} = \delta_{kl} \,, \tag{2.35}$$

where the final step assumes that the transformation is unitary. Thus a unitary transformation transforms an orthonormal basis into another orthonormal basis.

**Exercise:** Show that a unitary transformation preserves the length of a vector: $\langle x|x\rangle = \langle x'|x'\rangle$.

## 2.6 Quantum Mechanics and Hilbert Space

We can consider the states of a quantum mechanical system to be vectors in a vector space which is known formally as a *Hilbert space*. In fact, technically speaking the states correspond to *rays* rather than vectors, because the normalisation of a state contains an arbitrary phase. However, it is seldom necessary to make this distinction in our treatment.

A Hilbert space is defined as a *complete* vector space with an inner product. A formal discussion of completeness is beyond the scope of this course, but in essence it is the requirement that any state of the system can be described exactly by a vector in the space. It implies that there exists at least one orthonormal basis. The Hilbert space of quantum mechanics is a complex space, because when we combine states to form *superpositions* the coefficients can be complex numbers.

We can use many different bases to span this Hilbert space. In fact, if we take all the eigenstates of any Hermitian operator which can act on a quantum system, they are guaranteed to form a complete basis for the Hilbert space of the system. We will often use the eigenstates of a Hamiltonian to define a basis, but the operators associated with any measurement are also Hermitian, so their eigenstates provide alternative bases, which are useful when we discuss measurements.

If we write the states as functions, $\psi_\alpha(x)$, the inner product is defined by

$$\langle \alpha|\beta\rangle = \int_{\substack{\text{all}\\ \text{space}}} dx\, \psi^*_\alpha(x) \psi_\beta(x) \,. \tag{2.36}$$

Here I have, for clarity, used a one-dimensional notation with a single variable $x$. The expression generalises to multiple spatial dimensions in the obvious way. $\psi_\alpha(x)$ is the representation of the state $|\alpha\rangle$ in the basis of the eigenstates of position. We will often write $|\alpha\rangle = \psi_\alpha(x)$, but some people prefer to emphasise this by using $\langle x|\alpha\rangle = \psi_\alpha(x)$.

Our Hilbert space will often have infinite dimension, because for many systems there are an infinite number of eigenstates. But there are also important systems for which the number of dimensions is finite. For example, if we only consider the spin of a single particle with spin-$\frac{1}{2}$, the Hilbert space will be two dimensional. As two state systems, such as a spins, are the qubits of quantum information, they are very important in quantum information theory.

As a final comment, I have assumed in all our discussion of vector spaces that the basis states are discrete; they can be labelled with natural numbers $|e_1\rangle, |e_2\rangle \ldots$. This is not always the case. For example, the eigenstates of the position operator form a continuum ($x$ can take any real value). The same vector space structure applies in this case with modifications; for example, the equations with matrices become integral equations. We will not consider such bases in this course.

## 2.7 A Vector Space of Matrices

There are many different mathematical objects which can form the 'vectors' in a vector space. In this section, we will look at a vector space of matrices, which is something we will use later.

More specifically, consider the vector space which consists of all $2 \times 2$ Hermitian matrices. The most general such matrix is

$$H = \begin{pmatrix} a & c+id \\ c-id & b \end{pmatrix}. \tag{2.37}$$

This is a real vector space, because a sum like $\mu A + \nu B$ is Hermitian if only if the coefficients $\mu$ and $\nu$ are real. It has 4 dimensions - we need the 4 real numbers $a, b, c$ and $d$ to define an arbitrary $2 \times 2$ Hermitian matrix.

The conventional choice of inner product for the space is

$$\langle A|B\rangle = \frac{1}{2}\mathrm{Tr}\{AB\}\,. \tag{2.38}$$

A basis spanning the space consists of the unit matrix and the three *Pauli matrices*, that is

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.39}$$

It is straight forward to show that this is an orthonormal basis. For example

$$\sigma_x\sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \tag{2.40}$$

so $\langle\sigma_x|\sigma_y\rangle = \frac{1}{2}\mathrm{Tr}\{\sigma_x\sigma_y\} = 0$.

This means that we can expand any $2\times2$ Hermitian matrix as a linear combination of Pauli matrices and the unit matrix. This can be useful, as we can sometimes do a calculation for the Pauli matrices and deduce a more general result. The expansion coefficients can be found using Eq.(2.15). For example, with our general matrix Eq.(2.37), the coefficient of the unit matrix is

$$\langle\mathbb{1}|H\rangle = \frac{1}{2}\mathrm{Tr}\left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & c+id \\ c-id & b \end{pmatrix}\right\} = \frac{1}{2}\mathrm{Tr}\left\{\begin{pmatrix} a & c+id \\ c-id & b \end{pmatrix}\right\} = \frac{1}{2}(a+b)\,. \tag{2.41}$$

We find that

$$H = \begin{pmatrix} a & c+id \\ c-id & b \end{pmatrix} = \frac{1}{2}(a+b)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - d\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \frac{1}{2}(a-b)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.42}$$

## 2.8 Combining Vector Spaces

### 2.8.1 Tensor Product

In quantum mechanics, we sometimes want to combine two vector spaces to make a larger space. An exampe which we will look at later occurs if we have a system of two spin-$\frac{1}{2}$ particles. Each single

spin has a two dimensional Hilbert space, so when we combine them, we will get a four dimensional space – we need two states to describe each spin, so four states for the whole system.

We call such a combination of subspaces a *tensor-product*. If we label our two subspaces $\mathscr{U}$ and $\mathscr{V}$, we write the combined space as $\mathscr{W} = \mathscr{U} \otimes \mathscr{V}$. The dimension of $\mathscr{W}$ is given by

$$\dim \mathscr{W} = \dim \mathscr{U} \times \dim \mathscr{V} \,. \tag{2.43}$$

If $\mathscr{U}$ is in state $|\psi\rangle$ and $\mathscr{V}$ in state $|\phi\rangle$, then we write the state of the combined system $\mathscr{W}$ as

$$|\psi\rangle \otimes |\phi\rangle \,. \tag{2.44}$$

However, this is not the most general state in $\mathscr{W}$. Our state is a vector in a vector space, so we can make linear combinations like

$$\mu|\psi_1\rangle \otimes |\phi_1\rangle + \nu|\psi_2\rangle \otimes |\phi_2\rangle \,. \tag{2.45}$$

We will see later that Eq.(2.44) represents a *product state*, while the one in Eq.(2.45) is an *entangled state*.

If we have a basis $|u_i\rangle$ for the space $\mathscr{U}$ and $|v_i\rangle$ for $\mathscr{V}$, then we can write $|\psi\rangle = \sum_i a_i|u_i\rangle$ and $|\phi\rangle = \sum_j b_j|v_j\rangle$. The tensor product of these vectors can be written as

$$|\psi\rangle \otimes |\phi\rangle = \sum_{ij} a_i b_j |u_i\rangle \otimes |v_j\rangle \equiv \sum_{ij} a_i b_j |u_i\rangle|v_j\rangle = \sum_{ij} a_i b_j |u_i, b_j\rangle \,, \tag{2.46}$$

where we introduced common abbreviations for the tensor product notation. We see that the tensor products of the basis states $|u_i\rangle \otimes |v_j\rangle \equiv |u_i\rangle|v_j\rangle \equiv |u_i, v_j\rangle$ form a basis in $\mathscr{W}$. As an example, considering combining a pair of vectors in two-dimensional subspaces:

$$(a_1|u_1\rangle + a_2|u_2\rangle) \otimes (b_1|v_1\rangle + b_2|v_2\rangle) = a_1b_1|u_1, v_1\rangle + a_1b_2|u_1, v_2\rangle + a_2b_1|u_2, v_1\rangle + a_2b_2|u_2, v_2\rangle \,, \tag{2.47}$$

or, in column vector notation

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ a_2 b_1 \\ a_2 b_2 \end{pmatrix} . \tag{2.48}$$

Note this corresponds to the conventional ordering of the basis vectors in $\mathscr{W}$:

$$|u_1, v_1\rangle, |u_1, v_2\rangle, |u_2, v_1\rangle, |u_2, v_2\rangle \,. \tag{2.49}$$

The inner product of two vectors that are tensor products is

$$(\langle\psi_1| \otimes \langle\phi_1|)(|\psi_2\rangle \otimes |\phi_2\rangle) = \langle\psi_1|\psi_2\rangle\langle\phi_1|\phi_2\rangle \,. \tag{2.50}$$

Operators also obey the tensor product structure, with

$$(\hat{A} \otimes \hat{B})(|\psi\rangle \otimes |\phi\rangle) = (\hat{A}|\psi\rangle) \otimes (\hat{B}|\phi\rangle) \,, \tag{2.51}$$

and

$$(\hat{A} \otimes \hat{B})(\hat{C} \otimes \hat{D})(|\psi\rangle \otimes |\phi\rangle) = (\hat{A}\hat{C}|\psi\rangle) \otimes (\hat{B}\hat{D}|\phi\rangle) \,. \tag{2.52}$$

These tensor products of operators are not the most general operators in $\mathscr{W}$. They correspond to doing something in subspace $\mathscr{U}$ and something independent in subspace $\mathscr{V}$. A more general operator will not have this independence. However, such operators can always be decomposed as $C = \sum_j A_j \otimes B_j$. Whether this is a useful decomposition to make will probably depend on th enumber of terms in this sum.

General rules for tensor products of operators are

1. $\hat{A}\otimes 0 = 0$ and $0\otimes\hat{B} = 0$,
2. $\mathbb{1}_{\mathscr{U}}\otimes\mathbb{1}_{\mathscr{V}} = \mathbb{1}_{\mathscr{W}}$,
3. $(\hat{A}_1+\hat{A}_2)\otimes\hat{B} = \hat{A}_1\otimes\hat{B}+\hat{A}_2\otimes\hat{B}$,
4. $a\hat{A}\otimes b\hat{B} = (ab)\hat{A}\otimes\hat{B}$,
5. $\mathrm{Tr}\{\hat{A}\otimes\hat{B}\} = \mathrm{Tr}\{\hat{A}\}\mathrm{Tr}\{\hat{B}\}$,
6. $(\hat{A}\otimes\hat{B})^{-1} = \hat{A}^{-1}\otimes\hat{B}^{-1}$,
7. $(\hat{A}\otimes\hat{B})^{\dagger} = \hat{A}^{\dagger}\otimes\hat{B}^{\dagger}$.

Note that the last rule preserves the order of the operators. In other words, operators always act on their own space. Often, it is understood implicitly which operator acts on which subspace, and we will write $\hat{A}\otimes\mathbb{1} = \hat{A}$ and $\mathbb{1}\otimes\hat{B} = \hat{B}$. Alternatively, we can add subscripts to the operator, e.g., $\hat{A}_{\mathscr{U}}$ and $\hat{B}_{\mathscr{V}}$.

As a practical example, consider two two-dimensional operators

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \tag{2.53}$$

expressed in our orthonormal bases $\{|u_1\rangle, |u_2\rangle\}$ and $\{|v_1\rangle, |v_2\rangle\}$ for $A$ and $B$, respectively. The question is now: what is the matrix representation of $A\otimes B$ in the four dimensional tensor-product space? We will work in our tensor product basis Eq.(2.49).

We can construct the matrix representation of $A\otimes B$, using Eq.(2.18), by applying the operator to the basis vectors. For the two subspaces, we have

$$\begin{aligned} A|u_1\rangle &= A_{11}|u_1\rangle + A_{21}|u_2\rangle \qquad B|v_1\rangle = B_{11}|v_1\rangle + B_{21}|v_2\rangle\,, \\ A|u_2\rangle &= A_{12}|u_1\rangle + A_{22}|u_2\rangle \qquad B|v_2\rangle = B_{12}|v_1\rangle + B_{22}|v_2\rangle\,, \end{aligned} \tag{2.54}$$

which leads to

$$\begin{aligned} A\otimes B|u_1,v_1\rangle &= (A_{11}|u_1\rangle + A_{21}|u_2\rangle)\otimes(B_{11}|v_1\rangle + B_{21}|v_2\rangle) \\ A\otimes B|u_1,v_2\rangle &= (A_{11}|u_1\rangle + A_{21}|u_2\rangle)\otimes(B_{12}|v_1\rangle + B_{22}|v_2\rangle) \\ A\otimes B|u_2,v_1\rangle &= (A_{12}|u_1\rangle + A_{22}|u_2\rangle)\otimes(B_{11}|v_1\rangle + B_{21}|v_2\rangle) \\ A\otimes B|u_2,v_2\rangle &= (A_{12}|u_1\rangle + A_{22}|u_2\rangle)\otimes(B_{12}|v_1\rangle + B_{22}|v_2\rangle) \end{aligned} \tag{2.55}$$

Looking at the first line of Eq. (2.55), the basis vector $|u_1,b_1\rangle$ gets mapped to all basis vectors:

$$A\otimes B|u_1,v_1\rangle = A_{11}B_{11}|u_1,v_1\rangle + A_{11}B_{21}|u_1,v_2\rangle + A_{21}B_{11}|u_2,v_1\rangle + A_{21}B_{21}|u_2,v_2\rangle\,. \tag{2.56}$$

Combining this into matrix form leads to

$$A\otimes B = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix}. \tag{2.57}$$

Although we can always use this process to write tensor products of matrices as matrices in the full Hilbert space, the point of the notation is to avoid doing this, and maintain the products where possible. A typical situation where we may be able to do this is if we are carrying out separate operations on the subspaces. For example, we may have two qubits and make separate measurement

on each of them. The measurement operator (see later) will then have the form $\hat{A} \otimes \hat{B}$. In the full Hilbert space, this will be represented by a $4\times4$ matrix, for which we will need to find the eigenvalues and eigenvectors. However, if $|\psi_a\rangle$ and $|\phi_b\rangle$ are eigenvectors of the individual operators, respectively $\hat{A}$ and $\hat{B}$, with eigenvalues $\lambda_a$ and $\lambda_b$, we can use Eq.(2.51) to write

$$(A \otimes B)|\psi_a\rangle \otimes |\phi_b\rangle = (A|\psi_a\rangle) \otimes (B|\phi_b\rangle) = (\lambda_a|\psi_a\rangle) \otimes (\lambda_b|\phi_b\rangle) = \lambda_a\lambda_b\,|\psi_a\rangle \otimes |\phi_b\rangle\,. \tag{2.58}$$

We see that the eigenvalues of $A \otimes B$ are $\lambda_a\lambda_b$ and the eigenvectors are $|\psi_a\rangle \otimes |\phi_b\rangle$. It is much easier to solve two $2 \times 2$ eigenvalue problems than one $4 \times 4$.

### 2.8.2 Direct Sum

We will mainly use the tensor product to combine vector spaces. However there is another combination which is sometimes needed. Suppose we are working in some sub-space of the system we are describing, for example two states of an atom forming a qubit. To describe some manipulation of the qubit, we may need to introduce other states of the atom (we could be doing an operation on our qubit states which involves optical excitation to a third state). So we want to be able to add this state to our vector space, producing a larger space.

This sort of combination is known as the direct sum $\oplus$. Given two vector spaces $\mathcal{U}$ and $\mathcal{V}$ we can define $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$, where the symbol $\oplus$ is called the *sum direct*. The dimension of the space $\mathcal{W}$ is then the sum

$$\dim \mathcal{W} = \dim \mathcal{U} + \dim \mathcal{V}\,. \tag{2.59}$$

Contrast this with Eq. (2.43). Let $|\psi\rangle$ be a vector in $\mathcal{U}$ and $|\phi\rangle$ a vector in $\mathcal{V}$. A vector in $\mathcal{W}$ can be written as

$$|\Psi\rangle_{\mathcal{W}} = |\psi\rangle_{\mathcal{U}} + |\phi\rangle_{\mathcal{V}}\,, \tag{2.60}$$

where $|\psi\rangle_{\mathcal{U}}$ and $|\phi\rangle_{\mathcal{V}}$ are typically *not* normalized (i.e., they are not unit vectors).

We can also compose operators using the direct sum, so that $A \oplus B$ acts as follows

$$(A \oplus B)(|\psi\rangle_{\mathcal{U}} + |\phi\rangle_{\mathcal{V}}) = A|\psi\rangle_{\mathcal{U}} + B|\phi\rangle_{\mathcal{V}}. \tag{2.61}$$

and a general operator on $\mathcal{W}$ can be decomposed as $C = \sum_j A_j \oplus B_j$.

## 2.9 Problems

2.1. Consider the set of linear combinations of the functions $e^{ix}, e^{-ix}, \sin x, \cos x, x\sin x$. Show that this set satisfies all the conditions required to form a vector space. Find a basis which spans the space. What is its dimension?

2.2. Consider a state

$$|\psi\rangle = \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{5}}|2\rangle + \frac{1}{\sqrt{10}}|3\rangle$$

which is given in terms of three orthonormal eigenstates $|1\rangle$, $|2\rangle$ and $|3\rangle$ of an operator $\hat{B}$, satisfying $\hat{B}\,|n\rangle = n^2|n\rangle$.

Normalise $|\psi\rangle$ and calculate the value of $\langle\psi|\hat{B}|\psi\rangle$ for this state.

2.3. The Legendre polynomials $P_n(x)$ form a complete orthogonal set which can be used as a basis to represent functions defined in the range $-1 \geq x \geq 1$. They are conventionally normalised so that

$$\int_{-1}^{1} dx P_m(x) P_n(x) = \frac{2}{2n+1}\delta_{nm}.$$

Show that, if we expand a function $f(x)$ in a basis of Legendre polynomials as $f(x) = \sum_n a_n P_n(x)$, the expansion coefficients are given by the the formula

$$a_n = \frac{2n+1}{2}\int_{-1}^{1} dx P_n(x) f(x)$$

Consider the function $f(x) = x(1-x)$. Use this formula to calculate the first four coefficients in the Legendre polynomial expansion of the function, $a_0$, $a_1$, $a_2$ and $a_3$.

Show that the resulting expansion is exactly equal to the function. What can you deduce about $a_n$ for larger values of $n$?

The first four Legendre polynomials are

$$P_0(x) = 1,\ \ P_1(x) = x,\ \ P_2(x) = \tfrac{1}{2}(3x^2-1),\ \ P_3(x) = \tfrac{1}{2}(5x^3-3x).$$

2.4. $P_2$ is the (three dimensional) linear vector space of all second degree polynomials, that is functions of the form

$$p(x) = p_0 + p_1 x + p_2 x^2,$$

with the inner product defined as

$$\langle p|q\rangle = \int_{-1}^{1} p^*(x) q(x)\, dx$$

$\hat{D}$ is the differential operator

$$\hat{D}\, p(x) = p'(x)$$

Taking a basis $\{e_1(x) = 1, e_2(x) = x, e_3(x) = x^2\}$ (note this is *not* orthonormal), use the relationship

$$\hat{D}|e_j\rangle = \sum_i D_{ij}|e_i\rangle$$

to obtain $D$, the matrix representation of $\hat{D}$. Check your answer by applying it to the vector representing an arbitrary $p(x)$.

Use the Gram-Schmidt procedure to obtain an orthonormal basis from the one defined above. Note that the procedure as given in the lecture handout does not produce normalised basis states - you will have to normalise them.

Obtain the matrix representation of $\hat{D}$ in your new basis.

2.5. Consider the operator $\hat{A} = |1\rangle\langle 0| + |0\rangle\langle 1| + |2\rangle\langle 2|$.

(a) Write down the matrix representations for $\hat{A}$ using the orthonormal basis choices $\{|0\rangle, |1\rangle, |2\rangle\}$.

(b) Find the matrix representation of $\hat{A}$ in a basis consisting of $|2\rangle$ and the vectors $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. (Hint: first evaluate $\hat{A}|+\rangle$, $\hat{A}|-\rangle$ and $A|2\rangle$ ).

(c) Find the unitary matrix $U$ which transforms between the two bases in the previous parts.

(d) Show that the unitary transformation $A' = U^\dagger A U$ correctly transforms between the matrix representations $A$ and $A'$ from parts (a) and (b).

2.6. Consider the matrix $A = |1\rangle\langle 0| + |0\rangle\langle 0|$. What are the eigenvectors of $A$? Is $A$ diagonalisable? Is $A$ a normal matrix?

2.7. Consider the matrices $A = |2\rangle\langle 1| + |1\rangle\langle 0|$ and $B = |0\rangle\langle 2|$,

(a) find $A^2$;

(b) find $B^2$;

(c) find $AB$ and $BA$;

(d) find $(A+B)^2$.

You can do this by converting to standard matrix form, but try to do them using the outer product notation.

2.8. Use the inner product formula Eq.(2.15) to expand the Hermitian matrix

$$H = \begin{pmatrix} 2 & 1+2i \\ 1-2i & 3 \end{pmatrix}$$

in the basis of Pauli matrices. Verify that your result reproduces the original matrix.

2.9. Consider the vectors

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad \text{and} \quad |\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)\,.$$

(a) Find $|\Psi_1\rangle = |\phi\rangle \otimes |\phi\rangle - |\psi\rangle \otimes |\psi\rangle$ and $|\Psi_2\rangle = |\phi\rangle \otimes |\phi\rangle + |\psi\rangle \otimes |\psi\rangle$ in terms of the $\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}$

(b) Find the inner product $\langle\Psi_1|\Psi_2\rangle$ using $|\Psi_1\rangle$ and $|\Psi_2\rangle$ from above.

2.10. For the matrices

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

calculate $A \otimes B$, $B \otimes A$ and $\mathbb{1} \otimes A$. By direct calculation, show that $\mathrm{Tr}\{A \otimes B\} = \mathrm{Tr}\{A\}\mathrm{Tr}\{B\}$.

2.11. Consider the operator

$$M = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

(a) Find the eigenvalues and eigenvectors using Eq.(2.58).

(b) Express the operator as a $4 \times 4$ matrix in the full Hilbert space, and find the eigenvalues and eigenvectors of this. Verify you get the same result as in (a).

# 3 Operators

We have already met the idea of an operator in a vector space as something which transforms one vector into another. We found that when we define a basis, we can represent a vector by a matrix. In this section, we will look at some different kinds of operators which are important in quantum mechanics: Hermitian, Unitary and Projection operators.

## 3.1 Hermitian Conjugate and Hermitian Operators

The Hermitian conjugate or adjoint of the operator $\hat{A}$, written $\hat{A}^\dagger$, is defined by the requirement that

$$\langle\alpha|\hat{A}^\dagger|\beta\rangle = \langle\beta|\hat{A}|\alpha\rangle^* \tag{3.1}$$

for all states $|\alpha\rangle$ and $|\beta\rangle$. Here, we have used the matrix element of the operator between the states $|\alpha\rangle$ and $|\beta\rangle$,

$$\langle\alpha|\hat{A}|\beta\rangle = \int_{\substack{\text{all}\\\text{space}}} dx\, \psi_\alpha^*(x)\hat{A}\psi_\beta(x)\,. \tag{3.2}$$

This definition means that the matrix we can make from the matrix elements of the adjoint of $\hat{A}$ is the matrix adjoint of the one made from $\hat{A}$.

Suppose we want to evaluate the matrix element $\langle\beta|\hat{A}|\alpha\rangle$. We could do the calculation, following the definition, by operating with $\hat{A}$ on $|\alpha\rangle$, then find the braket of the result with $\langle\beta|$. However, if we know $\hat{A}^\dagger$, we can also get the result by first operating with $\hat{A}^\dagger$ on $|\beta\rangle$. If we then take the braket with $\langle\alpha|$, we get the complex conjugate of the result we want.

We can also then define the bra which is the adjoint of the ket $\hat{A}|\alpha\rangle$. This is

$$(\hat{A}|\alpha\rangle)^\dagger = \langle\alpha|\hat{A}^\dagger. \tag{3.3}$$

Operators acting on a bra act to the left, while on a ket, they act to the right. It is straightforward to show that, as for matrices, the adjoint of a combination of operators reverses their order: $(AB)^\dagger = B^\dagger A^\dagger$. This makes perfect sense in terms of the rule that the action is to the left on the bra.

An operator is said to be *Hermitian* if it is equal to its Hermitian conjugate, $\hat{A} = \hat{A}^\dagger$. Then, for any states $|\alpha\rangle$ and $|\beta\rangle$

$$\langle\alpha|\hat{A}|\beta\rangle = \langle\alpha|\hat{A}^\dagger|\beta\rangle, \tag{3.4}$$

so

$$\langle\alpha|\hat{A}|\beta\rangle = \langle\beta|\hat{A}|\alpha\rangle^* \tag{3.5}$$

Or, in integral form

$$\int dx\, \psi_\alpha^*(x)\hat{A}\psi_\beta(x) = \left[\int dx\, \psi_\beta^*(x)\hat{A}\psi_\alpha(x)\right]^* \tag{3.6}$$

Hermitian operators are very important in quantum mechanics, because they represent physical observables. A Hermitian operator will be represented by a Hermitian matrix.

**Exercise:** Show that the momentum operator $\hat{p} = -i\hbar d/dx$ is Hermitian.

## 3.2 Eigenvalues and Eigenstates

The operator eigenvalue problem is defined by the equation

$$\hat{A}|\alpha\rangle = \lambda_\alpha|\alpha\rangle \tag{3.7}$$

where the set of eigenvalues, $\lambda_\alpha$, and corresponding eigenstates, $|\rangle\alpha$, are to be determined. Clearly if $|\alpha\rangle$ is a solution, then so is $\mu|\alpha\rangle$, where $\mu$ is any scalar, so we conventionally normalise the eigenvalues by insisting $\langle\alpha|\alpha\rangle = 1$.

We already know that an operator relationship $\hat{A}|x\rangle = |y\rangle$ becomes, once a basis is chosen, the matrix equation $A|x\rangle = |y\rangle$. Thus the operator eigenvalue equation is equivalent to the matrix eigenvalue problem

$$A|\alpha\rangle = \lambda_\alpha|\alpha\rangle \tag{3.8}$$

The operator eigenstates thus have the same properties as those of the corresponding matrix. For a Hermitian operator, the eigenvalues must be real and the eigenstates orthogonal, $\langle\alpha|\beta\rangle = \delta_{\alpha\beta}$.

Generally, the order in which operators act is important: $\hat{A}\hat{B}|\psi\rangle \neq \hat{B}\hat{A}|\psi\rangle$. If they are in fact the same, we say that the operators *commute*. Then, the commutator

$$[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} = 0 \tag{3.9}$$

An important theorem states that if two operators commute, we can find functions which are simultaneously eigenstates of both.

**Proof**

Suppose $|\alpha\rangle$ is a non-degenerate eigenfunction of $\hat{A}$ with eigenvalue $\lambda_a$, as in Eq.3.7. Operating with $\hat{B}$ on both sides, we get

$$\hat{B}\hat{A}|\alpha\rangle = \hat{B}\lambda_a|\alpha\rangle = \lambda_a(\hat{B}|\alpha\rangle).$$

If $\hat{A}$ and $\hat{B}$ commute, we can write this as

$$\hat{A}(\hat{B}|\alpha\rangle) = \lambda_a(\hat{B}|\alpha\rangle).$$

Hence $\hat{B}|\alpha\rangle$ is an eigenvalue of $\hat{A}$ with eigenvalue $\lambda_a$. Since we have assumed that $|\alpha\rangle$ is non-degenerate, this means that $\hat{B}|\alpha\rangle$ must be some multiple of $|\alpha\rangle$, that is

$$\hat{B}|\alpha\rangle = \lambda_b|\alpha\rangle,$$

so $|\alpha\rangle$ is also an eigenfunction of $\hat{B}$.

The theorem also applies for degenerate eigenstates, but I will not prove it.

The matrix representation of an operator in the basis of its own eigenstates is particularly simple. The elements of the matrix are

$$A_{ij} = \langle\phi_i|\hat{A}|\phi_j\rangle = \langle\phi_i|\lambda_j|\phi_j\rangle = \lambda_j\langle\phi_i|\phi_j\rangle = \lambda_j\delta_{ij}$$

So the matrix is diagonal, with the eigenvalues along the diagonal:

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots \\ 0 & \lambda_2 & 0 & \dots \\ 0 & 0 & \lambda_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.10}$$

## 3.3 Unitary Operators

An operator $\hat{A}$ is unitary if and only if $\hat{A}^\dagger = \hat{A}^{-1}$. Note that it is common to use $\hat{U}$ to denote unitary operators. Since $\hat{U}^\dagger\hat{U} = \mathbb{1} = \hat{U}\hat{U}^\dagger$, all unitary operators are normal, and therefore diagonalisable.

Unitary operators are not generally Hermitian, so the eigenvalues, $\lambda_i$, may be complex, but will always have unit magnitude $|\lambda_i| = 1$.

**Proof**
We can write the spectral decomposition of $\hat{U}$ as

$$\hat{U} = \sum_i \lambda_i |x_i\rangle\langle x_i| .$$

Then the adjoint is

$$\hat{U}^\dagger = \sum_j \lambda_j^* |x_j\rangle\langle x_j| .$$

However, $\hat{U}^\dagger\hat{U} = \mathbb{1}$, so

$$\hat{U}^\dagger\hat{U} = \sum_{i,j} \lambda_i\lambda_j^* |x_j\rangle\langle x_j|x_i\rangle\langle x_i| = \sum_i |\lambda_i|^2 |x_i\rangle\langle x_i| = \mathbb{1} .$$

However, from the completeness relationship Eq.(2.27), the expansion coefficients in the unit operator must all be one, so $|\lambda_i| = 1$.

A unitary operator can be expressed as the exponent of an Hermitian operator, $\hat{H}$, according to $\hat{U} = \exp(i\hat{H})$. The exponent of an operator is *defined* by the power series expansion

$$\exp(\hat{A}) = \sum_{n=0}^{\infty} \frac{(\hat{A})^n}{n!} . \tag{3.11}$$

The $n^{\text{th}}$ power of an operator is straightforward: just multiply $\hat{A}$ $n$ times with itself. The expression in Eq. (3.11) is then well defined (provided the series converges). In general, we can construct any *analytic* function of an operator, that is, any function which has a power series expansion:

$$f(\hat{A}) = \sum_{n=0}^{\infty} f_n \hat{A}^n . \tag{3.12}$$

If $\hat{A}$ is diagonalizable with spectral decomposition $\hat{A} = \sum_i a_i |\phi_i\rangle\langle\phi_i|$ then one finds that

$$f(\hat{A}) = \sum_i f(a_i)|\phi_i\rangle\langle\phi_i|. \tag{3.13}$$

Therefore, $\hat{A}$ and $f(\hat{A})$ share the same eigenvectors, but the eigenvalues are related by $a_i \to f(a_i)$.

We our now in a position to prove that $\exp(i\hat{H})$ is unitary.

**Proof**
If $\hat{H}$ is Hermitian, then it has real eigenvalues $\{a_j\}$, so the operator $i\hat{H}$ has imaginary eigenvalues $ia_j$. Thus $\hat{U} = \exp(i\hat{H})$ will have eigenvalues $\exp(ia_j)$, which are complex numbers with unit modulus. It is therefore unitary.

Recall that $\hat{U}^\dagger$ has eigenvalues $\exp(ia_j)^* = \exp(-ia_j)$, so if $\hat{U} = \exp(i\hat{H})$ then $\hat{U}^\dagger = \exp(-i\hat{H})$.

## 3.4 Projection Operators

The idea of projecting onto a lower dimensional space should be familiar: a photograph is a projection of a three dimensional scene onto a two dimensional image. A bit more formally, the projection of a vector in three dimensions onto a two dimensional plane is the in-plane vector we get by dropping a perpendicular from the end of the original vector onto the plane. Its length will be the Cartesian inner product of the original vector and a unit vector, in the plane, which is parallel to the projection.

Since the projection takes us from one vector in our space to another, it can be described by an operator:

$$|y\rangle = \hat{P}|x\rangle \,, \tag{3.14}$$

where $|x\rangle$ is any vector in our space, and $|y\rangle$ is its projection in the lower dimensional subspace onto which we are projecting.

Clearly, once we have projected our vector, if we do so again, nothing more will happen. So a key property of a projection operator is that

$$\hat{P}^2 = \hat{P}\hat{P} = \hat{P} \,. \tag{3.15}$$

It follows from this that the eigenvalues of $\hat{P}$ must be $\lambda = 0$ or $\lambda = 1$.

**Proof**

If we have an eigenvector $|x\rangle$, so

$$\hat{P}|x\rangle = \lambda|x\rangle \,,$$

the linearity of $\hat{P}$ means that

$$\hat{P}^2|x\rangle = \hat{P}\lambda|x\rangle = \lambda^2|x\rangle \,.$$

However, for a projection operator $\hat{P}^2 = \hat{P}$, so we get

$$\hat{P}|x\rangle = \lambda^2|x\rangle \,.$$

Thus $\lambda^2 = \lambda$, which means that $\lambda = 0$ or $\lambda = 1$.

The eigenvectors corresponding to $\lambda = 1$ are unchanged following the projection, so they must be in the subspace onto which we are projecting – indeed they form a basis for this subspace. The eigenvectors corresponding to $\lambda = 0$ are entirely outside the subspace. We can thus use the spectral decomposition Eq.(2.29) to write $\hat{P}$ as

$$\hat{P} = \sum_{\alpha} |x_\alpha\rangle\langle x_\alpha| \,, \tag{3.16}$$

where the sum is over any basis of the subspace onto which we are projecting. The completness relationship, Eq.(2.27) is a special case of this. We can see from Eq.(3.16) that a projection operator must be Hermitian, $\hat{P} = \hat{P}^\dagger$.

The number of unit eigenvectors of a projector is equal to the dimension of the space onto which it projects. The trace of the projector is the sum of the eigenvalues, so the trace gives that dimension.

## 3.5 Problems

3.1. The parity operator $\hat{P}$ is defined in one-dimension so that

$$\hat{P}f(x) = f(-x)$$

for any function $f(x)$.

(a) Prove that $\hat{P}$ is an Hermitian operator.

(b) Prove that the eigenvalues of $\hat{P}$ are $\pm 1$.

(c) Show that any function $\phi(x)$ with definite parity, that is $\phi(-x) = \pm\phi(x)$, is an eigenfunction of $\hat{P}$.

(d) Show that, if

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$$

is the Hamiltonian for a one-dimensional system with a potential satisfying $V(x) = V(-x)$, then the commutator $[\hat{P}, \hat{H}] = 0$.

3.2. The exponent of a matrix $A$ is defined by the power series expansion

$$\exp(A) = \mathbb{1} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

Consider the matrix

$$A = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}$$

Calculate $A^2$ and $A^3$. Show that $A^3 = -\theta^2 A$. Hence show that

$$\exp\begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

3.3. Show that the matrix

$$P = \frac{1}{9}\begin{pmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{pmatrix}$$

is a projector. What is the dimension of the space it projects onto? Find an orthogonal basis which spans this space.

3.4. Show that the matrix

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

can be written in the form

$$A = 2P_2 - P_{-1}\,,$$

where $P_2$ and $P_{-1}$ are projectors which you should find.

3.5. Let $\hat{P}$ be a projector with *orthocomplement* $\hat{P}^\perp = \mathbb{1} - \hat{P}$ and let $\hat{U}$ and $\hat{V}$ be unitaries.

(a) Defining $\hat{W} = \hat{P}\otimes\hat{U} + \hat{P}^\perp\otimes\hat{V}$, find $\hat{W}\hat{W}^\dagger$. What kind of operator is $\hat{W}$?

(b) Letting $\hat{Q} = \hat{P}\otimes\hat{P} + \hat{P}^\perp\otimes\hat{P}^\perp$, find $\hat{Q}\hat{Q}$. What kind of operator is $\hat{Q}$?

3.6. A system consists of a pair of two level systems, $A$ and $B$, so the Hilbert space is a tensor product of a pair of two-dimensional spaces, $\mathscr{H}_A \otimes \mathscr{H}_B$, each spanned by the states $|0\rangle$ and $|1\rangle$.

Consider the operator

$$\hat{U} = \hat{P}_0 \otimes \mathbb{1} + \hat{P}_1 \otimes \sigma_x \,,$$

where the projectors are $\hat{P}_0 = |0\rangle\langle 0|$ and $\hat{P}_1 = |1\rangle\langle 1|$,

(a) Show that this operator is unitary.

(b) Find the matrix representation of this operator in a basis consisting of (in order) $|0\rangle \otimes |0\rangle$, $|0\rangle \otimes |1\rangle$, $|1\rangle \otimes |0\rangle$ and $|1\rangle \otimes |1\rangle$.

(c) Find the eigenvectors of $\hat{U}$.

(d) Find a factorisable 'input' state $|\Psi^{\text{in}}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, which becomes the maximally entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)$$

when $\hat{U}$ acts on it.

# 4 Quantum Mechanics

## 4.1 The Schrödinger Equation

The Schrödinger equation is usually introduced in the form of a partial differential equation. Using one dimensional notation (that is, $x$ rather than $\boldsymbol{r}$), it is

$$\left[\frac{\hat{p}^2}{2m} + V(x)\right]\psi = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\psi = i\hbar\frac{\partial\psi}{\partial t}\,. \tag{4.1}$$

More generally, we can write it in the language of operators and vectors as

$$\hat{H}|\psi\rangle = i\hbar\frac{\partial}{\partial t}|\psi\rangle\,. \tag{4.2}$$

Here $\hat{H}$ is the *Hamiltonian* of the physical system. $|\psi(t)\rangle$ is the *state* of the system, the vector, in the Hilbert space of the system, which represents it at time $t$. This form applies in any basis we use to expand the state of the system $|\psi\rangle$: The partial differential equation comes from working in a basis of the eigenstates of position, $|\psi(t)\rangle = \psi(x,t)$ (or $\langle x|\psi(t)\rangle = \psi(x,t)$). $\psi(x,t)$ is called the *wavefunction* of the system.

We will often consider separable solutions of the form $|\psi(t)\rangle = |\phi_n\rangle\exp(-iE_nt/\hbar)$, where $|\phi_n\rangle$ satisfys the *time independent* Schrödinger equation

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle\,. \tag{4.3}$$

This is an eigenvalue equation: the $|\phi_n\rangle$ are eigenvectors, and $E_n$ eigenvalues, of the operator $\hat{H}$. We can always separate out the time depence like this, provided that the Hamiltonian contains no terms explicitly involving $t$.

The state vector describing the system must be normalised. In vector notation, $\langle\psi|\psi\rangle = \langle\phi_n|\phi_n\rangle = 1$. In the position representation, $\phi_n(x) = \langle x|\phi_n\rangle$, this corresponds to the requirement

$$\int_{-\infty}^{\infty} dx\,|\phi_n(x)|^2 = 1. \tag{4.4}$$

The Hamiltonian is an Hermitian operator, so its eigenvalues are orthogonal, $\langle\phi_n|\phi_m\rangle = \delta_{nm}$, or

$$\int_{-\infty}^{\infty} dx\,\phi_n^*(x)\phi_m(x) = \delta_{nm}\,. \tag{4.5}$$

Note that, as in this section, I will tend to use one-dimensional notation (just the variable $x$) for the general development. This is intended to keep things simple, and it should be clear how it generalises to higher dimensions.

The state of the system at any time is a vector (or ray) in our Hilbert space. Thus the states at two different times, $|\psi(t)\rangle$ and $|\psi(0)\rangle$, will be will be connected by an operator:

$$|\psi(t)\rangle = \hat{U}(t)\,|\psi(0)\rangle\,. \tag{4.6}$$

If the Hamiltonian has no explicit time dependence, this operator is

$$\hat{U}(t) = \exp\left(-\frac{it}{\hbar}\hat{H}\right). \tag{4.7}$$

**Proof**
$|\psi(t)\rangle$ must satisfy the Schrödinger equation. Differentiating with respect to $t$, we see that it does, because

$$\frac{\partial}{\partial t}|\psi(t)\rangle = \frac{\partial \hat{U}}{\partial t}|\psi(0)\rangle = -\frac{i}{\hbar}\hat{H}\hat{U}|\psi(0)\rangle = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle\,.$$

As we saw in Sec.(3.3), an operator of this form is unitary. This means that the normalisation of the state is maintained when we use this operator to perform its time evolution. There are some cases where we can evaluate the exponent of $\hat{H}$ exactly, to solve a time evolution problem, but more frequently we use this as a formal solution in derivations.

## 4.2 Some Solutions to the Schrödinger Equation

In this, and the following sections, we will revise the solution of the Schrödinger equation for a couple of specific potentials. Although this course is not, for the most part, about solving the partial differential equation form, it is useful to have a few example systems to work with.

### 4.2.1 The Infinite Square Well

The first potential we shall consider is the infinite square well, $V(x) = 0$ for $|x| \le a/2$, $V(x) = \infty$ for $|x| > a/2$, the eigenfunctions are, for $|x| \le a/2$,

$$\begin{aligned}\phi_n(x) &= \sqrt{\frac{2}{a}}\cos\Big(\frac{n\pi x}{a}\Big) \quad (n \text{ odd})\\ \phi_n(x) &= \sqrt{\frac{2}{a}}\sin\Big(\frac{n\pi x}{a}\Big) \quad (n \text{ even},)\end{aligned} \tag{4.8}$$

and zero for $|x| > a/2$. The corresponding eigenvalues are $E_n = \hbar^2\pi^2n^2/2ma^2$. The cosine solutions have even *parity*, while the sines have odd parity. Pay attention to parity when evaluating integrals.

Some times the quantum well is defined to be between $x = 0$ and $x = a$. This clearly makes no difference to the energy levels, but the shift changes the form of the eigenfunctions to

$$\phi_n(x) = \sqrt{\frac{2}{a}}\sin\Big(\frac{n\pi x}{a}\Big) \quad 0 \le x \le a. \tag{4.9}$$

This form is valid for all $n$. Note that the parity for these functions corresponds to reflections about the point $x = a/2$.

### 4.2.2 The Harmonic Oscillator

For the harmonic oscillator $V(x) = \frac{1}{2}m\omega_c^2x^2$, the solutions are written in terms of *Hermite polynomials*, $H_n$:

$$\phi_n(x) = \frac{1}{(n!2^n a\sqrt{\pi})^{\frac{1}{2}}}H_n\Big(\frac{x}{a}\Big)\,e^{-x^2/2a^2}, \tag{4.10}$$

with $a = \sqrt{(\hbar/m\omega_c)}$. The eigenvalues are $E_n = (n + \frac{1}{2})\hbar\omega_c$.

The first three Hermite polynomials are:

$$H_0(x) = 1, \qquad H_1(x) = 2x \quad \text{and} \quad H_2(x) = 4x^2 - 2. \tag{4.11}$$

Hence

$$\phi_0(x) = \frac{1}{(a\sqrt{\pi})^{\frac{1}{2}}} e^{-x^2/2a^2}, \qquad \phi_1(x) = \frac{1}{(2a\sqrt{\pi})^{\frac{1}{2}}} 2(x/a) e^{-x^2/2a^2}, \qquad \phi_2(x) = \frac{1}{(8a\sqrt{\pi})^{\frac{1}{2}}} \left(4(x/a)^2 - 2\right) e^{-x^2/2a^2}. \tag{4.12}$$

For a three dimensional harmonic oscillator, with $V(\boldsymbol{r}) = \frac{1}{2}m\omega_c^2(x^2+y^2+z^2)$, the method of separation of variables can be used to show that the wavefunction can be written as a product of one dimensional solutions

$$\phi_{n_x,n_y,n_z}(\boldsymbol{r}) = \phi_{n_x}(x)\phi_{n_y}(y)\phi_{n_z}(z). \tag{4.13}$$

The corresponding eigenvalue is $E = (n_x + n_y + n_z + \frac{3}{2})\hbar\omega_c$

## 4.3 Raising and Lowering Operators for the Harmonic Oscillator

An elegant way of solving the harmonic oscillator problem uses *lowering and raising operators*. They form a prototype for a common method for solving quantum mechanics problems without calculating wavefunctions. We will use it again when we look at many particle systems, and it features strongly in treatments of quantum field theory.

The lowering operator ($\hat{a}$) and raising operator ($\hat{a}^\dagger$) are defined as

$$\hat{a} = \sqrt{\frac{m\omega_c}{2\hbar}}\,\hat{x} + i\frac{\hat{p}}{\sqrt{2m\omega_c\hbar}} \quad \text{and} \quad \hat{a}^\dagger = \sqrt{\frac{m\omega_c}{2\hbar}}\,\hat{x} - i\frac{\hat{p}}{\sqrt{2m\omega_c\hbar}}. \tag{4.14}$$

These are examples of operators which are not Hermitian. As the notation implies $\hat{a}^\dagger$ is the adjoint of $\hat{a}$.

Using $[\hat{x},\hat{p}] = i\hbar$, the commutator $[\hat{a},\hat{a}^\dagger] = 1$. We can write the Hamiltonian as

$$\hat{H} = (\hat{a}^\dagger\hat{a} + \tfrac{1}{2})\hbar\omega_c. \tag{4.15}$$

Hence, the eigenvectors of $\hat{H}$ are the same as those of the *number operator*, $\hat{n} = \hat{a}^\dagger\hat{a}$. These are the kets $|n\rangle$,

$$\hat{n}|n\rangle = \hat{a}^\dagger\hat{a}|n\rangle = n|n\rangle\,, \tag{4.16}$$

where $n$ is the corresponding eigenvalue. Note that we have not yet shown that $n$ must be an integer (though it is!).

When the raising and lowering operators act on the kets $|n\rangle$ they transform it into $|n+1\rangle$ and $|n-1\rangle$ respectively. That is

$$\hat{a}^\dagger|n\rangle \propto |n+1\rangle \quad \text{and} \quad \hat{a}|n\rangle \propto |n-1\rangle\,. \tag{4.17}$$

**Proof**

For the lowering case, we first act on either side of Eq.(4.16) with $\hat{a}$ to get

$$\hat{a}\hat{a}^\dagger(\hat{a}|n\rangle) = n(\hat{a}|n\rangle)\,.$$

Now using the commutator to swap the first two operators round, we find that

$$\hat{a}^\dagger\hat{a}(\hat{a}|n\rangle) = (n-1)(\hat{a}|n\rangle).$$

From this, we see that

$$\hat{a}|n\rangle = c\,|n-1\rangle$$

where $c$ is some unknown constant. The lowering operator acting on an eigenstate of $|n\rangle$ creates the state $|n-1\rangle$, with eigenvalue $n-1$.

We are now in a position to find the possible values of $n$. We know that the eigenvalues, $n$, satisfy

$$n = \langle n|\hat{n}|n\rangle = \langle n|\hat{a}^\dagger\hat{a}|n\rangle = |c|^2\langle n-1|n-1\rangle\,. \tag{4.18}$$

The right hand side of this equation must be greater than or equal to zero, so $n$ has a lower bound of zero. This means that the lowering process must terminates somewhere, that is, there is a state $|n_{\min}\rangle$ such that $\hat{a}|n_{\min}\rangle = 0$. Then

$$\hat{n}|n_{\min}\rangle = \hat{a}^\dagger\hat{a}|n_{\min}\rangle = 0 = n_{\min}|n_{\min}\rangle\,. \tag{4.19}$$

From this we see that $n_{\min} = 0$, so the allowed values of $n$ are non-negative integers $n = 0,1,2,\ldots$.

We can find $c$ from the requirement that $\langle n-1|n-1\rangle = 1$ in Eq.(4.18). This means that $|c|^2 = n$, so

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\,. \tag{4.20}$$

**Exercise:** Show that $\hat{a}^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle$

The raising and lowering operators are very useful for evaluating matrix elements involving states of the harmonic oscillator. As an example, we will calculate $\langle n'|x|n\rangle$.

From the definition of the raising and lowering operators, Eq.(4.14), the operator $\hat{x}$ can be written as

$$\hat{x} = \frac{1}{2}\sqrt{\frac{2\hbar}{m\omega_c}}\left(a + \hat{a}^\dagger\right) = \sqrt{\frac{\hbar}{2m\omega_c}}\left(a + \hat{a}^\dagger\right)$$

We know the action of the raising and lowering operators:

$$\hat{x}|n\rangle = \sqrt{\frac{\hbar}{2m\omega_c}}\left(\hat{a}|n\rangle + \hat{a}^\dagger|n\rangle\right) = \sqrt{\frac{\hbar}{2m\omega_c}}\left(\sqrt{n}|n-1\rangle + \sqrt{n+1}|n+1\rangle\right)$$

Then, using the orthogonality of the basis states,

$$\langle n'|x|n\rangle = \sqrt{\frac{\hbar}{2m\omega_c}}\left(\sqrt{n}\,\delta_{n',n-1} + \sqrt{n+1}\,\delta_{n',n+1}\right) \tag{4.21}$$

**Exercise:** Find the matrix representations of the raising and lowering operators $\hat{a}^\dagger$ and $\hat{a}$ in the basis of the eigenstates of the harmonic oscillator

## 4.4 Spin

All particles have an internal degree of freedom which is called spin (for some, the spin is zero). As its name suggests, this is mathematically something like angular momentum, though there is nothing physically corresponding to a rotation – indeeed, fundamental particles have no known structure, so it would be meaningless to think about them rotating.

Like angular momentum, spin is described by operators corresponding to three components $\hat{S}_x$, $\hat{S}_y$ and $\hat{S}_z$. These can be combined to give the total spin $\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$. The commutation relationship for the spin operators is

$$[\hat{S}_x, \hat{S}_y] = i\hbar \hat{S}_z \,, \tag{4.22}$$

along with its cyclic permuations.

Each component of the spin commutes with the total spin, so we can find simultaneous eigenstates of $\hat{S}^2$ and any one of the components. In defining a basis for spin, it is conventional to choose $\hat{S}_z$. We thus have basis states $|s,m\rangle$ which satisfy

$$\hat{S}^2|s,m\rangle = s(s+1)\hbar^2|s,m\rangle \,, \tag{4.23}$$

where $s$ takes half integer values $0, \frac{1}{2}, 1, \frac{3}{2} \ldots$. They also satisfy

$$\hat{S}_z|s,m\rangle = m\hbar|s,m\rangle \,, \tag{4.24}$$

where $m$ goes from $-s$ to $+s$ in integer steps, so there are $2s+1$ values. Note that these states are not eigenstates of the other components $\hat{S}_x$ and $\hat{S}_y$.

We will mainly be interested in spin $s = \frac{1}{2}$, which is the spin of an electron. Then $m$ takes two values, $-\frac{1}{2}$ and $\frac{1}{2}$. If the total spin is fixed, as it is for a fundamental particle, we can consider just a two dimensional Hilbert space, with basis states $|\frac{1}{2},\frac{1}{2}\rangle$ and $|\frac{1}{2},-\frac{1}{2}\rangle$. We normally abbreviate these by missing out the $s$ variable, and call them $|\frac{1}{2}\rangle$ and $|-\frac{1}{2}\rangle$ or $|\uparrow\rangle$ and $|\downarrow\rangle$.

For the spin-$\frac{1}{2}$, the operators for the components of the spin are represented, in this basis, by more-or-less the Pauli matrices: $\hat{S}_i = \frac{1}{2}\hbar\sigma_i$. Thus

$$\hat{S}_x = \frac{1}{2}\hbar\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_y = \frac{1}{2}\hbar\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{S}_z = \frac{1}{2}\hbar\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.25}$$

For $s = 1$, we get a set of $3 \times 3$ matrices etc.

The spin is an internal degree of freedom of a particle, which will also be able to nove in space. Hence the spin Hilbert space is really part of a larger Hilbert space made by a tensor product of the spatial and spin parts. However, there are circumstances where the physical operators are also of a tensor product form, in which case we can think of the spin part in isolation. For example, if we apply a magnetic field, the Hamiltonian will consist of tensor product of a spatial part, corresponding to the classical Lorentz force, and a spin part. For an electron, the spin part is

$$\hat{H} = \frac{1}{2} g\mu_B (B_x\sigma_x + B_y\sigma_y + B_z\sigma_z) \,, \tag{4.26}$$

where $\mu_B = |e|\hbar/2m_e$ is the *Bohr magneton*. The constant $g = 2(1 + \alpha/2\pi + \cdots) \approx 2.0023193043622 \pm 0.0000000000015$, is the *gyromagnetic ratio*. This is probably the most accurately measured quantity in physics.

If there are no other constraints, the direction of the magnetic field can always be chosen to correspond to the $z$ axis. Then

$$\hat{H} = \frac{1}{2} g\mu_B B\sigma_z \,. \tag{4.27}$$

Since this is diagonal, we can read off the eigenstates straight away. They are the basis states $|\uparrow\rangle$ and $|\downarrow\rangle$, with energies $E = \pm\frac{1}{2}g\mu_B B$.

## 4.5 Problems

4.1. At time $t = 0$, a one-dimensional harmonic oscillator with natural frequency $\omega_c$ is in the state

$$|\psi(0)\rangle = \frac{1}{\sqrt{5}}|0\rangle + \frac{2}{\sqrt{5}}|1\rangle,$$

where $|0\rangle$ and $|1\rangle$ are the ground and first excited state of the oscillator.

(a) Show that this state is normalised.

(b) Write down the state of the system at an arbitrary time $t$.

(c) Calculate the expectation value $\langle\psi(t)|\hat{x}|\psi(t)\rangle$.

4.2. Evaluate the following integrals:

(a) For the eigenfunctions of the harmonic oscillator,

$$\text{(i)} \int_{-\infty}^{\infty} dx\, \phi_0^*(x)\, x\, \phi_1(x) \quad \text{and} \quad \text{(ii)} \int_{-\infty}^{\infty} dx\, \phi_0^*(x)\, x\, \phi_2(x).$$

(b) For the eigenfunctions of the infinite quantum well,

$$\text{(i)} \int_{-a/2}^{a/2} dx\, \phi_1^*(x)\, x\, \phi_2(x) \quad \text{and} \quad \text{(ii)} \int_{-a/2}^{a/2} dx\, \phi_1^*(x)\, x^2\, \phi_1(x)$$

(c) For the harmonic oscillator eigenfunctions

$$\int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2\, \phi_0^*(x_1)\, \delta(x_1 - x_2)\, \phi_0(x_2).$$

4.3. Show that the raising and lowering operators satisfy the relationship for an operator and its adjoint

$$\langle\alpha|\hat{a}^\dagger|\beta\rangle = \langle\beta|\hat{a}|\alpha\rangle^*.$$

You may use the fact that, as we have already proved, the position and momentum operators are Hermitian.

4.4. Use the formula

$$\langle n'|x|n\rangle = \sqrt{\frac{\hbar}{2m\omega_c}}\left(\sqrt{n}\,\delta_{n',n-1} + \sqrt{n+1}\,\delta_{n',n+1}\right)$$

to evaluate the following matrix elements:

(a) $\langle 2|x|2\rangle$

(b) $\langle 1|x|2\rangle$

(c) $\langle 2|x|1\rangle$

(d) $\langle 0|x|2\rangle$

(e) $\langle 100|x|101\rangle$

4.5. By expressing the position operator $\hat{x}$ in terms of the raising and lowering operators, $\hat{a}$ and $\hat{a}^\dagger$, show that the matrix elements of $\hat{x}^2$ between simple harmonic oscillator states are:

$$\langle n'|\hat{x}^2|n\rangle = \frac{\hbar}{2m\omega_c}\left([(n+1)(n+2)]^{\frac{1}{2}}\,\delta_{n',n+2} + (2n+1)\,\delta_{n',n} + [n(n-1)]^{\frac{1}{2}}\,\delta_{n',n-2}\right).$$

Use this result to write down the matrix representation of $\hat{x}^2$ in this basis. (The matrix is infinite, so just give the first few rows and columns in the top left corner.)

4.6. A particle of mass $m$ moves in a two dimensional harmonic oscillator potential,

$$V(x,y) = \tfrac{1}{2} m\omega_c^2 (x^2 + y^2).$$

The eigenstates can be written as products of one dimensional harmonic oscillator eigenstates, $\phi_n$, in the $x$ and $y$ directions:

$$|n_x, n_y\rangle = \phi_{n_x}(x)\phi_{n_y}(y),$$

$\hat{a}_x$ and $\hat{a}_x^\dagger$ are lowering and raising operators for the $x$ component of the wavefunction:

$$\hat{a}_x \phi_{n_x}(x) = \sqrt{n_x}\,\phi_{n_x-1}(x) \quad \text{and} \quad \hat{a}_x^\dagger \phi_{n_x}(x) = \sqrt{n_x+1}\,\phi_{n_x+1}(x),$$

with

$$\hat{a}_x = \sqrt{\frac{m\omega_c}{2\hbar}}\,\hat{x} + i\frac{\hat{p}_x}{\sqrt{2m\omega_c\hbar}} \quad \text{and} \quad \hat{a}_x^\dagger = \sqrt{\frac{m\omega_c}{2\hbar}}\,\hat{x} - i\frac{\hat{p}_x}{\sqrt{2m\omega_c\hbar}}.$$

$\hat{a}_y$ and $\hat{a}_y^\dagger$ are defined similarly for the $y$ component.

(a) Give the $n_x$ and $n_y$ values corresponding to the ground and the (degenerate) first excited levels.

(b) Show that the angular momentum operator $\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$ can be written as

$$\hat{L}_z = i\hbar(\hat{a}_y^\dagger \hat{a}_x - \hat{a}_x^\dagger \hat{a}_y)$$

(c) Evaluate the result of this operator acting on the states of the ground and first excited levels obtained in part (a).

(d) Construct eigenstates of $\hat{L}_z$ from linear combinations of the two states corresponding to the first excited level.

4.7. Show that the eigenvalues of the full spin-$\frac{1}{2}$ Hamiltonian, Eq.(4.26), are $\pm\frac{1}{2}g\mu_B|B|$.

4.8. Show that each component of the spin commutes with the total spin, that is $[\hat{S}_x, \hat{S}^2] = 0$. Note that this is surprisingly tricky – you will have to do quite a lot of messing around with commutators.

# 5 Dynamics of a Two State Systems

In this section, we will look at a few methods for solving the time dependent Schrödinger equation for a two state system. There are many systems where only two state are important, and the approximation of neglecting all the other states works very well. An obvious physics example is the spin degree of freedom of a spin-$\frac{1}{2}$ particle such as an electron. If we can ignore spin-orbit coupling, the Hilbert space is two-dimensional, with just two basis states, which we conventionally take to be the spin-up and spin-down states.

In quantum computing, a qubit with two states is the basic unit of information. Physically, this may be a spin, but there are many other examples, including the polarisation states of a photon, a pair of energy levels in an atom, or two different values of the quantised flux through a superconducting loop. Quantum information processing is all about manipulating qubits with external fields, so the dynamics of such two state systems plays a central role.

As an example, we will consider the Hamiltonian

$$\hat{H} = \hbar\omega\sigma_x = \hbar\omega \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{5.1}$$

with the system initially in the state

$$|\psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{5.2}$$

The methods introduced here will work for any $2 \times 2$ Hamiltonian, as long as it has no time dependence. Most generalise to large Hilbert spaces, though the calculations become more difficult.

## 5.1 Expanding in the Eigenstates

The most straight-forward way of solving the matrix form of the Schrödinger equation is to express the general solution as a linear combination of the eigenstates:

$$|\psi(t)\rangle = \sum_n a_n |\phi_n\rangle e^{-iE_n t/\hbar}. \tag{5.3}$$

If we know $|\psi(0)\rangle$, the initial state of the system, we can use this to find the coeffients $a_n$, which gives the general solution for all times.

For a $2 \times 2$ matrix, there are just two eigenvalues, so the sum contains just two terms. For our example, $E_0 = \hbar\omega$ and $E_1 = -\hbar\omega$, so

$$|\psi(t)\rangle = a_0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-i\omega t} + a_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\omega t} \tag{5.4}$$

(the normalisation of the eigenvectors can be absorbed into the coefficients). We next use the initial state to find these coefficients. Putting $t = 0$, we have

$$|\psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = a_0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + a_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{5.5}$$

Hence $a_0 + a_1 = 1$ and $a_0 - a_1 = 0$, which has solution $a_0 = a_1 = \frac{1}{2}$. Then

$$|\psi(t)\rangle = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-i\omega t} + \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\omega t} = \begin{pmatrix} \cos\omega t \\ -i\sin\omega t \end{pmatrix}. \tag{5.6}$$

This method generalises readily to larger matrices. It is easier then to use Eq.(2.15) to obtain the coefficients, $a_n = \langle\phi_n|\psi(0)\rangle$.

## 5.2 Finding the Time Evolution Operator, $\hat{U}(t)$.

An alternative way to find the time dependence of the state is to calculate the time evolution operator, $\hat{U}(t)$, from Eq.(4.7):

$$\hat{U}(t) = \exp\left(-\frac{it}{\hbar}\hat{H}\right). \tag{5.7}$$

For our $\hat{H}$, this is

$$\hat{U}(t) = \exp(-i\omega t\sigma_x) = \mathbb{1} + (-i\omega t\sigma_x) + \frac{1}{2!}(-i\omega t\sigma_x)^2 + \dots . \tag{5.8}$$

We know that $\sigma_x^2 = \mathbb{1}$, so $\sigma_x^3 = \sigma_x$, $\sigma_x^4 = \mathbb{1}$ etc. Hence

$$\begin{aligned}\hat{U}(t) &= \left(1 - \frac{1}{2!}(\omega t)^2 + \frac{1}{4!}(\omega t)^4 + \dots\right)\mathbb{1} - i\left((\omega t) - \frac{1}{3!}(\omega t\sigma_x)^3 + \dots\right)\sigma_x \\ &= \cos\omega t\,\mathbb{1} - i\sin\omega t\,\sigma_x\,.\end{aligned} \tag{5.9}$$

We can now find the time dependence of any initial state. For our example

$$|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle = \left[\cos\omega t\begin{pmatrix}1&0\\0&1\end{pmatrix} - i\sin\omega t\begin{pmatrix}0&1\\1&0\end{pmatrix}\right]\begin{pmatrix}1\\0\end{pmatrix} = \cos\omega t\begin{pmatrix}1\\0\end{pmatrix} - i\sin\omega t\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}\cos\omega t\\ -i\sin\omega t\end{pmatrix}, \tag{5.10}$$

as we found before.

This method generalises to any two level Hamiltonian, because we can always express it as a sum over Pauli matrices and the unit matrix. It is not too difficult to show that, if we define the Pauli vector (a vector of Pauli matrices)

$$\boldsymbol{\sigma} = \sigma_x\,\boldsymbol{i} + \sigma_y\,\boldsymbol{j} + \sigma_z\,\boldsymbol{k}\,, \tag{5.11}$$

so that

$$a_x\sigma_x + a_y\sigma_y + a_z\sigma_z = \boldsymbol{a}\cdot\boldsymbol{\sigma}\,, \tag{5.12}$$

then

$$\exp(i\boldsymbol{a}\cdot\boldsymbol{\sigma}) = \exp ia(\boldsymbol{n}\cdot\boldsymbol{\sigma}) = \cos a\,\mathbb{1} + i\sin a\,(\boldsymbol{n}\cdot\boldsymbol{\sigma})\,. \tag{5.13}$$

With this notation, the most general $2\times 2$ Hermitian matrix can be written (Section 2.7) as

$$H = \hbar(b_0\,\mathbb{1} + \boldsymbol{b}\cdot\boldsymbol{\sigma})\,. \tag{5.14}$$

Then

$$U(t) = \exp(-iHt/\hbar) = \exp(-ib_0t)\exp(-it\boldsymbol{b}\cdot\boldsymbol{\sigma})\,, \tag{5.15}$$

so $\boldsymbol{a} = -\boldsymbol{b}t$. Note that the exponent only separates into a product like this because $\mathbb{1}$ commutes with the Pauli matrices. In general $\exp(A+B) \neq \exp A \exp B$ for operators.

This approach can be extended to larger Hermitian matrices, by defining generalisations of the Pauli matrices. However, here we will look at a method of finding the exponent of any (diagonalisable) matrix using a change of basis.

Suppose we have a matrix, $A$, and diagonalise it by performing a similarity transformation to get $D = S^{-1}AS$, or $A = SDS^{-1}$. Then

$$A^2 = (SDS^{-1})(SDS^{-1}) = SD^2S^{-1}. \tag{5.16}$$

Similarly, for any power $A^n = SD^nS^{-1}$. Hence, using $\mathbb{1} = SS^{-1} = S\mathbb{1}S^{-1}$,

$$\exp(A) = S\mathbb{1}S^{-1} + SDS^{-1} + \frac{1}{2!}SD^2S^{-1} + \frac{1}{3!}SD^3S^{-1} + \cdots = S\exp(D)S^{-1}. \tag{5.17}$$

For a diagonal matrix, the powers are just diagonal matrices with the same power of the diagonal elements, that is, if

$$D = \begin{pmatrix} \lambda_0 & 0 & \cdots \\ 0 & \lambda_1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad \text{then} \quad (D)^n = \begin{pmatrix} \lambda_0^n & 0 & \cdots \\ 0 & \lambda_1^n & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{5.18}$$

Hence

$$\exp(D) = \mathbb{1} + \begin{pmatrix} \lambda_0 & 0 & \cdots \\ 0 & \lambda_1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} \lambda_0^2 & 0 & \cdots \\ 0 & \lambda_1^2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} + \cdots = \begin{pmatrix} \exp(\lambda_0) & 0 & \cdots \\ 0 & \exp(\lambda_1) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{5.19}$$

Then

$$\exp(A) = S\begin{pmatrix} \exp(\lambda_0) & 0 & \cdots \\ 0 & \exp(\lambda_1) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}S^{-1}. \tag{5.20}$$

Now let us apply this to the problem we are considering. As the matrix is Hermitian, the required $S$ is unitary, with columns made from the eigenvectors of $H$. Thus

$$H = \hbar\omega\sigma_x = \hbar\omega\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{1}{2}\hbar\omega\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{5.21}$$

Then

$$U(t) = \exp\left(\frac{-iHt}{\hbar}\right) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} e^{-i\omega t} & 0 \\ 0 & e^{i\omega t} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} \cos\omega t & -i\sin\omega t \\ -i\sin\omega t & \cos\omega t \end{pmatrix}, \tag{5.22}$$

as before.

## 5.3 The Bloch Sphere

The Bloch sphere is a convenient way of visualising the time dependence of the state of a two level system. The most general for for the state of such a system is $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $\alpha$ and $\beta$ are arbitrary complex numbers. This corresponds to 4 real parameters. However, a physical state must be normalised, which gives a constraint $|\alpha|^2 + |\beta|^2 = 1$, and the overall phase has no significance, so we actually only need two parameters. There are many ways to choose these, but the most natural is to define real parameters $\theta$ and $\phi$ such that

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle. \tag{5.23}$$

If we think of $\theta$ and $\phi$ as representing the angular part of spherical polar coordinates, any state of a two level system can be represented by a point on the surface of a sphere. This is known as the *Bloch sphere*. The 'north pole' ($\theta = 0$) represents the state $|0\rangle$ and the south pole $|1\rangle$. $\theta = \pi/2$ is the equator, which is traversed by varying the $\phi$ value from 0 to $2\pi$.

In the example solved above, comparing the result for $|\psi t\rangle$ with this expression, we see that $\theta = 2\omega t$ and $\phi = 3\pi/2$. The trajectory then is a circle corresponding to a line of longitude.

## 5.4 Problems

5.1. Consider the Hamiltonian

$$H = \hbar\omega \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$$

and the initial state

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Find the time evolution of this state using (a) the expansion in eigenstates and (b) the Pauli expansion, verifying that you get the same result in each case.

5.2. Show that the square of $\boldsymbol{a}.\boldsymbol{\sigma}$, Eq.(5.12), is equal to $|a|^2\mathbb{1}$. The rest of the proof of Eq.(5.13) is then similar to the special case in the notes.

5.3. In this question we consider a particle with spin $s = 1$, which has a three dimensional Hilbert space, so we are dealing with $3 \times 3$ matrices. We will chose a magnetic field acting in the $y$ direction, for which the Hamiltonian takes the form

$$H = \mu_B B \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & i & 0 \\ -i & 0 & i \\ 0 & -i & 0 \end{pmatrix}.$$

If the initial state of the system is

$$|\psi(0)\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

find the time dependent state $|\psi(t)\rangle$.

5.4. Find the unitary $U(t) = \exp(-iHt/\hbar)$ for the matrix $H$ from the previous question. Check that it gives the same result when acting on the same input state.

# 6 Operators and Measurements

## 6.1 The Born Rule

Suppose we have represented the operator corresponding to a measurement of a physical quantity by a matrix, $A$, and the state by a vector $|\phi\rangle$. The possible results of the measurement are the eigenvalues of $A$, $\lambda_a$, with corresponding eigenvectors $|a\rangle$. These different results occur randomly, with the probability of obtaining the result $\lambda_a$ given by

$$p_a = |\langle a|\phi\rangle|^2 . \tag{6.1}$$

This process encapsulates the idea of quantum uncertainty: there is a randomness implicit in the measurement process which means that, except in special cases, we cannot predict the result of a measurement, only the probabilities of various results which occur at random.

After the measurement has been made, the state of the system *collapses* into the eigenvector $|a\rangle$ which corresponds to the $\lambda_a$ obtained by the measuring apparatus. This description of the measurement process, involving the collapse of the wavefunction, is fundamental to the *Copenhagen interpretation* of quantum mechanics.

If we expand $|\phi\rangle$ in the basis of eigenstates of $\hat{A}$,

$$|\phi\rangle = \sum_a \langle a|\phi\rangle |a\rangle, \tag{6.2}$$

then

$$\hat{A}|\phi\rangle = \sum_a \lambda_a \langle a|\phi\rangle |a\rangle. \tag{6.3}$$

Hence the matrix element

$$\langle\phi|\hat{A}|\phi\rangle = \sum_a \lambda_a \langle a|\phi\rangle\langle\phi|a\rangle = \sum_a \lambda_a |\langle a|\phi\rangle|^2 = \sum_a \lambda_a p_a . \tag{6.4}$$

This is just the expectation value of the measurement, $\overline{\lambda}$, so

$$\overline{\lambda} = \langle\hat{A}\rangle = \langle\phi|\hat{A}|\phi\rangle . \tag{6.5}$$

**Exercise:** Suppose we have a system with $s = 1$ in the spin state

$$|\phi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

and measure $\hat{S}_z$, which is represented by the matrix

$$S_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{6.6}$$

What are the possible results and their probabilities? What is the expectation value?

The theory given above assumes that the measured quantity, $a$, can only take discrete outcomes. We need to think a bit more about how to handle cases where the result is a continous variable, such as measurements of position ($\hat{x}$) or momentum ($\hat{\mathbf{p}}$). The expectation value looks the same:

$$\langle\hat{x}\rangle = \langle\phi|\hat{x}|\phi\rangle , \tag{6.7}$$

where $\hat{x} = x$ if $|\phi\rangle$ is in the position representation, $|\phi\rangle = \phi(x)$. The probabilities of diffent outcomes works slightly differently. We start from the idea that an eigenstate of $\hat{x}$, with eigenvalue $x$, is a Dirac delta function:

$$|x\rangle = \delta(x' - x)\,, \tag{6.8}$$

where $x'$ is the variable. Then,

$$\langle x|\phi\rangle = \int dx'\, \delta(x' - x)\phi(x') = \phi(x)\,. \tag{6.9}$$

Using our previous result, this would give the probability of getting result $x$ as $p_x = |\phi(x)|^2$. This is not quite right: for a continuous variable, the probability of getting exactly a given value is zero. What we have calculated is the *probability density*. The probability of getting a result in the infinitesimal interval $x \to x + dx$ is

$$p_{x\to x+dx} = |\phi(x)|^2\, dx\,. \tag{6.10}$$

## 6.2 Measurements as Projections

The process of the collapse of the state during a measurement can (almost) be described by the projection operator $\hat{P}_a = |a\rangle\langle a|$,

$$\hat{P}_a|\phi\rangle = \langle a|\phi\rangle\, |a\rangle\,. \tag{6.11}$$

This is wrong, because the final state is not correctly normalised. Since $\hat{P}_a$ is Hermitian and $\hat{P}_a^2 = \hat{P}_a$, the normalisation of this state is

$$\langle\phi|\hat{P}_a^\dagger \hat{P}_a|\phi\rangle = \langle\phi|\hat{P}_a\hat{P}_a|\phi\rangle = \langle\phi|\hat{P}_a|\phi\rangle\,. \tag{6.12}$$

Hence the correctly normalised collapse process is given by

$$|\phi\rangle \to \frac{\hat{P}_a|\phi\rangle}{\sqrt{\langle\phi|\hat{P}_a|\phi\rangle}}\,. \tag{6.13}$$

We also see that

$$\langle\phi|\hat{P}_a|\phi\rangle = \langle a|\phi\rangle\,\langle\phi|a\rangle = |\langle a|\phi\rangle|^2 = p_a\,, \tag{6.14}$$

the probability of the result $a$. Eq.(6.13) and Eq.(6.14) tells us what to do if the eigenstates of $\hat{A}$ are degenerate. We simply replace $\hat{P}_a$ by the projector onto the subspace corresponding to the degenerate set of eigenvectors,

$$\hat{P}_a = \sum_{\lambda_i = \lambda_a} |a_i\rangle\langle a_i|\,. \tag{6.15}$$

Using the projector formalism, we now have expressions for the probability of a result and the subsequent state of the system which work for both degenerate and non-degenerate cases.

Note that a projector cannot be unitary (except for the trivial case $\hat{P} = \mathbb{1}$), since it has zero eigenvalues, so the collapse process cannot be described by the Schrödinger equation. It is not possible to describe both the system and the measurement apparatus as a bigger quantum system. This is central to the *measurement problem* in the Copenhagen interpretation. Measurements are special, but there is no clear criterion for what constitutes a measurement: is a cat dying a measurement, or do we have to look in the box first?

The collapse following measurement means that measurement processes are an important way of initialising a quantum system into a particular state. If the state we want is an eigenstate for a particular measurement, we perform that measurement on the system, and if the outcome is right, we know that the system is in the required state. If it is the wrong outcome, the measurement has to be repeated until the right value is obtained.

The expression for the expectation value can be written in terms of an operator trace:

$$\langle \hat{A} \rangle = \langle \phi | \hat{A} | \phi \rangle = \mathrm{Tr}\left\{ |\phi\rangle\langle\phi| \hat{A} \right\} . \tag{6.16}$$

**Proof**

We can calculate the trace in any basis we chose, so, using the eigenvectors $|a\rangle$ as the basis

$$\mathrm{Tr}\left\{ |\phi\rangle\langle\phi| \hat{A} \right\} = \sum_a \langle a|\phi\rangle\langle\phi|\hat{A}|a\rangle = \sum_a \langle a|\phi\rangle \lambda_a \langle\phi|a\rangle = \sum_a \lambda_a |\langle a|\phi\rangle|^2 = \sum_a \lambda_a p_a ,$$

as required.

Putting $\hat{A} = \hat{P}_a$, gives

$$p_a = \langle \phi | \hat{P}_a | \phi \rangle = \mathrm{Tr}\left\{ |\phi\rangle\langle\phi| \hat{P}_a \right\} \tag{6.17}$$

## 6.3 Measurements using the Tensor Product Formalism

When we consider measurements on a combined system, say of two qubits, the tensor product formalism can be very useful for simplifying calculations. Often we will be carrying out indpendent measurements on each subsystem, so the operator corresponding to the measurement will be of the form $\hat{A}_1 \otimes \hat{A}_2$. For a two qubit system, this can be represented by a $4 \times 4$ matrix, and the calculation completed exactly as above. However, finding eigenvalues and eigenvectors of a $4 \times 4$ matrix is tedious, and it is better to stick with the tensor product form. Let us look at a particular example, when we have two qubits in the state

$$|\phi\rangle = \frac{1}{\sqrt{2}} (|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle) = \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) , \tag{6.18}$$

where the second form is just a compression of the notation. Suppose we do a simultaneous measurment of $\sigma_x$ on each qubit, so the meaasurement operator is $\hat{A} = \sigma_x \otimes \sigma_x$. We saw in Section 2.8.1 that the eigenvalues and eigenvectors of this operator can be written in terms of those of $\sigma_x$ as

$$\begin{aligned}
&\lambda = +1 \times +1 = 1 \qquad |\psi_{++}\rangle = |+\rangle \otimes |+\rangle = \frac{1}{2}(|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \\
&\lambda = +1 \times -1 = -1 \qquad |\psi_{+-}\rangle = |+\rangle \otimes |-\rangle = \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle) \\
&\lambda = -1 \times +1 = -1 \qquad |\psi_{-+}\rangle = |-\rangle \otimes |+\rangle = \frac{1}{2}(|00\rangle + |01\rangle - |10\rangle - |11\rangle) \\
&\lambda = -1 \times -1 = 1 \qquad |\psi_{--}\rangle = |-\rangle \otimes |-\rangle = \frac{1}{2}(|00\rangle - |01\rangle - |10\rangle + |11\rangle) .
\end{aligned} \tag{6.19}$$

We can now work out the probabilities for the various results. For example, the probability of getting $= 1$ for both measurements is

$$p_{++} = |\langle \psi_{++} | \phi \rangle|^2 = \left( \frac{1}{2\sqrt{2}} \right)^2 |(\langle 00| + \langle 01| + \langle 10| + \langle 11|)(|01\rangle + |10\rangle)|^2 = \frac{1}{8}|1+1|^2 = \frac{1}{2} . \tag{6.20}$$

Similarly

$$p_{+-} = |\langle \psi_{+-} | \phi \rangle|^2 = \left( \frac{1}{2\sqrt{2}} \right)^2 |(\langle 00| - \langle 01| + \langle 10| - \langle 11|)(|01\rangle + |10\rangle)|^2 = \frac{1}{8}|-1+1|^2 = 0 , \tag{6.21}$$

and $p_{-+} = 0$ and $p_{--} = 1/2$.

## 6.4 Problems

6.1. A system is initially in the state

$$|\psi_0\rangle = \frac{1}{\sqrt{7}}\left(\sqrt{2}|\phi_1\rangle + \sqrt{3}|\phi_2\rangle + |\phi_3\rangle + |\phi_4\rangle\right),$$

where the $|\phi_n\rangle$ are eigenstates of the system's Hamiltonian, $\hat{H}$, satisfying $\hat{H}|\phi_n\rangle = n^2 E_H|\phi_n\rangle$.

(a) If the energy is measured, what values will be obtained and with what probabilities?

(b) Consider an operator $\hat{A}$, whose action on $|\phi_n\rangle$ is defined by $\hat{A}|\phi_n\rangle = (n+1)a_0|\phi_n\rangle$. If $\hat{A}$ is measured, what values will be obtained and with what probabilities?

(c) Suppose the measurement of the energy yields $4E_H$. If $\hat{A}$ is measured immediately afterwards, what value will be obtained?

6.2. The state of a system, $|\psi\rangle$ can be expanded in an orthonormal basis $|\phi_i\rangle$ ($i = 1, 2$) as

$$|\psi\rangle = \frac{1}{\sqrt{5}}\left(|\phi_1\rangle + 2|\phi_2\rangle\right).$$

A measurement is performed on the system corresponding to an operator which is represented in this basis by the matrix

$$A = \begin{pmatrix} 1 & 8 \\ 8 & -11 \end{pmatrix}$$

(a) What are the possible results of the measurement?

(b) With what probabilities will these results occur?

6.3. An observable, in a Hilbert space with dimension 3, has values '2' and '−1'. It corresponds to the operator $A = 2P_2 - P_{-1}$, where, $P_2$ and $P_{-1}$ are represented by matrices

$$P_2 = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad P_{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

A system in this Hilbert space is represented by the state

$$|\psi\rangle = \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

(a) Verify that $P_2$ and $P_{-1}$ are projectors.

(b) If a measurement of the observable $A$ is made on this system, determine the probabilities of the two values being obtained.

(c) In each case, find the state of the system after the measurement has been made.

6.4. Consider a spin-1 particle with an orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle\}$ and the operator

$$M = i|2\rangle\langle 1| + |0\rangle\langle 0| - i|1\rangle\langle 2| \tag{6.22}$$

(a) Let $|\psi_\theta\rangle = (|0\rangle + |1\rangle + e^{i\theta}|2\rangle)/\sqrt{3}$ and find $M|\psi_\theta\rangle$. For what values of $\theta$ is this an eigenstate of $M$?

(b) What is the expectation of $M$ with respect to state $|\psi_\theta\rangle$. Confirm that the expectation value is a real number.

(c) Show that $P_+ = (\mathbb{1} + M)/2$ and $P_- = (\mathbb{1} - M)/2$ are projectors, and find a decomposition of $M$ in terms of projectors.

(d) If we measure $M$ on state $|\psi_\theta\rangle$ and obtain outcome "+1", what is the normalised quantum state after the measurement.

6.5. Consider the observable $M = \sigma_z \otimes \mathbb{1}$, where $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$, acting in the Hilbert space of a system consisting of two qubits.

(a) If we perform a measurement of $M$ on state $|\psi\rangle$, what are the projectors for the different measurement outcomes? Use the notation $|00\rangle$ etc for the two qubit states.

(b) If the system is in the state $|\psi\rangle = \cos(\theta)|\Psi^+\rangle + \sin(\theta)|\Psi^-\rangle$, where

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \tag{6.23}$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\,, \tag{6.24}$$

what are the possible collapsed states after measuring $M$ and with what probabilities do they occur?

# 7 Mixed States and the Density Operator

## 7.1 Pure and Mixed States

So far we have considered that the state vector $|\psi\rangle$ contains a complete description of our system. In reality, however, we very rarely know for sure what the state of the system actually is. Even if we perform a measurement to create a particular state, any measurement apparatus is necessarily imperfect, so there will always be a finite probability that the system will be in a different state. If we want to predict the outcome of an experiment on this system, we have to take into account the possibilities that it is not in the state we intended. We call such a mixture of states a *mixed state*, in contrast to a *pure state* where we know the state vector with certainty.

Note that this is not the same as the quantum uncertainty for a system in a pure state – the fact that we cannot predict the results of a measurement even if we know the state for sure. Assuming we know the probabilities of each of the different states, we could, in principle, work out the results of the experiment for each possible state, and take an average weighted by these probabilities. However, there is a nicer method which describes a system in a mixed state by an operator, the *density operator*.

Note also that the classical uncertainty is subjective: I could have a system for which I know the state (with some level of precision), then give it to you without telling you what state it is in. As far as you are concerned, it could be in any state, so your prediction for the results of the measurement would be different from mine – and presumably less accurate. Again, this is different from the quantum uncertainty. If we both know the state of the system, we may get different results when we perform a measurement, but our theoretical predictions would be the same.

It is important to be clear that we cannot describe this situation by a superposition of two or more different states. As an example, consider that Alice prepares a qubit in the state $|0\rangle$ or in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ depending on the outcome of a balanced (50:50) coin toss. If Alice notes the result of the coin toss, she will know the state of the system. However, suppose she gives the qubit to Bob without passing on this information. How does Bob describe the state before any measurement? First, we *cannot* say that the state is $\frac{1}{2}|0\rangle + \frac{1}{2}|+\rangle$, (indeed, this is not even normalized) because that would just be a different pure state.

## 7.2 The Density Operator

The key to the solution is to find a description that gives the expectation values correctly. Eq.(6.16) gives us a clue. If we replace $|\phi\rangle\langle\phi|$ by

$$\hat{\rho} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +|\,, \tag{7.1}$$

we get

$$\overline{\lambda} = \mathrm{Tr}\left\{\hat{\rho}\hat{A}\right\} = \frac{1}{2}\mathrm{Tr}\left\{|0\rangle\langle 0|\hat{A}\right\} + \frac{1}{2}\mathrm{Tr}\left\{|+\rangle\langle +|\hat{A}\right\} = \frac{1}{2}\langle 0|\hat{A}|0\rangle + \frac{1}{2}\langle +|\hat{A}|+\rangle\,. \tag{7.2}$$

This is the correctly weighted probability, taking into account the expectation values for the $|0\rangle$ and $|+\rangle$ states.

The result generalises. We describe a system in a mixed state with this classical uncertainty using a *density operator* instead of a vector. This is the operator

$$\hat{\rho} = \sum_k p_k |\phi_k\rangle\langle\phi_k|\,, \tag{7.3}$$

where the $p_k$ are the probabilities of different states $|\phi_k\rangle$ in the mixture. These must sum up to one ($\sum_k p_k = 1$) and the $|\phi_k\rangle$ are normalized states, but they do not need to be complete or orthogonal, they just have to include all the possible states in our mixture. Note that we can also describe our pure state by a density operator: there is only one state in the sum, and it must have probability 1, so $\hat{\rho} = |\phi\rangle\langle\phi|$.

Of course, we can express an operator in any basis we chose, so our density operator may not have the simple form of Eq.(7.3). However, we can always find a diagonal representation

$$\hat{\rho} = \sum_\alpha \lambda_\alpha |\alpha\rangle\langle\alpha|, \tag{7.4}$$

where $|\alpha\rangle$ form a complete orthonormal basis, $0 \le \lambda_\alpha \le 1$, and $\sum_\alpha \lambda_\alpha = 1$. We can interpret these $\lambda_\alpha$ as probabilities that the system is in the pure state $|\alpha\rangle$. If the $|\phi_k\rangle$ In Eq.(7.3) are orthogonal, the $|\alpha\rangle$ are equivalent to the $|\phi_k\rangle$.

In general for an operator to be a valid density operator, it must have the following properties:

1. It must be Hermitian: $\hat{\rho}^\dagger = \hat{\rho}$,
2. It must give normalised probabilities: $\mathrm{Tr}\{\hat{\rho}\} = 1$,
3. The probabilities cannot be negative, so $\hat{\rho}$ must be positive: $\langle\psi|\hat{\rho}|\psi\rangle \ge 0$ for any $|\psi\rangle$.

A mixed state can be slightly mixed or very mixed and there are many ways to quantify the mixedness. One simple way is called the purity $P(\hat{\rho}) = \mathrm{Tr}\{\hat{\rho}^2\}$. If $\hat{\rho}$ is pure then $\hat{\rho}^2 = \hat{\rho}$ and so $P(\hat{\rho}) = \mathrm{Tr}\{\hat{\rho}\} = 1$. In contrast, for the maximally mixed state $\hat{\rho} = \mathbb{1}/D$ in a Hilbert space of dimension $D$, we have

$$P(\hat{\rho}) = \mathrm{Tr}\left\{\hat{\rho^2}\right\} = \mathrm{Tr}\{\mathbb{1}^2\}/D^2 = D/D^2 = 1/D. \tag{7.5}$$

**Exercise:** show that $P(\hat{U}\,\hat{\rho}\,\hat{U}^\dagger) = P(\hat{\rho})$ whenever $\hat{U}$ is unitary. Unitary evolution, such as that due to the action of the Hamiltonian, maintains the purity of a mixed state.

As an example, let us calculate the purity of the density matrix in Eq.(7.1). We have

$$\hat{\rho} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +| = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{4}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|) = \frac{1}{4}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}. \tag{7.6}$$

Thus

$$\hat{\rho}^2 = \frac{1}{16}\begin{pmatrix} 10 & 4 \\ 4 & 2 \end{pmatrix} \tag{7.7}$$

so the purity is $\mathrm{Tr}\{\hat{\rho}^2\} = 3/4$.

We obviously want to know how the density operator evolves in time due to the action of the Hamiltonian. The evolution of the can be derived directly from the Schrödinger equation $i\hbar\partial|\phi\rangle/\partial t = \hat{H}|\phi\rangle$:

$$\begin{aligned} i\hbar\frac{d\hat{\rho}}{dt} &= i\hbar\frac{d}{dt}\sum_k p_k|\phi_k\rangle\langle\phi_k| \\ &= i\hbar\sum_k\left\{\frac{dp_k}{dt}|\phi_k\rangle\langle\phi_k| + p_k\left[\left(\frac{d}{dt}|\phi_k\rangle\right)\langle\phi_k| + |\phi_k\rangle\left(\frac{d}{dt}\langle\phi_k|\right)\right]\right\} \\ &= i\hbar\frac{\partial\hat{\rho}}{\partial t} + \hat{H}\hat{\rho} - \hat{\rho}H \\ &= [\hat{H},\hat{\rho}] + i\hbar\frac{\partial\hat{\rho}}{\partial t}. \end{aligned} \tag{7.8}$$

This is sometimes known as the *von Neumann equation*. In most problems, the second term will not be present – it only appears when the probabilities of being in different state of the mixture change with time, so the $p_k$ have an explicit time dependence.

We also need to extend the rules which describe how we calculate the outcomes of measurements to deal with mixed states described by density operators. We have already shown that the expectation value of an operator, $A$, can be written as a trace.

$$\langle \hat{A} \rangle = \mathrm{Tr}\left\{\hat{\rho}\hat{A}\right\} , \tag{7.9}$$

where $\rho$ is the density operator representing the mixed state. Similarly, the expression for the probability of outcome $a$, Eq.(6.17), generalises to

$$p_a = \mathrm{Tr}\left\{\hat{\rho}\hat{P}_a\right\} . \tag{7.10}$$

For a pure state $|\phi\rangle$, we know that the (unnormalised) state after the measurement is $\hat{P}_a|\phi\rangle$. Thus

$$\hat{\rho} = |\phi\rangle\langle\phi| \to \hat{P}_a|\phi\rangle\langle\phi|\hat{P}_a = \hat{P}_a\hat{\rho}\hat{P}_a . \tag{7.11}$$

Of course, this $\hat{\rho}$ will not be normalised, with unit trace. We must divide by $\mathrm{Tr}\left\{\hat{P}_a\hat{\rho}\hat{P}_a\right\} = \mathrm{Tr}\left\{\hat{\rho}\hat{P}_a\right\}$, so the correct evolution is

$$\hat{\rho} \to \frac{\hat{P}_a\hat{\rho}\hat{P}_a}{\mathrm{Tr}\left\{\hat{\rho}\hat{P}_a\right\}} . \tag{7.12}$$

This also generalises to the case where $\hat{\rho}$ represents a mixed state.

We now have a theoretical framework for measurements which works for degenerate or non-degenerate measurement results, and for pure and mixed system states. This is not quite the complete story, because we have assumed that our measurement apparatus is perfect, and never gives a wrong result. Our results can be generalised to include this – the theory of imperfect measurements –but we will not cover this here.

## 7.3 Solving the von Neumann Equation

In this section, we will look at solving the von Neumann equation to find the time dependence of the density operator for a two state system. As an example, we will solve the problem in Section 5 using this formalism. Thus $H = \hbar\omega\sigma_x$, and the initial state is a pure state, $|\psi(0)\rangle = (1\,0)^T$.

The approach we will use is to expand the density operator in a basis of the Pauli matrices:

$$\hat{\rho}(t) = a_0(t)\mathbb{1} + a_x(t)\sigma_x + a_y(t)\sigma_y + a_z(t)\sigma_z . \tag{7.13}$$

Actually, this is too general, because we know that, for a valid density operator, $\mathrm{Tr}\left\{\hat{\rho}\right\} = 1$. Since the Pauli matrices have zero trace, this means that $a_0 = 1/2$, and it can have no time dependence. We can thus write, slightly redefining the coefficients,

$$\hat{\rho}(t) = \frac{1}{2}\left(\mathbb{1} + a_x(t)\sigma_x + a_y(t)\sigma_y + a_z(t)\sigma_z\right) . \tag{7.14}$$

We will first write the initial state in this form. For the pure state $|\psi(0)\rangle = (1\,0)^T$, we have

$$\begin{aligned}\hat{\rho}(0) = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} &= \frac{1}{2}\left(\mathbb{1} + a_x(0)\sigma_x + a_y(0)\sigma_y + a_z(0)\sigma_z\right)\\ &= \frac{1}{2}\left(\left(\begin{matrix}1 & 0\\ 0 & 1\end{matrix}\right) + a_x(0)\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + a_y(0)\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} + a_z(0)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right) .\end{aligned} \tag{7.15}$$

It is easy to see that this requires $a_z(0) = 1$ and $a_x(0) = a_y(0) = 0$ (remember this is a real vector space, so the coefficients must be real). These can also be obtained from the inner product formula for the coefficients, Eq.(2.15).

Now we substitute the expansion into the Von Neumann equation. There is no explicit time dependence for the density operator, so we get

$$i\hbar\frac{d\hat{\rho}}{dt} = [\hat{H},\hat{\rho}] = \frac{1}{2}\hbar\omega[\sigma_x,(\mathbb{1} + a_x\sigma_x + a_y\sigma_y + a_z\sigma_z)]\,. \tag{7.16}$$

Evaluating the commutators, we have $[\sigma_x,\mathbb{1}] = 0$, $[\sigma_x,\sigma_x] = 0$, $[\sigma_x,\sigma_y] = 2i\sigma_z$ and $[\sigma_x,\sigma_z] = -2i\sigma_y$, so

$$i\hbar\frac{1}{2}\left(\frac{da_x}{dt}\sigma_x + \frac{da_y}{dt}\sigma_y + \frac{da_z}{dt}\sigma_z\right) = i\hbar\omega\left(a_y\sigma_z - a_z\sigma_y\right) \tag{7.17}$$

Equating coefficients of the Pauli matrics, we get

$$\frac{da_x}{dt} = 0 \qquad \frac{da_y}{dt} = -2\omega a_z \qquad \frac{da_x}{dt} = 2\omega a_y\,. \tag{7.18}$$

Thus $a_x$ is a constant, which we know to be zero from the initial conditions. The other two are best solved by differentiating the second and substituting from the third:

$$\frac{d^2a_y}{dt^2} = -2\omega\frac{da_z}{dt} = -4\omega^2 a_z\,. \tag{7.19}$$

Thus $a_y(t) = A\cos 2\omega t + B\sin 2\omega t$, and from the initial condition, $a_y(0) = 0$, we know that $A = 0$. Then

$$\frac{da_y}{dt} = 2\omega B\cos 2\omega t = -2\omega a_z\,, \tag{7.20}$$

so the initial condition $a_z(0) = 1$ requires $B = -1$. Putting the bits together, we get

$$\begin{aligned}\hat{\rho}(t) &= \frac{1}{2}\left(\mathbb{1} - \sin 2\omega t\,\sigma_y + \cos 2\omega t\,\sigma_z\right) = \frac{1}{2}\begin{pmatrix}1+\cos 2\omega t & i\sin 2\omega t\\ -i\sin 2\omega t & 1-cos2\omega t\end{pmatrix}\\ &= \begin{pmatrix}\cos^2\omega t & i\sin\omega t\cos\omega t\\ -i\sin\omega t\cos\omega t & \sin^2\omega t\end{pmatrix}.\end{aligned} \tag{7.21}$$

This is the density matrix corresponding to the pure state we found before, Eq.(5.6).

## 7.4 The Bloch Sphere Again

Previously, we defined the Bloch sphere by using a parametrisation of the state as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle\,..\tag{7.22}$$

This is not the usual definition; it is normally defined in terms of the expectation values of the Pauli matrices. To see the equivalence, we calculate these expectation values for the above state. We get

$$\langle\sigma_x\rangle = \left(\cos\left(\frac{\theta}{2}\right) \quad e^{-i\phi}\sin\left(\frac{\theta}{2}\right)\right)\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\cos\left(\frac{\theta}{2}\right)\\ e^{i\phi}\sin\left(\frac{\theta}{2}\right)\end{pmatrix} = \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)(e^{-i\phi} + e^{i\phi}) = \sin\theta\cos\phi\,. \tag{7.23}$$

Similarly, $\langle\sigma_y\rangle = \sin\theta\sin\phi$ and $\langle\sigma_z\rangle = \cos\theta$. These are the $x$, $y$ and $z$ coordinates of the point $(\theta,\phi)$ in spherical polars, with radial coordinate $r = 1$.

With this definition, we can extend our Bloch sphere description to mixed states described by a density matrix, using Eq.(7.9). While pure states are always on the surface of the sphere, mixed states occupy the interior. As an example, for the density matrix in Eq.(7.1), we have

$$\hat{\rho} = \frac{1}{4}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}, \tag{7.24}$$

so

$$\langle \sigma_x \rangle = \mathrm{Tr}\left\{\hat{\rho}\sigma_x\right\} = \mathrm{Tr}\left\{\frac{1}{4}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right\} = \frac{1}{4}\mathrm{Tr}\left\{\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}\right\} = \frac{1}{2}. \tag{7.25}$$

Similarly, $\langle \sigma_y \rangle = 0$ and $\langle \sigma_z \rangle = 1/2$. The point representing this is a distance $r = 1/\sqrt{2}$ from the origin.

## 7.5 Problems

7.1. An observable, $A$, in a two dimensional Hilbert space is described by the matrix operator

$$A = \begin{pmatrix} 2 & 2i \\ -2i & -1 \end{pmatrix}.$$

(a) Find the eigenvalues, $a_j$, and eigenvectors of this matrix.

(b) Hence express the matrix in the form $A = a_1 P_1 + a_2 P_2$, where $P_1$ and $P_2$ are projectors.

A qubit is initialised such that there is a probability 4/5 that it is in the state

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

The other possibility, ocurring with probability 1/5, is that the qubit is in a state which is completely unknown to the experimenter.

(c) Find the density matrix which describes the experimenter's knowledge of this qubit.

(d) If the observable $A$ is measured, with what probabilities will each of the $a_j$ occur?

(e) What is the density matrix describing the state of the system after each measurement?

(f) Show that the density matrices after the measurement correspond to pure states. Identify the pure states.

7.2. Show that the density matrix

$$\rho = \frac{1}{25}\begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}$$

represents a pure state and find that state.

7.3. An observable in a three dimensional Hilbert space is represented by the operator

$$\hat{A} = |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 2| + |2\rangle\langle 1|.$$

Calculate the probability that the outcome of the measurement will be zero for states described by the following density matrices:

$$(a)\,\rho = \frac{1}{4}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (b)\,\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad (c)\,\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

7.4. Let $\hat{\rho}$ be a density operator represented by the matrix $\rho(0) = \frac{1}{2}(\mathbb{1} + a_x\sigma_x + a_y\sigma_y + a_z\sigma_z)$. Use the expansion in eigenstates to calculate the matrix $\rho(t)$ given the Hamiltonian $H = \hbar\omega\sigma_z$. Describe the path of this state on the Bloch sphere.

# 8 Many Particle Systems

This section is concerned with systems containing many particles, considering in particular the case where they are indistinguishable, for example, all electrons. We will look at two approaches to the description of such many particles state. The first can be summarised as 'saying which state each particle is in', while the second corresponds to 'saying how many particles are in each state'.

## 8.1 The Many Particle Wavefunction

We will start with the wavefunction description of a many particles state. The wavefunction is a function of all the particle coordinates (and spins, which we shall ignore for now): $\psi(x_1, x_2, \ldots x_N; t)$. The interpretation in terms of probability is extended in a straight forward way; the probability $P(x_1, x_2, \ldots x_N; t)$ of finding *simultaneously* each particle $i$ in a volume element $dx_i$ about $x_i$ is

$$P(x_1, x_2, \ldots x_N; t)\, dx_1\, dx_2 \ldots dx_N = |\psi(x_1, x_2, \ldots x_N; t)|^2\, dx_1\, dx_2 \ldots dx_N. \tag{8.1}$$

If, for example, we are only interested in the probability for particle 1, we get this by integrating the above expression over all the other particle coordinates. The normalisation condition comes from integrating over all the coordinates (and summing over spins if included):

$$\int dx_1 \int dx_2 \ldots \int dx_N\, |\psi(x_1, x_2, \ldots x_N; t)|^2 = 1. \tag{8.2}$$

It may be possible for an operators representing physical a observable to be associated with a single-particle, acting only on the coordinates associated with that particle. Other operators act on two (or more) particles.

An example of a single particle operator is the momentum operator: we can define a momentum for the $i^{\text{th}}$ particle $\hat{p}_i = -i\hbar\nabla_i$. The operator for the total momentum of all the particles is then $\hat{P} = \sum_i \hat{p}_i$.

Most interactions, such as the Coulomb interaction, act on pairs of particles, so they are two-particle operators of the form $\hat{V}(x_i, x_j)$. The operator for the total interaction energy is then

$$\hat{V} = \frac{1}{2} \sum_{ij} \hat{V}(x_i, x_j)\,. \tag{8.3}$$

where the factor 1/2 prevents double counting in the sum.

## 8.2 Non-Interacting Particles

If we have a system of non-interacting particles, the Hamiltonian can be written as

$$\hat{H} = \sum_i \frac{1}{2m_i} \hat{p}_i^2 + V_i(x_i) = \sum_i \hat{H}_i\,. \tag{8.4}$$

The method of separation of variables can then be used to show that the eigenfunctions are of the form

$$\psi(x_1, x_2 \ldots) = \phi_a(x_1)\, \phi_b(x_2) \ldots\,, \tag{8.5}$$

with eigenvalues $E = E_a + E_b + \ldots$, where the $\phi_a(x_i)$ and $E_a$ etc are the eigenstates of the single particle Schrödinger equation corresponding to Hamiltonian $\hat{H}_i$.

**Exercise:** Show that this separation works.

We can also express this in the language of tensor product states. The Hilbert space for the many particle system is the tensor product of the spaces for each particle: $\mathscr{U} = \mathscr{U}_1 \otimes \mathscr{U}_2 \otimes \ldots \mathscr{U}_N$. The non-interacting Hamiltonian then takes the form $\hat{H} = \hat{H}_1 \otimes \hat{H}_2 \otimes \ldots \hat{H}_N$, and the separable states are

$$|a,b,c\ldots\rangle = |a\rangle \otimes |b\rangle \otimes |c\rangle \ldots . \tag{8.6}$$

Of course, we will often be dealing with states which are not separable like this. But we can always use these tensor products of single particle states as a basis to describe more interesting physics. Recall that states which do not factorise are said to be entangled.

This product form is only correct if the particles are *distinguishable*, for example, in the hydrogen atom when we are dealing with an electron and a proton. It does not work if the particles are truly indistinguishable, for example, both electrons. The problem occurs because, with these simple products, the state $\phi_a(x_1)\phi_b(x_2)$ and $\phi_b(x_1)\phi_a(x_2)$ are two distinct states. In fact, they correspond to having an electron in state $a$, and another in state $b$, so they are really the same state.

## 8.3 Indistinguishable Particles

If two particles cannot be distinguished in any way, it is essential that the results predicted for any measurement must be unchanged when we swap the labels $1,2\ldots$ which we arbitrarily assigned to the coordinates in the expressions for wavefunctions, Hamiltonian *etc*. This is not quite the same as saying that the wavefunction must be unchanged; there are, in fact, important phase factors associated with the swap, which do not affect any measurements. The following discussion is restricted to a system with just two particles, but it generalises to many particle systems, where we must have the same symmetries for swapping *any* pair of identical particles.

To formalise the treatment, we introduce a *particle exchange operator*, $\hat{P}_{12}$, which has the effect of interchanging the labels 1 and 2 in the function which it acts on. Any valid operator associated with measurements must not distinguish between the particles, so, for example, $\hat{H}(1,2) = \hat{H}(2,1)$. Now consider the operation

$$\hat{P}_{12}\hat{H}(1,2)\psi(1,2) = \hat{H}(2,1)\psi(2,1) = \hat{H}(1,2)\hat{P}_{12}\psi(1,2) \Rightarrow [\hat{P}_{12},\hat{H}(1,2)] = 0,$$

so the exchange operator commutes with the Hamiltonian (and any other meassurement operator). This means that we can always find eigenstates of $\hat{H}$ which are also eigenstates of $\hat{P}_{12}$.

The eigenvalues of $\hat{P}_{12}$ are found by solving $\hat{P}_{12}\psi(1,2) = \lambda\psi(1,2)$. Operating on this equation again with $\hat{P}_{12}$, we get

$$(\hat{P}_{12})^2\psi(1,2) = \lambda\hat{P}_{12}\psi(1,2) = \lambda^2\psi(1,2).$$

But two exchanges are equivalent to doing nothing, so $(\hat{P}_{12})^2 = \mathbb{1}$, which means that $\lambda^2 = 1$, and $\lambda = \pm 1$.

We postulate that a physically correct wavefunction must be either symmetric or antisymmetric when any pair of identical particles is exchanged. Considerations of relativistic generalisations of the Schrödinger equation lead to the *spin-statistics theorem*, which states that for particles with integer spin $(0,1,\ldots)$, the wavefunction must be symmetric, while for half-integer spins $(\frac{1}{2},\frac{3}{2},\ldots)$ antisymmetry is required. The two types of particles are known, respectively, as bosons and fermions.

## 8.4 Non-Interacting Indistinguishable Particles

Returning to the problem of two non-interacting particles, it is now clear that the wavefunction $\psi(x_1,x_2) = \phi_a(x_1)\phi_b(x_2)$ does not have the right symmetry if the particles are identical. For indistinguishable particles, the functions $\phi_a(x_1)\phi_b(x_2)$ and $\phi_a(x_2)\phi_b(x_1)$ are degenerate eigenstates for

the two particle system. These are used to create the symmetric and antisymmetric combinations, which are eigenstates not just of the Hamiltonian, but also the particle exchange operator:

$$\psi(x_1,x_2) = \frac{1}{\sqrt{2}}[\phi_a(x_1)\phi_b(x_2) + \phi_a(x_2)\phi_b(x_1)] \qquad \text{(symmetric - bosons)}$$
$$\psi(x_1,x_2) = \frac{1}{\sqrt{2}}[\phi_a(x_1)\phi_b(x_2) - \phi_a(x_2)\phi_b(x_1)] \qquad \text{(antisymmetric - fermions),} \tag{8.7}$$

where the factor of $1/\sqrt{2}$ is required for normalisation. It is necessary to be a bit more careful in the case $a = b$, so the two particles are in the same state. Then the boson form is $\psi(x_1,x_2) = \phi_a(x_1)\phi_a(x_2)$. For fermions, the two terms cancel, so it is not possible to find a wavefunction for two fermions in the same state; this is the *Pauli exclusion principle*.

We can again write these states in the tensor product notation, as

$$|a,b\rangle = \frac{1}{\sqrt{2}}(|a\rangle \otimes |b\rangle + |b\rangle \otimes |a\rangle) \tag{8.8}$$

and

$$|a,b\rangle = \frac{1}{\sqrt{2}}(|a\rangle \otimes |b\rangle - |b\rangle \otimes |a\rangle)\,. \tag{8.9}$$

Note that the notation $|a,b\rangle$ is generally used to indicates a correctly symmetrised combination of the single particle states, as in these equations, for the type of particle they are describing.

**Exercise:** Show that the wavefunctions $\psi(x_1,x_2)$ of Eq.(8.7) are correctly normalised if the single particle wavefunctions $\phi_a$ and $\phi_b$ are normalised.

The symmetry has a profound effect on the eigenstates of many particle systems. For example, the ground state of two non-interacting bosons, like that for distinguishable particles, consists of both particles in the lowest energy level; for fermions we can only put one particle in the lowest level, with the other going into the next level up, leading to a higher total energy. Also, the degeneracy of the many particle system is affected: for distinguishable particles, the states $\phi_a(x_1)\phi_b(x_2)$ and $\phi_a(x_2)\phi_b(x_1)$ would form a degenerate pair, while there is only one legal combination for identical particles of each type.

If we put $x_1 = x_2 = x$, for fermions, $\psi(x,x) = 0$; there is no possibility of finding two fermions in the same place. By contrast, for bosons $|\psi(x,x)|^2 = 2|\phi_a(x)|^2\,|\phi_b(x)|^2$, which gives *twice* the probability of finding the particles in the same place as for distinguishable particles. This shows that the symmetrisation of the wavefunction for indentical particles introduces *correlations* even when there are no interactions. In the next section, we shall see these comments are not quite right, and have to be modified slightly when spin is taken into account.

### 8.4.1 Adding Spin to the Picture

If the wavefunctions consist of a spin part as well as a spatial part, the same symmetry principles apply. In this section, I shall consider mainly the example of two spin-$\frac{1}{2}$ particles, which, by the spin-statistics theorem, are necessarily fermions with a wavefunction which is antisymmetric for particle exchange. I will use the tensor product notation to combine the spatial and spin parts of the state for each particle, writing states like $\phi_a(x) \otimes |\uparrow\rangle$ and $\phi_b(x) \otimes |\downarrow\rangle$. However, I will combine the states for the two particles just by multiplying them together.

For the two particles, in spatial states $a$ and $b$ there are four single particle states to consider: $\phi_a(x_i) \otimes |\uparrow\rangle$, $\phi_a(x_i) \otimes |\downarrow\rangle$ $\phi_b(x_i) \otimes |\uparrow\rangle$ and $\phi_b(x_i) \otimes |\downarrow\rangle$. By considering all the possible pairs, these can

be used to create four anti-symmetric states, with one particle in $a$, the other in $b$. An example of such a state would be

$$\Psi(x_1,x_2)=\frac{1}{\sqrt{2}}\left((\phi_a(x_1)\otimes|\uparrow\rangle)(\phi_b(x_2)\otimes|\downarrow\rangle)-(\phi_b(x_1)\otimes|\downarrow\rangle)(\phi_a(x_2)\otimes|\uparrow\rangle)\right). \tag{8.10}$$

However, it is unusual to write the states of the two electron system in this way. The normal set of four states we would chose are the *singlet* state

$$\Psi_s(x_1,x_2)=\frac{1}{\sqrt{2}}\left(\phi_a(x_1)\phi_b(x_2)+\phi_a(x_2)\phi_b(x_1)\right)\otimes\frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle-|\downarrow\rangle|\uparrow\rangle). \tag{8.11}$$

and the three *triplet* states

$$\Psi_t(x_1,x_2)=\frac{1}{\sqrt{2}}\left(\phi_a(x_1)\phi_b(x_2)-\phi_a(x_2)\phi_b(x_1)\right)\otimes\begin{cases}|\downarrow\rangle|\downarrow\rangle\\ \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle+|\downarrow\rangle|\uparrow\rangle)\\ |\uparrow\rangle|\uparrow\rangle\end{cases}. \tag{8.12}$$

We have separated the states into a spatial start and spin part. Note that, if we multiply the two parts out, we get the same states we used in Eq.(8.10). This is just a different set of combinations of these states. The spin states we have created are eigenstates of the total spin operator, $\hat{S}^2=\hat{S}_1^2+\hat{S}_2^2$. For the singlet state, the total spin is $s=0$ (recall that the eigenvalues of $\hat{S}^2$ are written as $s(s+1)\hbar^2$), while the triplet states correspond to $s=1$. The three triplet states are the eigenstates of $\hat{S}_z$ corresponding to $m=-1,0,+1$. You will demonstrate this in one of the problems for this section.

The states in Eq.(8.11) and Eq.(8.12) are orthogonal to each other and normalised. They are also antisymmetric for particle exchange, as we require for fermions. However they achieve this in different ways: the singlet state has a symmetric spatial part and an anti-symmetric spin part, while for the triplet states, the spatial part is antitsymmetric and the spin part symmetric. If the spatial parts of the states are the same, that is $a=b$, then the triplet state is not allowed, because its wavefunction is zero. The singlet state, with its symmetric spatial part, is still possible; two fermions can go in the same spatial state provided they have opposite spins.

The previous comments about correlations also need to be revised when spin is considered. We can say for sure that if the particles have the same spin, then $\psi(x_1,x_2)=0$, so they cannot be found at the same place. If they have opposite spin, it depends on whether they are in the triplet or singlet state: for the triplet $\psi(x_1,x_2)=0$ again, but for the singlet the symmetric spatial state means that the probability is enhanced, like the boson case discussed previously. This is important when we add interactions to the picture. The Coulomb interaction between two electrons is repulsive, so the different symmetries mean that the triplet states tends to be lower energy than the corresponding singlet configuration, as the electrons have a higher probability of being close together.

### 8.4.2 Many Particle Systems

The considerations for bosons and fermions also apply when there are more than two identical particles. The rule then is that we must have the correct symmetry or anti-symmetry when we exchange *any* pair of particles. So for three fermions, if we start from $\psi(1,2,3)$, this must be combined with $-\psi(2,1,3)$, $+\psi(2,3,1)$ and $-\psi(3,2,1)$. For bosons the signs in the combinations are all positive.

The way to achieve this for non-interacting bosons is to write the wavefunction as a linear combination of every possible permutation of particles and states. So for three bosons

$$\begin{aligned}\psi(x_1,x_2,x_3)=\frac{1}{\sqrt{3!}}\big[&\phi_a(x_1)\phi_b(x_2)\phi_c(x_3)+\phi_a(x_1)\phi_b(x_3)\phi_c(x_2)+\phi_a(x_3)\phi_b(x_1)\phi_c(x_2)+\\ &\phi_a(x_2)\phi_b(x_1)\phi_c(x_3)+\phi_a(x_2)\phi_b(x_3)\phi_c(x_1)+\phi_a(x_3)\phi_b(x_2)\phi_c(x_1)\big]\end{aligned} \tag{8.13}$$

For three fermions, the same form could be used, but with alternating signs. However, the wavefunction is almost always written in the form of a *Slater determinant*

$$|a,b,c\rangle = \psi(x_1,x_2,x_3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \phi_a(x_1) & \phi_b(x_1) & \phi_c(x_1) \\ \phi_a(x_2) & \phi_b(x_2) & \phi_c(x_2) \\ \phi_a(x_3) & \phi_b(x_3) & \phi_c(x_3) \end{vmatrix}, \tag{8.14}$$

as this gets the signs right automatically. The generalisation to $N$ particles in $N$ states is obvious (the prefactor is $1/\sqrt{N!}$).

Exactly the same forms can be used when we have spin as well as spatial parts to the wavefunction. It is just necessary to consider the labels $a$, $b$ etc to contain both spin and spatial quantum numbers, and the coordinates $x_i$ include the spin degree of freedom. It is still possible to make states which are combinations of spatial and spin parts, but these are not in the Slater determinant form.

Again, for bosons we have to be careful if two or more states are the same. Considering the case of all the particles in the same state $a$, we get a wavefunction of the form:

$$\psi(x_1,x_2,\ldots x_N) = \phi_a(x_1)\phi_a(x_2)\ldots\phi_a(x_N).$$

There is nothing to prevent all the particles going into the same state, even for macroscopically large $N$. Indeed, this is exactly what happens in a Bose-Einstein condensate. For fermions, the Slater determinant form of the wavefunction evaluates to zero if two of the state labels are the same, so it is not possible to write down a Slater determinant wavefunction which violates the exclusion principal.

Although these wavefunctions can only be eigenstates for non-interacting particles, they are frequently used as basis states for solving problems with interactions. If we make sure that the basis states have the correct symmetry, then any state we expand using this basis will have the same symmetry.

### 8.4.3 Example: States for Two Electrons in a Quantum Well

We will now look at what all this means for the states of system of two electrons placed in an infinite quantum well of width $a$. We will neglect the effects of the Coulomb interaction between the electrons. We will also assume that there is no magnetic field (or spin-orbit interactions), so the two spin states of an electron have the same energy.

As a reminder, the single particle spatial states are

$$\phi_n(x) = \begin{cases} \sqrt{\frac{2}{a}}\cos\left(\frac{n\pi x}{a}\right) & (n \text{ odd}) \\ \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi x}{a}\right) & (n \text{ even},) \end{cases} . \tag{8.15}$$

The corresponding single particle energies are $E_n = \hbar^2\pi^2n^2/2ma^2$.

The ground state for two electrons will consist of both in the single particle state with $n = 1$. This means that they have to be in the singlet state, with

$$\begin{aligned} \Psi(x_1,x_2) &= \phi_1(x_1)\phi_1(x_2) \otimes \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle) \\ &= \frac{2}{a}\cos\left(\frac{\pi x_1}{a}\right)\cos\left(\frac{\pi x_2}{a}\right) \otimes \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle) . \end{aligned} \tag{8.16}$$

The energy of the state is $2E_1 = \hbar^2\pi^2/ma^2$. It is not degenerate.

For the first excited state, we leave one electron in the $n = 1$ state, but the other must be promoted to the $n = 2$ state. As the single particle states are different, we can make both triplet and singlet states:

$$\Psi_s(x_1,x_2) = \frac{\sqrt{2}}{a}\left(\cos\left(\frac{\pi x_1}{a}\right)\sin\left(\frac{2\pi x_2}{a}\right) - \cos\left(\frac{\pi x_2}{a}\right)\sin\left(\frac{2\pi x_1}{a}\right)\right) \otimes \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle)\,. \tag{8.17}$$

and

$$\Psi_t(x_1,x_2) = \frac{\sqrt{2}}{a}\left(\cos\left(\frac{\pi x_1}{a}\right)\sin\left(\frac{2\pi x_2}{a}\right) + \cos\left(\frac{\pi x_2}{a}\right)\sin\left(\frac{2\pi x_1}{a}\right)\right) \otimes \begin{cases} |\downarrow\rangle|\downarrow\rangle \\ \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle) \\ |\uparrow\rangle|\uparrow\rangle \end{cases}\,. \tag{8.18}$$

These states have energy $E_1 + E_2 = 5\hbar^2\pi^2/2ma^2$, and they are four-fold degenerate.

## 8.5 Problems

8.1. Determine the symmetry of the following functions under the action of the exchange operator which swaps pairs of labels.

(a) $\psi(x_1,x_2) = 4(x_1-x_2)^2 + \dfrac{1}{x_1^2+x_2^2}$

(b) $\psi(x_1,x_2) = -\dfrac{3(x_1-x_2)}{2(x_1-x_2)^2+7}$

(c) $\psi(x_1,x_2,x_3) = 6x_1x_2x_3 + \dfrac{x_1^2+x_2^2+x_3^2-1}{2x_1^3+2x_2^3+2x_3^3+5}$

(d) $\psi(x_1,x_2) = \dfrac{1}{x_1+3}e^{-|x_2|}$

8.2. Two identical spin-half particles with mass $m$ are confined in a two-dimensional harmonic oscillator potential $\frac{1}{2}m\omega_c^2(x^2+y^2)$. If the total spin state is a triplet,

(a) Find the two lowest energy levels.

(b) What are the degeneracies of these levels?

8.3. Show that $\sigma_x|\uparrow\rangle = |\downarrow\rangle$. Similarly, work out the effects of all the Pauli operators, $\sigma_x$, $\sigma_y$ and $\sigma_z$, on $|\uparrow\rangle$ and $|\downarrow\rangle$.

Taking the total spin operator for two spin-$\frac{1}{2}$ particles, $\hat{\boldsymbol{S}}^2 = (\hat{\boldsymbol{S}}^{(1)} + \hat{\boldsymbol{S}}^{(2)})^2$, express this in terms of the Pauli matrices for the individual particles and show that

$$\hat{\boldsymbol{S}}^2 = \frac{3}{2}\hbar^2\mathbb{1} + \frac{1}{2}\hbar^2(\sigma_x^{(1)}\sigma_x^{(2)} + \sigma_y^{(1)}\sigma_y^{(2)} + \sigma_z^{(1)}\sigma_z^{(2)})$$

Now calculate what $\hat{\boldsymbol{S}}^2$ does to the singlet and triplet spin combinations in Eq.(8.11) and Eq.(8.12). Show that they are eigenstates of $\hat{\boldsymbol{S}}^2$ and find the corresponding eigenvalues.

8.4. Two electrons are placed in a one-dimensional box of length $a$, represented by the infinite potential well

$$V(x) = \begin{cases} 0 & (|x| \le a/2) \\ \infty & (|x| > a/2) \end{cases} .$$

The electrons are both in the same spin state (the combined state is $|\uparrow\rangle|\uparrow\rangle$ or $|\downarrow\rangle|\downarrow\rangle$).

(a) Neglecting the effects of the Coulomb interaction between the electrons, write down the spatial part of the ground state wavefunction $\psi(x_1, x_2)$ of the two particle system.

(b) A measurement is made to determine in which side ($x < 0$ or $x > 0$) of the box each of the electrons is located. Calculate, for the ground state, the probability that both electrons are in the same side of the box.

(c) If the electrons were in fact distinguishable particles, what would be the probability that both were in the same side of the box?

8.5. Two non-interacting spin-$\frac{1}{2}$ particles are moving in a one-dimensional harmonic oscillator potential, for which the eigenstates are $\phi_n(x)$. One particle is in the state $n = 0$, the other, $n = 1$.

(a) Write down appropriate forms for the spatial parts of the wavefunctions, $\Psi(x_1, x_2)$, corresponding to the singlet and triplet spin states. Give your answers in terms of the single particle wavefunctions $\phi_0$ and $\phi_1$.

(b) A weak interaction potential $V(x_1 - x_2)$ is switched on. Give expressions for $\langle\Psi|V(x_1 - x_2)|\Psi\rangle$ for the states from part (a).
This corresponds to an approximation – first order perturbation theory – for the change in energy of the states caused by the interaction.

(c) If $V(x_1 - x_2) = V_0\,\delta(x_1 - x_2)$, evaluate the changes in energy for the singlet and triplet states.

(d) Explain physically your answer for the triplet state.

# 9 Second Quantisation

In the previous section, I introduced the 'saying which state each particle is in' description of many particle states. Now we will meet the 'saying how many particles are in each state' approach, which is known more formally as *second quantisation*.

## 9.1 Creation and Annihilation operators

Consider some single-particle Hermitian operator $\hat{A}$ with eigenvalues $a_j$. On physical grounds, and regardless of distinguishability, we require that $n_j$ particles in the eigenstate $|a_j\rangle$ of $\hat{A}$ must have a value $n_j \times a_j$ for the observable $A$. We can repeat this for all eigenvalues $a_j$, and obtain a potentially infinite set of basis vectors

$$|n_1, n_2, n_3, \ldots\rangle\,, \tag{9.1}$$

for all integer values of $n_j$, including zero. You should convince yourself that this exhausts all the possible ways any number of particles can be distributed over the eigenvalues $a_j$. The eigenvalues of $A$ can be bounded or unbounded, and discrete or continuous. It may even be degenerate. For simplicity we consider here an unbounded, non-degenerate discrete spectrum.

A special state is given by

$$|\varnothing\rangle = |0, 0, 0, \ldots\rangle, \tag{9.2}$$

which indicates the state of no particles, or the *vacuum*.

The numbers $n_j$ are called the occupation number, and any physical state can be written as a superposition of the states in Eq.(9.1)

$$|\Psi\rangle = \sum_{n_1,n_2,n_3,\ldots=0}^{\infty} c_{n_1,n_2,n_3,\ldots} |n_1, n_2, n_3, \ldots\rangle. \tag{9.3}$$

The basis states $|n_1, n_2, n_3, \ldots\rangle$ span a linear vector space called a *Fock space* $\mathscr{F}$. It is the direct sum of the Hilbert spaces for zero particles $\mathscr{H}_0$, one particle $\mathscr{H}_1$, two particles, etc.:

$$\mathscr{F} = \mathscr{H}_0 \oplus \mathscr{H}_1 \oplus \mathscr{H}_2 \oplus \mathscr{H}_3 \oplus \cdots \tag{9.4}$$

Since $|\Psi\rangle$ can now include a superposition over different particle numbers, we require operators that change the particle number. These are the creation and annihilation operators, $\hat{a}^\dagger$ and $\hat{a}$ respectively. Up to a proportionality constant that we will determine later, the action of these operators is defined by

$$\begin{aligned} \hat{a}_j^\dagger |n_1, n_2, \ldots, n_j, \ldots\rangle &\propto |n_1, n_2, \ldots, n_j + 1, \ldots\rangle\,, \\ \hat{a}_j |n_1, n_2, \ldots, n_j, \ldots\rangle &\propto |n_1, n_2, \ldots, n_j - 1, \ldots\rangle\,. \end{aligned} \tag{9.5}$$

So the operator $\hat{a}_j^\dagger$ creates a particle in a state with eigenvalue $a_j$, and the operator $\hat{a}_j$ removes a particle in a state with eigenvalue $a_j$. These operators are each others' Hermitian adjoint, since removing a particle is the time reversal of adding a particle. Clearly, when an annihilation operator attempts to remove particles that are not there, the result must be zero:

$$\hat{a}_j |n_1, n_2, \ldots, n_j = 0, \ldots\rangle = 0\,. \tag{9.6}$$

The vacuum is then defined as the state that gives zero when acted on by any annihilation operator: $\hat{a}_j|\varnothing\rangle = 0$ for any $j$. Notice how we have so far sidestepped the problem of particle swapping; we exclusively used aspects of the *total* particle number.

What are the basic properties of these creation and annihilation operators? In particular, we are interested in their commutation relations. We will now derive these properties from what we have determined so far. First, note that we can create two particles with eigenvalues $a_i$ and $a_j$ in the system in any order, and the only difference this can make is in the normalisation of the state:

$$\hat{a}_i^\dagger \hat{a}_j^\dagger |\Psi\rangle = \lambda \hat{a}_j^\dagger \hat{a}_i^\dagger |\Psi\rangle \,, \tag{9.7}$$

where $\lambda$ is some (maybe complex) number. The state $|\Psi\rangle$ is certainly not zero, we require that

$$\hat{a}_i^\dagger \hat{a}_j^\dagger - \lambda \hat{a}_j^\dagger \hat{a}_j^\dagger = 0 \,. \tag{9.8}$$

Since $k$ and $l$ are just dummy variables, we equally have

$$\hat{a}_j^\dagger \hat{a}_i^\dagger - \lambda \hat{a}_i^\dagger \hat{a}_j^\dagger = 0 \,. \tag{9.9}$$

We now substitute Eq. (9.9) into Eq. (9.8) to eliminate $\hat{a}_j^\dagger \hat{a}_i^\dagger$. This leads to

$$(1-\lambda^2)\, \hat{a}_i^\dagger \hat{a}_j^\dagger = 0 \,, \tag{9.10}$$

and therefore

$$\lambda = \pm 1 \,. \tag{9.11}$$

The relation between different creation operators can thus take two forms. They can obey a commutation relation when $\lambda = +1$:

$$\hat{a}_i^\dagger \hat{a}_j^\dagger - \hat{a}_j^\dagger \hat{a}_i^\dagger = \left[\hat{a}_i^\dagger, \hat{a}_j^\dagger\right] = 0 \,, \tag{9.12}$$

or they can obey an anti-commutation relation when $\lambda = -1$:

$$\hat{a}_i^\dagger \hat{a}_j^\dagger + \hat{a}_j^\dagger \hat{a}_i^\dagger = \left\{\hat{a}_i^\dagger, \hat{a}_j^\dagger\right\} = 0 \,, \tag{9.13}$$

While creating the particles in different temporal order is not the same as swapping two particles, it should not come as a surprise that there are two possible situations (the commutation relation and the anti-commutation relation). We encountered two possibilities in our previous approach as well, where we found that many-particle states are either symmetric or anti-symmetric. In fact, creation operators that obey the commutation relation produce symmetric states, while creation operators that obey the anti-commutation relation produce anti-symmetric states. We also see that the creation operators described by the anti-commutation relations naturally obey Pauli's exclusion principle. Suppose that we wish to create two identical particles in the same eigenstate $|a_j\rangle$. The anti-commutation relations say that $\{\hat{a}_i^\dagger, \hat{a}_i^\dagger\} = 0$, so

$$\hat{a}_i^{\dagger 2} = 0 \,. \tag{9.14}$$

Any higher powers of $\hat{a}_i^\dagger$ will also be zero, and we can create at most one particle in the state $|a_i\rangle$.

Taking the adjoint of the commutation relations for the creation operators gives us the corresponding relations for the annihilation operators

$$\hat{a}_i \hat{a}_j - \hat{a}_j \hat{a}_i = \left[\hat{a}_i, \hat{a}_j\right] = 0 \,, \tag{9.15}$$

or

$$\hat{a}_i \hat{a}_j + \hat{a}_j \hat{a}_i = \left\{\hat{a}_i, \hat{a}_j\right\} = 0 \,. \tag{9.16}$$

The remaining question is now what the (anti-) commutation relations are for products of creation and annihilation operators.

We proceed along similar lines as before. Consider the operators $\hat{a}_i$ and $\hat{a}_j^\dagger$ with $i \neq j$, and apply them in different orders to a state $|\Psi\rangle$.

$$\hat{a}_i \hat{a}_j^\dagger |\Psi\rangle = \mu \hat{a}_j^\dagger \hat{a}_i |\Psi\rangle \,. \tag{9.17}$$

The same argument as before leads to $\mu = \pm 1$. For different $i$ and $j$ we therefore find

$$\left[\hat{a}_i, \hat{a}_j^\dagger\right] = 0 \qquad \text{or} \qquad \left\{\hat{a}_i, \hat{a}_j^\dagger\right\} = 0 \,. \tag{9.18}$$

Now let us consider the case $i = j$. For the special case where $|\Psi\rangle = |\varnothing\rangle$, we find

$$\left(\hat{a}_i \hat{a}_i^\dagger - \mu \hat{a}_i^\dagger \hat{a}_i\right) |\varnothing\rangle = \hat{a}_i \hat{a}_i^\dagger |\varnothing\rangle = |\varnothing\rangle \,, \tag{9.19}$$

based on the property that $\hat{a}_i |\varnothing\rangle = 0$. So we find for the two possible values of $\mu$

$$\hat{a}_i \hat{a}_i^\dagger - \hat{a}_i^\dagger \hat{a}_i = 1 \qquad \text{or} \qquad \hat{a}_i \hat{a}_i^\dagger + \hat{a}_i^\dagger \hat{a}_i = 1 \tag{9.20}$$

which is equivalent to

$$\left[\hat{a}_i, \hat{a}_i^\dagger\right] = 1 \qquad \text{or} \qquad \left\{\hat{a}_i, \hat{a}_i^\dagger\right\} = 1 \,. \tag{9.21}$$

To summarise, we have two sets of algebras for the creation and annihilation operators. The algebra in terms of the commutation relations is given by

$$\left[\hat{a}_i, \hat{a}_j\right] = \left[\hat{a}_i^\dagger, \hat{a}_j^\dagger\right] = 0 \qquad \text{and} \qquad \left[\hat{a}_i, \hat{a}_j^\dagger\right] = \delta_{ij} \,. \tag{9.22}$$

This algebra describes particles that obey Bose-Einstein statistics, or bosons. The algebra in terms of anti-commutation relations is given by

$$\left\{\hat{a}_i, \hat{a}_j\right\} = \left\{\hat{a}_i^\dagger, \hat{a}_j^\dagger\right\} = 0 \qquad \text{and} \qquad \left\{\hat{a}_i, \hat{a}_j^\dagger\right\} = \delta_{ij} \,. \tag{9.23}$$

This algebra describes particles that obey Fermi-Dirac statistics, or fermions.

## 9.2 The Number Operator

We next consider a new observable, the number operator, which gives us the total number of particles in the system. We denote this by $\hat{n}$, and we see that it must be additive over all particle numbers for the different eigenvalues of $A$:

$$\hat{n} = \sum_j \hat{n}_j \,, \tag{9.24}$$

where $\hat{n}_j$ is the number operator for the eigenstate $|a_j\rangle$, that is

$$\hat{n}_j |n_j\rangle = n_j |n_j\rangle. \tag{9.25}$$

For both fermions and bosons, we have

$$\hat{a}_j^\dagger |n_j\rangle = c_j |n_j + 1\rangle, \tag{9.26}$$

where $c_j$ is a constant of proportionality. Thus

$$\hat{a}_j^\dagger \hat{n}_j |n_j\rangle = n_j c_j |n_j + 1\rangle \tag{9.27}$$

while

$$\hat{n}_j \hat{a}_j^\dagger |n_j\rangle = (n_j+1)c_j |n_j+1\rangle. \tag{9.28}$$

From this we see that $\hat{n}_j$ must satisfy

$$[\hat{n}_j, \hat{a}_j^\dagger] = \hat{a}_j^\dagger. \tag{9.29}$$

For both types of particles, this is satisfied if we take

$$\hat{n}_j = \hat{a}_j^\dagger \hat{a}_j. \tag{9.30}$$

We can now deduce the proportionality constants in Eq.(9.26). Taking its adjoint, we have

$$\langle n_j | \hat{a}_j = c_j^* \langle n_j + 1|, \tag{9.31}$$

We thus get

$$\langle n_j | \hat{a}_j \hat{a}_j^\dagger | n_j \rangle = |c_j|^2 \langle n_j+1 | n_j+1\rangle = |c_j|^2. \tag{9.32}$$

Using the commutator/anticommutator, this gives

$$|c_j|^2 = \langle n_j | (\mathbb{1} \pm \hat{n}_j) | n_j \rangle = 1 \pm n_j \tag{9.33}$$

In the case of bosons, we take $c_j = \sqrt{n_j+1}$, giving

$$\hat{a}_j^\dagger |n_j\rangle = \sqrt{n_j+1}|n_j+1\rangle, \tag{9.34}$$

which then requires

$$\hat{a}_j |n_j\rangle = \sqrt{n_j}|n_j-1\rangle. \tag{9.35}$$

These operators satisfy all the commutation relations in Eqs.(9.22).

For fermions, we have to be a bit more careful with the phases of the square roots, as the anticommutation relationships of Eqs.(9.23) are not trivial when $i \neq j$. In order to satsify them, we need to add some phases, so that the sign of a state depends on the order in which the particles are created. The conventional choice is to define

$$\nu_j = \sum_{k<j} n_k. \tag{9.36}$$

Then we take

$$\begin{aligned} \hat{a}_j |0\rangle_j = 0 \qquad &\text{and} \qquad \hat{a}_j^\dagger |0\rangle_j = (-1)^{\nu_j} |1\rangle_j, \\ \hat{a}_j |1\rangle_j = (-1)^{\nu_j} |0\rangle_j \qquad &\text{and} \qquad \hat{a}_j^\dagger |1\rangle_j = 0. \end{aligned} \tag{9.37}$$

To see how this works, we will consider creating two fermions in states $i$ and $j$, with $j > i$. We start with the state $|n_i = 0, n_j = 0\rangle$, and, for convenience, assume that there are no other particles in the system. We will do the process two ways: in the first we act with the creation operaators in the order $\hat{a}_i^\dagger \hat{a}_j^\dagger$, and in the second $\hat{a}_j^\dagger \hat{a}_i^\dagger$. At the start $\nu_i = \nu_j = 0$. Thus

$$\hat{a}_j^\dagger |00\rangle = (-1)^0 |01\rangle = |01\rangle \qquad \hat{a}_i^\dagger |00\rangle = (-1)^0 |10\rangle = |10\rangle. \tag{9.38}$$

Now $\nu_i = 0$ but $\nu_j = 1$ (as $i < j$). So the action of the second operator is different. We have

$$\hat{a}_i^\dagger \hat{a}_j^\dagger |00\rangle = \hat{a}_i^\dagger |01\rangle = (-1)^0 |11\rangle = |11\rangle \qquad \hat{a}_j^\dagger \hat{a}_i^\dagger |00\rangle = \hat{a}_j^\dagger |10\rangle = (-1)^1 |11\rangle = -|11\rangle. \tag{9.39}$$

This gives the correct anticommutator:

$$(\hat{a}_i^\dagger \hat{a}_j^\dagger + \hat{a}_j^\dagger \hat{a}_i^\dagger)|00\rangle = 0. \tag{9.40}$$

## 9.3 Observables based on creation and annihilation operators

So far, we have considered only the basis states of many particles for a single observable $\hat{A}$. What about other observables, in particular those that do not commute with $A$? We can make a similar construction. Suppose an observable $\hat{B}$ has eigenvalues $b_j$. We can construct creation and annihilation operators $\hat{b}_j^\dagger$ and $\hat{b}_j$ that act according to

$$\begin{aligned}\hat{b}_j^\dagger|m_1,m_2,\ldots,m_j,\ldots\rangle &= \sqrt{m_j+1}|m_1,m_2,\ldots,m_j+1,\ldots\rangle\,,\\ \hat{b}_j|m_1,m_2,\ldots,m_j,\ldots\rangle &= \sqrt{m_j}|m_1,m_2,\ldots,m_j-1,\ldots\rangle\,.\end{aligned} \tag{9.41}$$

where $m_j$ is the number of particles with value $b_j$. Typically, the basis states of two observables are related via a unitary transformation

$$|b_i\rangle = \sum_j U_{ji}|a_j\rangle, \tag{9.42}$$

where $U_{ij} = \langle a_i|b_j\rangle$ is a unitary matrix.

How does this relate the creation and annihilation operators? To answer this, let us look at the single particle states. We can write

$$|b_i\rangle = \sum_j U_{ji}|a_j\rangle = \sum_j U_{ji}\hat{a}_j^\dagger|\varnothing\rangle = \hat{b}_i^\dagger|\varnothing\rangle\,. \tag{9.43}$$

This suggests that the creation and annihilation operators are related by

$$\hat{b}_i^\dagger = \sum_j U_{ji}\hat{a}_j^\dagger \qquad \text{and} \qquad \hat{b}_i = \sum_j U_{ji}^*\,\hat{a}_j\,. \tag{9.44}$$

We have only shown that this correspondence is correct when operating on the vacuum state, but it does indeed work for any $N$ particle states.

We can construct any system operators using the creation and annihilation operators. In order to do this, we need to classify operators as one-particle, two-particle etc, as we did in the previous section. For a one particle operator the operator for the system can be written as the sum over operators which act on individual particles:

$$\hat{A} = \sum_\alpha \hat{A}_\alpha. \tag{9.45}$$

Suppose we have used this operator to define our basis $|a_i\rangle$, so

$$\hat{A}|a_i\rangle = a_i|a_i\rangle\,. \tag{9.46}$$

For a system of identical particles, we can see on physical grounds that the effect on the system in state $|\Psi\rangle = |n_1,n_2,\ldots\rangle$ must be

$$\hat{A}|\Psi\rangle = \sum_i n_i a_i|\Psi\rangle\,. \tag{9.47}$$

But $\hat{n}_i|\Psi\rangle = n_i|\Psi\rangle$, so $\hat{A}$ can be written as a sum over states:

$$\hat{A} = \sum_i a_i\hat{n}_i = \sum_i a_i\,\hat{a}_i^\dagger\hat{a}_i\,. \tag{9.48}$$

We also want to be able to express other single particle operators in a similar way. For the operator $\hat{B}$, we would have

$$\hat{B} = \sum_i b_i\,\hat{b}_i^\dagger\hat{b}_i\,. \tag{9.49}$$

Using Eq.(9.44), we get

$$\hat{B} = \sum_{ijk} b_i U_{ji} \hat{a}_j^\dagger U_{ki}^* \hat{a}_k \ = \sum_{jk} B_{jk} \hat{a}_j^\dagger \hat{a}_k \ , \tag{9.50}$$

where

$$\begin{aligned} B_{jk} &= \sum_i b_i U_{ji} U_{ki}^* = \sum_i b_i \langle a_j | b_i \rangle \langle b_i | a_k \rangle = \langle a_j | \left( \sum_i b_i |b_i\rangle\langle b_i| \right) |a_k\rangle \\ &= \langle a_j | \hat{B} | a_k \rangle = \int dx \, \psi_i^*(x) \hat{A} \psi_j(x) . \end{aligned} \tag{9.51}$$

In a similar way, we can write a two particle operator, for example an interaction

$$\hat{V} = \frac{1}{2} \sum_{\alpha\beta} V(x_\alpha, x_\beta) \tag{9.52}$$

in second quantised form as

$$\hat{V} = \frac{1}{2} \sum_{ijkl} V_{ij,kl} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k \tag{9.53}$$

with

$$V_{ij,kl} = \langle a_i, a_j | \hat{V} | a_k, a_l \rangle = \int dx_\alpha \int dx_\beta \, \psi_i^*(x_\alpha) \psi_j^*(x_\beta) V(x_\alpha, x_\beta) \psi_k(x_\alpha) \psi_l(x_\beta) . \tag{9.54}$$

I will not prove this result here.

An important consequence of these results is that we can find any single particle properties of a system by calculating expectation values of $\hat{a}_i^\dagger \hat{a}_j$, and any two particle properties from $\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k$.

## 9.4 Problems

9.1. Calculate the Slater determinant for three electrons and show that no two electrons can be in the same state.

9.2. The number operator

(a) Show that the number operator $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$ satisfies Eq.(9.29), for both fermionic and bosonic operators.

(b) Show that, for both bosons and fermions, the number operators for two different modes commute, $[\hat{n}_i, \hat{n_j}] = 0$ for $i \neq j$.

(c) Show that the number operator takes the same form in any basis, that is

$$\hat{n} = \sum_i \hat{b}_i^\dagger \hat{b}_i = \sum_i \hat{a}_i^\dagger \hat{a}_i$$

when the $a$ and $b$ operators are related by the transformation Eq.(9.44).

9.3. Verify that the operators given by Eqs.(9.34,9.35) and Eq.(9.37) satisfy all the commutation and anticommutation relations for bosonic and fermionic operators.

9.4. Consider a fermionic system with just two energy levels, labelled 1 and 2. Write down, in the occupation number basis, the four states which span the Fock space for this system. Construct the matrices representing the annihilation and creation operators $\hat{a}_1, \hat{a}_2, \hat{a}_1^\dagger$ and $\hat{a}_2^\dagger$. Chose the phases to follow the convention given in Eq.(9.37).

# 10 Entanglement

## 10.1 Composite Systems and Entanglement

Suppose we have two systems, described by Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$, respectively. We can choose orthonormal bases for each system:

$$\mathscr{H}_A : \left\{|\phi_1^A\rangle, |\phi_2^A\rangle, \ldots, |\phi_N^A\rangle\right\} \quad \text{and} \quad \mathscr{H}_B : \left\{|\phi_1^B\rangle, |\phi_2^B\rangle, \ldots, |\phi_M^B\rangle\right\}. \tag{10.1}$$

The respective dimensions of $\mathscr{H}_A$ and $\mathscr{H}_B$ are $N_A$ and $N_B$. We can construct $N_A \times N_B$ basis states for the composite system via $|\phi_j^A\rangle$ and $|\phi_k^B\rangle$. This implies that the total Hilbert space of the composite system can be spanned by the tensor product

$$\left\{|\phi_j^A\rangle \otimes |\phi_k^B\rangle\right\}_{jk} \quad \text{on} \quad \mathscr{H}_{A+B} = \mathscr{H}_A \otimes \mathscr{H}_B. \tag{10.2}$$

An arbitrary pure state on $\mathscr{H}_{A+B}$ can be written as

$$|\Psi\rangle = \sum_{jk} c_{jk} |\phi_j^A\rangle \otimes |\phi_k^B\rangle \equiv \sum_{jk} c_{jk} |\phi_j^A, \phi_k^B\rangle. \tag{10.3}$$

For example, the system of two qubits can be written on the basis $\{|0,0\rangle, |0,1\rangle, |1,0\rangle, |1,1\rangle\}$.

We say $|\psi\rangle$ is a separable, or product state, if there exist vectors $|\psi_A\rangle$ in $\mathscr{H}_A$ and $|\psi_B\rangle$ in $\mathscr{H}_B$ such that $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. If no such decomposition exists, we say the state is entangled. Note that it is not always obviously whether a state is entangled. For instance, consider the state

$$|\Psi\rangle = \frac{1}{2}(|0\rangle_A \otimes |0\rangle_B + |0\rangle_A \otimes |1\rangle_B + |1\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B). \tag{10.4}$$

At first glance it may appear entangled. However, we can factorise it as follows

$$|\Psi\rangle = \frac{1}{2}((|0\rangle_A + |1\rangle_A) \otimes (|0\rangle_B + |1\rangle_B)) \tag{10.5}$$

## 10.2 Measures of Entanglement

Some useful measures of entanglement can be found by taking the *partial trace* over one of the subsystems. We define this using the density operator formalism. Suppose we have an entangled state of two qubits

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B). \tag{10.6}$$

The corresponding density operator for this pure state is

$$\hat{\rho} = |\Psi\rangle\langle\Psi| = \frac{1}{2}(|0\rangle_A\langle 0| \otimes |0\rangle_B\langle 0| + |0\rangle_A\langle 1| \otimes |0\rangle_B\langle 1| + |1\rangle_A\langle 0| \otimes |1\rangle_B\langle 0| + |1\rangle_A\langle 1| \otimes |1\rangle_B\langle 1|). \tag{10.7}$$

Taking the partial trace over the $B$ qubit, we sum over the terms which are diagonal, $|0\rangle_B\langle 0|$ and $|1\rangle_B\langle 1|$, dropping the others, to get

$$\hat{\rho}_A = \mathrm{Tr}_B\{\hat{\rho}\} = \frac{1}{2}(|0\rangle_A\langle 0| + |1\rangle_A\langle 1|). \tag{10.8}$$

This is an operator which acts entirely in $\mathscr{H}_A$ (we could write it as $\hat{\rho}_A \otimes \mathbb{1}$ if we wanted to think of it as an operator in $\mathscr{H}_{A+B}$). Although we started from a pure state in $\mathscr{H}_{A+B}$, $\hat{\rho}_A$ is not a pure state. Though this is obvious in this case, we can check by calculating the purity of the state,

$$P(\hat{\rho}_A) = \mathrm{Tr}\{\hat{\rho}_A^2\} = \mathrm{Tr}\left\{\frac{1}{4}(|0\rangle\langle 0| + |1\rangle\langle 1|)\right\} = \frac{1}{2}. \tag{10.9}$$

The only case where we get a pure state when we trace over one subsystem is if there is no entanglement in the initial state. The purity we calculate does not depend on which subsystem we chose to take the partial trace, so it makes a good measure of entanglement, varying from $\frac{1}{2}$ for a maximally entangled state (of qubits) to 1 for an unentangled, factorisable state.

Another, very commonly used, measure of entanglement is based on entropy. We again take the partial trace over one subsystem, to get $\hat{\rho}_A$. We then calculate the *von Neumann entropy*

$$S(\hat{\rho}_A) = -\mathrm{Tr}\left\{\hat{\rho}_A \log_2(\hat{\rho}_A)\right\} . \tag{10.10}$$

Again this is independent of the subsystem over which we take the partial trace. Of course, it is awkward to take the logarithm of an operator, but the trace is the same in any basis, so we can choose to work in the basis for which $\hat{\rho}_A$ is diagonal. Then

$$\rho_A = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \end{pmatrix} \quad \text{and} \quad \log_2(\rho_A) = \begin{pmatrix} \log_2\lambda_1 & & 0 \\ & \log_2\lambda_2 & \\ 0 & & \ddots \end{pmatrix} . \tag{10.11}$$

Thus

$$S(\hat{\rho}_A) = -\sum_i \lambda_i \log_2\lambda_i = -\sum_i p_i \log_2 p_i \, , \tag{10.12}$$

where the second form comes from identifying the eigenvalues of the density matrix as the probabilities of states in the corresponding mixture, Eq.(7.4). This is just the *Shannon entropy*, from information theory, associated with the randomness in the mixture.

The $\hat{\rho}_A$ in Eq.(10.8) is already in a diagonal form, so we can read off the repeated eigenvalue $\lambda_{1,2} = \frac{1}{2}$. Thus $S(\hat{\rho}_A) = 1$, which is the maximum value the entanglement entropy for a qubit can take (more generally, the maximum value is $\log_2 N$, where $N$ is the dimension of the subsystem Hilbert space). The minimum is zero, which occurs for any factorisable state.

## 10.3 Measurements on Entangled States

Suppose we make an entangled state of two qubits and then separate them and take measurements on each subsytem. Traditionally, Alice takes the measurements on qubit $A$ and Bob measures qubit $B$. We have looked previously (Section 6.3) at how to predict the results of such measurements using the Born rule directly. Here, we will do it slightly differently, using the projector formalism with the vector product notation.

We will do the calculations for the entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle\otimes|1\rangle - |1\rangle\otimes|0\rangle\right) . \tag{10.13}$$

First, let us look at the case where Alice and Bob both measure $\sigma_z$. The eigenstates of $\sigma_z$ are $|0\rangle$, with eigenvalue $+1$, and $|1\rangle$, with eigenvalue $-1$. If Alice and Bob both measure $+1$, the measurement eigenstate is $|0\rangle\otimes|0\rangle$, so the probability of this result is

$$p_{++} = |\langle\Psi|(|0\rangle\otimes|0\rangle)|^2 = \frac{1}{2}|(\langle 0|\otimes\langle 1| - \langle 1|\otimes\langle 0|)(|0\rangle\otimes|0\rangle)|^2 = \frac{1}{2}|\langle 0|0\rangle\langle 1|0\rangle - \langle 1|0\rangle\langle 0|0\rangle|^2 = 0 . \tag{10.14}$$

For the $+-$ result, compressing the notation, we have

$$p_{+-} = |\langle\Psi|01\rangle|^2 = \frac{1}{2}|\langle 01|01\rangle - \langle 10|01\rangle|^2 = \frac{1}{2} . \tag{10.15}$$

In a similar way, we get $p_{--} = 0$ and $p_{-+} = 1/2$.

Now suppose both Alice and Bob measure $\sigma_x$, for which the eigenstates are $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. For the outcome $++$, the eigenstate is $|++\rangle = |+\rangle \otimes |+\rangle$, so

$$p_{++} = |\langle\Psi|++\rangle|^2 = \frac{1}{2}|\langle 01|++\rangle - \langle 10|++\rangle|^2 = \frac{1}{2}|\langle 0|+\rangle\langle 1|+\rangle - \langle 1|+\rangle\langle 0|+\rangle|^2 = 0\,. \tag{10.16}$$

For the $+-$ result, we have

$$p_{+-} = |\langle\Psi|+-\rangle|^2 = \frac{1}{2}|\langle 01|+-\rangle - \langle 10|+-\rangle|^2 = \frac{1}{2}|\langle 0|+\rangle\langle 1|-\rangle - \langle 1|+\rangle\langle 0|-\rangle|^2 == \frac{1}{8}|1\times(-1) - 1\times 1|^2 = \frac{1}{2}\,. \tag{10.17}$$

Again, $p_{--} = 0$ and $p_{-+} = 1/2$.

Whatever direction they choose for the measurement, as long it is the same for each, the results are perfectly anti-correlated: if Alice gets $+1$, Bob gets $-1$, and if Alice gets $-1$, Bob gets $-1$. This is Einstein's 'spooky action at a distance'. Remember, for a single spin, the results are random, but, whatever result Alice gets, Bob's will be the opposite. This is even true if they are far apart, and not causally connected (in the sense of relativistic space time). The experiment has been done (using entangled photons), with careful timing of the measurements, such that it would not be possible for the information about Alice's result to reach Bob, travelling at the speed of light, before his measurement is made. At first sight, this suggests that it may be possible to use such measurements to transmit messages at faster than light speed. However, as Alice's results are random, not under her control, Bob knowing Alice's result does not allow a message to be exchanged.

## 10.4 Measurements on a Single Subsystem

The effect of tracing over one subsystem is to remove all the information about that system from our description of the state. We might wish to do this if we are only interested in the subsystem which remains – if any measurement we do will only involve operators acting on that subsystem. Suppose we have a measurement operator $\hat{M}_A$ which acts only on the $A$ subsystem, so in $\mathscr{H}_{A+B}$ it can be written $\hat{M} = \hat{M}_A \otimes \mathbb{1}$. Then our measurement theory tells us that the expectation value for the measurement is

$$\langle\hat{M}\rangle = \mathrm{Tr}\left\{\hat{\rho}(\hat{M}_A \otimes \mathbb{1})\right\} = \mathrm{Tr}\left\{\hat{\rho}_A \hat{M}_A\right\}\,. \tag{10.18}$$

**Proof**

Using our tensor product basis, we can always expand the density matrix for the entire system, as in Eq.(10.7), in the form

$$\hat{\rho} = \sum_{ijkl} \rho_{ijkl} |\phi_i^A\rangle\langle\phi_j^A| \otimes |\phi_k^B\rangle\langle\phi_l^B| \,. \tag{10.19}$$

Taking the partial trace over subsystem $B$ gives

$$\hat{\rho}_A = \sum_{ijk} \rho_{ijkk} |\phi_i^A\rangle\langle\phi_j^A| \,. \tag{10.20}$$

Considering now the expectation value,

$$\hat{\rho}(\hat{M}_A \otimes \mathbb{1}) = \sum_{ijkl} \rho_{ijkl} (|\phi_i^A\rangle\langle\phi_j^A|\hat{M}_A) \otimes |\phi_k^B\rangle\langle\phi_l^B| \,. \tag{10.21}$$

We can take the trace in two stages: first take the partial trace over qubit $B$, as we did before, then trace over qubit $A$. The first step works exactly the same as without the $\hat{M}$:

$$\mathrm{Tr}_B \left\{\hat{\rho}(\hat{M}_A \otimes \mathbb{1})\right\} = \sum_{ijk} \rho_{ijkk} (|\phi_i^A\rangle\langle\phi_j^A|\hat{M}_A) = \hat{\rho}_A \hat{M}_A \,. \tag{10.22}$$

The second step gives the full trace

$$\langle\hat{M}\rangle = \mathrm{Tr}\left\{\hat{\rho}(\hat{M}_A \otimes \mathbb{1})\right\} = \mathrm{Tr}\left\{\hat{\rho}_A \hat{M}_A\right\} \,. \tag{10.23}$$

We will make use of this in the next section to deal with open systems, where we will treat the system we are interested in as one subsystem, and the environment (rest of the universe) as the other. If we are only interested in measurements on our system, tracing out the environment gives us a density matrix which has all the information we need. Of course, we need to find a way of calculating the time dependence of this density matrix without solving the Schrödinger equation for the entire universe.

## 10.5 Problems

10.1. Consider the entangled state

$$|\Psi\rangle = \frac{1}{5}(3|0\rangle \otimes |0\rangle + 4|1\rangle \otimes |1\rangle) \,.$$

(a) Find the reduced density matrices, $\rho_A$ and $\rho_B$, corresponding to taking the trace over one of the subsystems.

(b) Calculate the entanglement entropies $S(\rho_A)$ and $S(\rho_B)$, showing that they are the same.

(c) If Alice and Bob both measure the quantity $\sigma_x$, calculate $p_{++}$, $p_{--}$, $p_{+-}$ and $p_{-+}$.

10.2. Calculate the entanglement entropy for the state

$$|\Psi\rangle = \frac{1}{\sqrt{3}}(|00\rangle + |10\rangle + |11\rangle) \,.$$

(Here, we use $|00\rangle$ for $|0\rangle \otimes |0\rangle$).

10.3. Consider a state $|\phi\rangle = \cos(\theta)|\Phi^+\rangle + \sin(\theta)|\Phi^-\rangle$ where $|\Phi^\pm\rangle$ are the *Bell states*

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$
$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)\,.$$

(a) Find the density operator $\hat{\rho} = |\phi\rangle\langle\phi|$ then take the partial trace to find the state on system $\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho})$.

(b) Find the operator $\hat{\rho}_A^2$ and calculate the purity of $\hat{\rho}_A$.

(c) For what values of $\theta$ is $|\phi\rangle$ maximally entangled and for what values is it unentangled?

10.4. For the entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)\,,$$

Find $p_{++}$, $p_{--}$, $p_{+-}$ and $p_{-+}$ if Alice and Bob both measure $\sigma_y$.

10.5. For a measurement at an arbitrary angle, $\theta$, in the $x-y$ plane, the measurement operator is

$$\sigma_\theta = \cos\theta\,\sigma_x + \sin\theta\,\sigma_y\,.$$

(a) Find the eigenvectors of this operator, corresponding to the eigenvalues $\pm 1$.

(b) Suppose we have the entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\,.$$

Alice and Bob carry out measurements at angles $\theta_A$ and $\theta_B$. Show that correlation

$$p_{++} = \frac{1}{2}\sin^2\left(\frac{\theta_A - \theta_B}{2}\right).$$

# 11 Open Quantum Systems

Much of quantum mechanics is concerned with closed systems, where we can (effectively) treat the system as isolated from its environment. Of course, this is an approximation, but in many cases it is a very good one. In this section, we will look at the how to find the time evolution of of a system which is *open*, meaning that we are interested in the effects of interactions between system and its environment. The strategy we will follow is to consider a larger, *composite* system, consisting of the system and environment. We will then take the partial trace over the environment, to get a density matrix describing just the system. If we only measure the system, this contains all the information needed to predict the results of any experiment.

As we have learned, if the interactions between the system and environment create entanglement, our density matrix will always evolve into one describing a mixed state. Thus, when a system prepared in a pure state interacts with its environment, it will generally end up in a mixed state, a process we call *decoherence*.

## 11.1 A Simple Example of Open Systems Dynamics

Consider a Hilbert space $\mathscr{H}_S \otimes \mathscr{H}_E$ where $\mathscr{H}_S$ is the system of interest and $\mathscr{H}_E$ is the environment. For this example, we will take both our system and 'environment' to be single qubits. We will start with the environment qubit in the state $|0\rangle$ and assume that, at $t = 0$, the system and environment are uncorrelated, so

$$\hat{\rho}(t=0) = \hat{\rho}_0 \otimes |0\rangle\langle 0|. \tag{11.1}$$

Clearly $\mathrm{Tr}_E\left\{\hat{\rho}(t=0)\right\} = \hat{\rho}_0$.

Now consider evolution under the Hamiltonian $\hat{H} = -\hbar\sigma_x^S \otimes \sigma_x^E$ where $\sigma_x$ is the Pauli operator. After a time $t$ the joint state evolved under the unitary

$$\begin{aligned}\hat{U} &= \exp(it\,\sigma_x^S \otimes \sigma_x^E) \\ &= \cos(t)\mathbb{1} + i\sin(t)\sigma_x^S \otimes \sigma_x^E\,.\end{aligned} \tag{11.2}$$

Then, recalling that $\sigma_x|0\rangle = |1\rangle$ etc,

$$\begin{aligned}\hat{\rho}(t) =& \hat{U}\hat{\rho}(0)\hat{U}^\dagger \\ =& \cos(t)^2\hat{\rho}_0 \otimes |0\rangle\langle 0| + i\cos(t)\sin(t)(\sigma_x^S\hat{\rho}_0 \otimes |1\rangle\langle 0| - \hat{\rho}_0\sigma_x^S \otimes |0\rangle\langle 1|) \\ &+ \sin(t)^2\sigma_x^S\hat{\rho}_0\sigma_x^S \otimes |1\rangle\langle 1|\,.\end{aligned} \tag{11.3}$$

Taking the partial trace over the environment, we have

$$\hat{\rho}_S(t) = \mathrm{Tr}_E\left\{\hat{\rho}(t)\right\} = \cos(t)^2\hat{\rho}_0 + \sin(t)^2\sigma_x\hat{\rho}_0\sigma_x\,. \tag{11.4}$$

Consider the case when $\hat{\rho}_0 = |0\rangle\langle 0|$, we now have

$$\hat{\rho}_S(t) = \cos(t)^2|0\rangle\langle 0| + \sin(t)^2|1\rangle\langle 1|. \tag{11.5}$$

At time $t = \pi/4$, the state is maximally mixed. All information about the state has leaked into the environment and is lost. However, at $t = \pi$, we have $\hat{\rho}_S(t) = |0\rangle\langle 0|$. The system has returned to its original state, so the information was not permanently lost! We say such an evolution is highly *non-Markovian*. The system only returns to its original state because of the very simple choice of environment, with only one degree of freedom. For sufficiently complex environments, information will be irreversibly lost to the environment.

Of course, what we really want to do is not calculate the full time evolution of the system and its environment, but come up with a treatment for the time evolution of the system, involving only the system variables. There are two ways to proceed. One, which we will not follow, is to create more complicated model environments (though obviously not as complicated as the universe!) and interactions, and try to find approximate equations for the density operator describing the system. Instead, we will figure out the most general equations for the evolution of a system subject to Markovian evolution, consistent with the rules of quantum mechanics. For a particular system and environment interaction, we then tailor the terms containing the environment interactions to the physics we need.

## 11.2 Quantum operations

The central idea of our treatment is that of a quantum *operation* or *channel*. We will use $\mathscr{E}$ to describe quantum operations (which are also called channels or super operators). An operation converts an operator (e.g. a density operator) into a new operator. If an operation $\mathscr{E}$ acts on a state $\hat{\rho}$, we denote the output as $\mathscr{E}(\hat{\rho})$. As quantum mechanics is a linear theory, operations must act linearly, so $\mathscr{E}(\hat{\rho}_1+\hat{\rho}_2)=\mathscr{E}(\hat{\rho}_1)+\mathscr{E}(\hat{\rho}_2)$ etc.

Operations are a general tool for describing time evolution. We will use $\mathscr{E}_t$ to denote the operation resulting from evolution for a time $t$, so that

$$\hat{\rho}_t=\mathscr{E}_t(\hat{\rho}_{t=0}). \tag{11.6}$$

A Markovian process is one where the environment has no memory, which means that the corresponding operations must satisfy

$$\mathscr{E}_{t+\tau}=\mathscr{E}_t\mathscr{E}_\tau \quad \text{that is} \quad \mathscr{E}_{t+\tau}(\hat{\rho})=\mathscr{E}_t(\mathscr{E}_\tau(\hat{\rho}))\,. \tag{11.7}$$

For unitary evolutions on closed systems have

$$\mathscr{E}_t(\hat{\rho})=\exp(-it\hat{H}/\hbar)\hat{\rho}\exp(it\hat{H}/\hbar), \tag{11.8}$$

so

$$\begin{aligned}\mathscr{E}_t\mathscr{E}_\tau(\hat{\rho}) &= \exp(-it\hat{H}/\hbar)\exp(-i\tau\hat{H}/\hbar)\hat{\rho}\exp(i\tau\hat{H}/\hbar)\exp(it\hat{H}/\hbar), \\ &= \exp(-i(t+\tau)\hat{H}/\hbar)\hat{\rho}\exp(i(t+\tau)\hat{H}/\hbar), \\ &= \mathscr{E}_{t+\tau}(\hat{\rho})\,.\end{aligned} \tag{11.9}$$

Therefore, all unitary evolutions on closed systems are Markovian.

However,unitary evolution in an open system is not necessarily lead to Markovian dynamics. Recall the dynamics given by Eq. 11.5, which in operation notation would be

$$\mathscr{E}_t(\hat{\rho})=\cos^2(t)\hat{\rho}+\sin^2(t)\sigma_x\hat{\rho}\sigma_x\,. \tag{11.10}$$

You should verify that this is not Markovian, for instance by confirming that $\mathscr{E}_{\pi/2}\mathscr{E}_{\pi/2}\neq\mathscr{E}_\pi$.

Nevertheless, many non-unitary decoherence processes are Markovian. Consider the so-called dephasing channel

$$\mathscr{E}_t(\hat{\rho})=\left(\frac{1+e^{-t}}{2}\right)\hat{\rho}+\left(\frac{1-e^{-t}}{2}\right)\sigma_x\hat{\rho}\sigma_x\,. \tag{11.11}$$

You will verify that this is Markovian in the problems.

## 11.3 The Kraus representation

For a quantum operation to represent a physical evolution, it must output a valid density matrix whenever we input a valid density matrix. Therefore, it must satisfy the following:

**1.** If $\hat{\rho} = \hat{\rho}^\dagger$ then $\mathscr{E}(\hat{\rho}) = \mathscr{E}(\hat{\rho})^\dagger$; (Hermiticity)

**2.** $\mathrm{Tr}\{\hat{\rho}\} = \mathrm{Tr}\{\mathscr{E}(\hat{\rho})\}$ for all $\hat{\rho}$; (trace preserving)

**3.** if $\hat{\rho} \geq 0$ then $\mathscr{E}(\hat{\rho}) \geq 0$; (positivity preserving)

In fact positivity is too weak a condition, because it would allow non-physical operations such as taking the transpose of $\hat{\rho}$. We actually need a stronger restriction, called *complete positivity* to ensure our operation is fully consistent with quantum mechanics. We will not go into this here.

A powerful tool in the treatment of quantum channels is the Kraus decomposition. Let us assume that it is possible to find operators $\hat{A}_k$, such that the channel can decomposed as

$$\mathscr{E}(\hat{\rho}) = \sum_k \hat{A}_k \hat{\rho} \hat{A}_k^\dagger. \tag{11.12}$$

Notice that Eq.(11.4) has this form, with $\hat{A}_1 = \cos(t)\mathbb{1}$ and $\hat{A}_2 = \sin(t)\sigma_x$.

Let us now consider our three properties. Maintaining Hermiticity is trivial:

$$\mathscr{E}(\hat{\rho})^\dagger = \sum_k (\hat{A}_k \hat{\rho} \hat{A}_k^\dagger)^\dagger = \sum_k \hat{A}_k \hat{\rho}^\dagger \hat{A}_k^\dagger \,, \tag{11.13}$$

so if $\hat{\rho}$ is hermitian, so is $\mathscr{E}(\hat{\rho})$.

For trace preservation, we require

$$\mathrm{Tr}\{\mathscr{E}(\hat{\rho})\} = \mathrm{Tr}\left\{\sum_k \hat{A}_k \hat{\rho} \hat{A}_k^\dagger\right\} = \mathrm{Tr}\left\{\sum_k \hat{A}_k^\dagger \hat{A}_k \hat{\rho}\right\}. \tag{11.14}$$

Hence trace preservation requires the Krauss operators to satisfy

$$\sum_k \hat{A}_k^\dagger \hat{A}_k = \mathbb{1}. \tag{11.15}$$

Next, recall that positivity means that $\langle\psi|\hat{\rho}|\psi\rangle \geq 0$ for all $|\psi\rangle$. We have

$$\langle\psi|\mathscr{E}(\hat{\rho})|\psi\rangle = \sum_k \langle\psi|\hat{A}_k \hat{\rho} \hat{A}_k^\dagger|\psi\rangle \,. \tag{11.16}$$

Defining $|\varphi_k\rangle = A_k^\dagger|\psi\rangle$, so that

$$\langle\psi|\mathscr{E}(\hat{\rho})|\psi\rangle = \sum_k \langle\varphi_k|\hat{\rho}|\varphi_k\rangle. \tag{11.17}$$

Hence if $\hat{\rho} \geq 0$, so $\langle\varphi_k|\rho|\varphi_k\rangle \geq 0$ for all $|\varphi_k\rangle$, then so is $\mathscr{E}(\hat{\rho})$.

Therefore, any channel with a Kraus representation (with $\sum_k \hat{A}_k^\dagger \hat{A}_k = \mathbb{1}$) will be Hermiticity, trace and positivity preserving. In fact, this is enough to guarantee complete positivity too.

## 11.4 Markovian Processes and the Lindblad equation

We return now to study of Markovian processes; those physical channels where the environment has no memory and so $\mathscr{E}_t\mathscr{E}_\tau = \mathscr{E}_{t+\tau}$. If $\mathscr{E}_t$ is continuous in time, then we can take derivatives and write

$$\frac{d\hat{\rho}_t}{dt} = \mathscr{L}(\hat{\rho}_t), \tag{11.18}$$

where we have introduced an new superoperator $\mathscr{L}$ to describe the time-derivative. We have assumed Markovianity in the above equation by taking $\mathscr{L}$ to be time-independent. This way the change in the state $d\hat{\rho}_t$ depends only on the physics of the process and the state of the system at time $t$.

This simple differential equation has the formal solution $\hat{\rho}_t = \exp(\mathscr{L}t)(\hat{\rho}_0)$ where the exponential of an operation is defined using a Taylor series. Expressed in terms of operations, we have shown that $\mathscr{E}_t = \exp(\mathscr{L}t)$.

We have already encountered one important class of $\mathscr{L}$ that generate physical operations. For any Hamiltonian $H$ we can define $\mathscr{L}_H$ by

$$\mathscr{L}_H(\hat{\rho}) = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}]. \tag{11.19}$$

then Eq.(11.18) reduces to the Von Neumann equation, Eq.(7.8).

More generally, we have the Lindblad equation which is the direct extension of the von Neumann for open systems. We want to describe an infinitesimal evolution of $\hat{\rho}$, in order to give the continuum evolution later on. We therefore consider a channel with Krauss decomposition

$$\hat{\rho}' = \hat{\rho} + \delta\hat{\rho} = \sum_k \hat{A}_k \hat{\rho} \hat{A}_k^\dagger . \tag{11.20}$$

Since $\delta\hat{\rho}$ is very small, one of the Kraus operators must be close to the identity, not proportional to $\delta t$. Without loss of generality we choose this to be $\hat{A}_0$, and then we can write

$$\hat{A}_0 = \mathbb{1} + (\hat{L}_0 - i\hat{K})\delta t \quad \text{and} \quad \hat{A}_k = \hat{L}_k\sqrt{\delta t}, \tag{11.21}$$

where we introduced the Hermitian operators $\hat{L}_0$ and $\hat{K}$, and the remaining $L_k$ are not necessarily Hermitian. We could have written $\hat{A}_0 = \mathbb{1} + \hat{L}_0\delta t$ and keep $\hat{L}_0$ general (non-Hermitian as well), but it will be useful later on to explicitly decompose it into Hermitian parts. We can now write

$$\begin{aligned} \hat{A}_0\hat{\rho}\hat{A}_0^\dagger &= \hat{\rho} + \left[(\hat{L}_0 - i\hat{K})\hat{\rho} + \hat{\rho}(\hat{L}_0 + i\hat{K})\right]\delta t + O(\delta t^2) \\ \hat{A}_k\hat{\rho}\hat{A}_k^\dagger &= \hat{L}_k\hat{\rho}\hat{L}_k^\dagger\delta t . \end{aligned} \tag{11.22}$$

We can substitute this into Eq.(11.20), to obtain up to first order in $\delta t$

$$\delta\hat{\rho} = \left[\left(\hat{L}_0\hat{\rho} + \hat{\rho}\hat{L}_0\right) - i(\hat{K}\hat{\rho} - \hat{\rho}\hat{K}) + \sum_{k\neq 0}\hat{L}_k\hat{\rho}\hat{L}_k^\dagger\right]\delta t . \tag{11.23}$$

We now obtain the differential equation for the evolution by by dividing by $\delta t$ and taking the limit $\delta t \to 0$:

$$\frac{d\hat{\rho}}{dt} = -i[\hat{K}, \hat{\rho}] + \{\hat{L}_0, \hat{\rho}\} + \sum_{k\neq 0}\hat{L}_k\hat{\rho}\hat{L}_k^\dagger , \tag{11.24}$$

where $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$ is the anti-commutator of $\hat{A}$ and $\hat{B}$.

For a closed system, Eq.(11.24) must then reduce to the von Neumann equation for the density operator $\hat{\rho}$ in Eq.(7.8), and we see that all $\hat{L}_k$ including $\hat{L}_0$ are zero, and $\hat{K}$ is proportional to the Hamiltonian $\hat{K} = \hat{H}/\hbar$.

We have not yet imposed the trace conserving condition, Eq.(11.15). Applying this to $\hat{A}_0$ and $\hat{A}_k$, we get

$$(\mathbb{1} + (\hat{L}_0 + i\hat{K})\delta t)(\mathbb{1} + (\hat{L}_0 - i\hat{K})\delta t) + \sum_{k\neq 0} \hat{L}_k^\dagger \hat{L}_k \delta t = \mathbb{1}\,, \tag{11.25}$$

so, correct to $O(\delta t)$,

$$\hat{L}_0 = -\frac{1}{2}\sum_{k\neq 0} \hat{L}_k^\dagger \hat{L}_k\,. \tag{11.26}$$

Making these substitutions, we finally get

$$\frac{d\hat{\rho}}{dt} = \mathscr{L}(\hat{\rho}) = \frac{1}{i\hbar}[H,\hat{\rho}] + \frac{1}{2}\sum_k \left(2\hat{L}_k\hat{\rho}\hat{L}_k^\dagger - \{\hat{L}_k^\dagger\hat{L}_k,\hat{\rho}\}\right). \tag{11.27}$$

The $k$ sum here excludes $k = 0$: we do not think of $\hat{L}_0$ as one of the Lindblad operators, because it is written in terms of all the others.

The operators $\hat{L}_k$ are chosen such that they model the relevant physical processes. This may sound vague, but in practice it will be quite clear. For example, modelling a transition $|1\rangle \to |0\rangle$ due to interactions with the environment will require a single Lindblad operator

$$\hat{L}_1 = \gamma|0\rangle\langle 1|\,, \tag{11.28}$$

where $\gamma$ is a real parameter indicating the strength of the transition.

## 11.5 Solving the Lindblad Equation

In this section we will consider the time evolution of a qubit, initially in the pure state $|0\rangle$, with $\hat{H} = 0$ but with environment interactions given by the Lindblad operators $\hat{L}_1 = \sqrt{\gamma}|0\rangle\langle 1|$ and $\hat{L}_2 = \sqrt{\gamma}|1\rangle\langle 0|$ (with $\gamma$ real).

We first calculate the various operator combinations we will need:

$$\begin{aligned}
\hat{L}_1^\dagger\hat{L}_1 &= \gamma|1\rangle\langle 0|0\rangle\langle 1| = \gamma|1\rangle\langle 1| \\
\hat{L}_2^\dagger\hat{L}_2 &= \gamma|0\rangle\langle 0| \\
\hat{L}_1^\dagger\hat{L}_1 + \hat{L}_1^\dagger\hat{L}_1 &= \gamma(|1\rangle\langle 1| + |0\rangle\langle 0|) = \gamma\mathbb{1}\,.
\end{aligned} \tag{11.29}$$

Hence the Lindblad equation becomes (remembering $\hat{H} = 0$)

$$\begin{aligned}
\frac{d\hat{\rho}}{dt} &= \frac{1}{2}\left(2\hat{L}_1\hat{\rho}\hat{L}_1^\dagger + 2\hat{L}_2\hat{\rho}\hat{L}_2^\dagger - \{(\hat{L}_1^\dagger\hat{L}_1 + \hat{L}_2^\dagger\hat{L}_2),\hat{\rho}\}\right). \\
&= \gamma\left(|0\rangle\langle 1|\hat{\rho}|1\rangle\langle 0| + |1\rangle\langle 0|\hat{\rho}|0\rangle\langle 1| - \hat{\rho}\right).
\end{aligned} \tag{11.30}$$

To solve this, let us use the property that the Pauli matrices plus $\mathbb{1}$ form a complete basis for any $2\times 2$ Hermitian operator. We can thus write

$$\hat{\rho}(t) = \frac{1}{2}\left(\mathbb{1} + a_x(t)\sigma_x + a_y(t)\sigma_y + a_z(t)\sigma_z\right). \tag{11.31}$$

which correctly has $\mathrm{Tr}\{\hat{\rho}\} = 1$ (recall $\mathrm{Tr}\{\sigma_i\} = 0$). With this parameterisation, we get

$$\begin{aligned}\langle 1|\hat{\rho}|1\rangle &= \frac{1}{2}(1-a_z)\\ \langle 0|\hat{\rho}|0\rangle &= \frac{1}{2}(1+a_z)\,,\end{aligned} \tag{11.32}$$

so

$$|0\rangle\langle 1|\hat{\rho}|1\rangle\langle 0| + |1\rangle\langle 0|\hat{\rho}|0\rangle\langle 1| = \frac{1}{2}(1-a_z)|0\rangle\langle 0| + \frac{1}{2}(1+a_z)|1\rangle\langle 1| = \frac{1}{2}(\mathbb{1} - a_z\sigma_z)\,. \tag{11.33}$$

Then, the Lindblad equation becomes

$$\frac{1}{2}\left(\frac{da_x}{dt}\sigma_x + \frac{da_y}{dt}\sigma_y + \frac{da_z}{dt}\sigma_z\right) = \frac{1}{2}\gamma\left[(\mathbb{1} - a_z\sigma_z) - \mathbb{1} - a_x\sigma_x - a_y\sigma_y - a_z\sigma_z\right]\,. \tag{11.34}$$

Equating coefficients of the Pauli matrices

$$\frac{da_x}{dt} = -\gamma a_x\,, \quad \frac{da_y}{dt} = -\gamma a_y\,, \quad \frac{da_z}{dt} = -2\gamma a_z\,. \tag{11.35}$$

Solving,

$$a_x = a_x^0 e^{-\gamma t}\,, \quad a_y = a_y^0 e^{-\gamma t} \quad a_z = a_z^0 e^{-2\gamma t}\,. \tag{11.36}$$

For our initial state, $|0\rangle$,

$$\hat{\rho}(0) = |0\rangle\langle 0| = \frac{1}{2}(\mathbb{1} + \sigma_z)\,, \tag{11.37}$$

so $a_x^0 = a_y^0 = 0$ and $a_z^0 = 1$. Thus

$$\hat{\rho}(t) = \frac{1}{2}(\mathbb{1} + e^{-2\gamma t}\sigma_z) = \frac{1}{2}(1+e^{-2\gamma t})|0\rangle\langle 0| + \frac{1}{2}(1-e^{-2\gamma t})|1\rangle\langle 1|\,. \tag{11.38}$$

This starts off as a pure state, $|0\rangle\langle 0|$, but ends up in the maximally mixed state $\frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$. In fact, whatever state we start off in, we will always end up with this mixed state – the decoherence process removes all information about the initial state.

## 11.6 Problems

11.1. Consider the dephasing operation

$$\mathscr{E}_t(\hat{\rho}) = \left(\frac{1+e^{-t}}{2}\right)\hat{\rho} + \left(\frac{1-e^{-t}}{2}\right)\sigma_z\hat{\rho}\sigma_z$$

(a) Show that the dephasing operation is Markovian;

(b) Show that the dephasing operation is trace preserving;

(c) For $\hat{\rho}_{t=0} = |+\rangle\langle +|$, where $|+\rangle$ is the $\sigma_x$ eigenstate $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, find $\hat{\rho}_t = \mathscr{E}_t(\hat{\rho}_{t=0})$;

11.2. Consider an amplitude damping operation on a 2-dimension Hilbert space

$$\mathscr{E}_{\gamma,\beta}(\hat{\rho}) = \hat{K}_1\hat{\rho}\hat{K}_1^\dagger + \hat{K}_2\hat{\rho}\hat{K}_2^\dagger$$

with Kraus operators of the form

$$\begin{aligned}\hat{K}_1 &= \gamma|0\rangle\langle 1|,\\ \hat{K}_2 &= |0\rangle\langle 0| + \beta|1\rangle\langle 1|.\end{aligned}$$

(a) For what values of $\gamma$ and $\beta$ is the operation trace preserving?

(b) Find the effect of this operation on the initial state $\hat{\rho} = |+\rangle\langle +|$, where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

11.3. A single spin interacts with an environment. The Hamiltonian for the spin is $H = \hbar\omega|\uparrow\rangle\langle\uparrow|$ and there are two Linblad operators, $L_+ = \sqrt{f_+}\sigma^+$ and $L_- = \sqrt{f_-}\sigma^-$, where $\sigma^\pm$ are the spin raising and lowering operators $\frac{1}{2}(\sigma_x \pm i\sigma_y)$.

Obtain the Linblad equation for the system and determine its steady state solution, that is, the $\hat{\rho}$ for which there is no time evolution.

If the environment consists of a thermal bath at temperature $T$, what can you say about $f_+$ and $f_-$?

11.4. Solve the Lindblad equation with $\hat{H} = \hbar\omega\sigma_x$ and a single Lindblad operator $\hat{L} = \sqrt{\frac{\gamma}{2}}\sigma_z$. The system is initially in the pure state $|0\rangle$.

11.5. A harmonic oscillator has Hamiltonian $\hat{H} = (\hat{a}^\dagger\hat{a} + \frac{1}{2})\hbar\omega$, where $\hat{a}^\dagger$ and $\hat{a}$ are the standard raising and lowering operators and $\omega$ is the oscillation frequency. The coupling to its environment is described by Lindblad operators $\hat{L}_1 = \sqrt{2\gamma}\hat{a}$ and $\hat{L}_2 = \sqrt{2f\gamma}\hat{a}^\dagger$.

(a) Write down the Lindblad equation for the density matrix, $\rho$, representing the state of the oscillator.

(b) Defining the occupation probability $p_n = \langle n|\hat{\rho}|n\rangle$, show that

$$\frac{dp_n}{dt} = 2\gamma(n+1)p_{n+1} + 2\gamma f n p_{n-1} - 2\gamma(n + f(n+1))p_n.$$

(c) Show that in the steady state $p_n = \alpha\nu^n$, and find values for the constants $\alpha$ and $\nu$ in terms of $f$ and $\gamma$.

(d) Determine the mean occupation number of the oscillator $\langle n\rangle$.