

# *Advanced Quantum Theory*

**Joan Vaccaro**

**Howard Wiseman**

## Preamble

Quantum mechanics is a modern scientific theory that describes the motion of objects and all of the interactions between them, except for gravitational ones. It was discovered over a 25-year-long period between 1900 and 1925 during which a large number of (mostly) famous physicists made significant contributions. These so-called 'founding fathers' included Albert Einstein, Erwin Schrödinger, Niels Bohr, Werner Heisenberg, Wolfgang Pauli and Max Born.

Arguably, it is one of humanity's greatest intellectual achievements. Certainly, it is our best current theory of the nature of the physical world. It is one of the most widely used theories in physics today and has applications in areas such as solid-state physics, astronomy, optics, medical physics and thermodynamics.

It is a very **diverse** theory and, for example, describes things ranging from processes in the interiors of stars to the constituents of atomic nuclei. It is also highly **accurate**. No accepted experiment to date disagrees with it and, in addition, some of its predictions have been verified to up to about one part in  $10^{11}$  (eg. its prediction for the magnetic moment of the electron).

It is used in industry in the design of high-technology devices such as lasers, computer chips and light sensors. Nobel laureate and physicist Leon Lederman has estimated that quantum mechanics underpins a third of the gross domestic product of the United States of America. This amounts to 4.5 trillion dollars Australian (AUD 4,500,000,000)!

Quantum mechanics tells us that the world we inhabit is radically different from what everyday experience suggests. For instance, it suggests that, at least at the microscopic level, there is genuine randomness. That is, that certain events, such as whether a given electron is measured to be spin up or down, are sometimes uncaused and so occur for no reason whatsoever. It also suggests that, in certain special circumstances, two distant microscopic objects can be instantaneously connected to each other (entanglement). We shall explore these points during the course.

Finally, it is just beginning to be used to create futuristic new technologies. These so-called **quantum technologies** include quantum computers, quantum teleporters and quantum cryptography. We shall discuss some of them in more detail later.

# Contents

1. **Dirac's bra-ket notation**  
kets, bras, operators, inner products, Hilbert spaces
2. **Quantum states**  
the position and momentum representations  
the density operator and mixed states
3. **Measurement theory**  
von Neumann projections, general measurement, POMs
4. **Quantum entanglement**  
composite systems, correlations, entanglement, Schrödinger's cat and many worlds
5. **Nonlocality and Bell's theorem**  
reality and locality in quantum interpretations, Bell's inequality, Bell's theorem, what does this violation mean?
6. **Angular momentum operators**  
commutation relations, eigenstates
7. **Addition of angular momentum**  
Clebsch-Gordan coefficients, Wigner 3-j symbols
8. **Irreducible tensor operators**  
The Wigner-Eckart theorem
9. **Time evolution**  
the time-evolution operator, "Pictures" of time evolution, Schrödinger picture, Heisenberg picture, interaction frame
10. **Dynamical perturbation theory**  
why use perturbation theory?, 1st order time-dependent perturbation theory, Higher-order time-dependent perturbation theory

# 1. Dirac's bra-ket notation

Paul Dirac introduced the names ‘bra’ and ‘ket’ in the 1920’s from parts of the word ‘bracket’. An expectation value is given by the ‘bracket’  $\langle \dots \rangle$  which is constructed from a ‘bra’  $\langle |$  and a ‘ket’  $| \rangle$ .

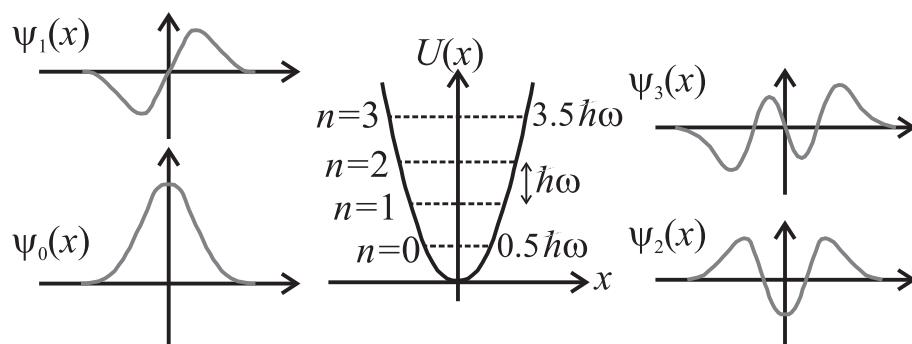
## a. Relationship between kets and wavefunctions

### Harmonic oscillator example

The harmonic oscillator is rather like *the particle in a one dimensional box* that you discovered in First Year Physics. The main difference is that, whereas the walls of the box have infinite potential energy, the potential energy of the harmonic oscillator rises more slowly as illustrated in the figure below. In fact the potential energy is just like that of a classical harmonic oscillator  $U(x) = \frac{1}{2}kx^2$  where  $k$  is the spring constant and  $x$  is the position. It turns out that the energy spectrum of the harmonic oscillator is discrete with values:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega$$

where  $\omega$  is the frequency of oscillation. A few of the energy eigenfunctions  $\psi_n(x)$  are illustrated in the figure.



**Fig 1.1:** The energy eigenfunctions and the spectrum of a simple harmonic oscillator.

The energy eigenfunctions are **orthonormal** (orthogonal and normalised) in the sense that

$$\int \psi_n^*(x)\psi_m(x)dx = \delta_{n,m} . \quad (1.1)$$

They form a **complete orthonormal basis** and so an arbitrary wave function  $\Psi(x)$  can be expanded as a series involving the eigenfunctions  $\psi_m(x)$  as follows:

$$\Psi(x) = \sum_{n=0}^{\infty} c_n \psi_n(x) \quad (1.2)$$

for **coefficients**<sup>†</sup>  $c_n$ . The coefficients are found by the inner product

$$c_n = \int \psi_n^*(x)\Psi(x)dx . \quad (1.3)$$

### Problem

Using the expansion Eq. (1.2) for  $\Psi(x)$  and the orthonormality relation Eq. (1.1) check that the right hand side of Eq. (1.3) does indeed equal  $c_n$ .

This means the expansion Eq. (1.2) is actually given by

### Wave function expansion

$$\Psi(x) = \sum_n \underbrace{\left[ \int \psi_n^*(x')\Psi(x')dx' \right]}_{\text{inner product}} \psi_n(x) \quad (1.4)$$

### Comment

We could write the last result as

$$\Psi(x) = \int \mathbf{1}(x, x')\Psi(x')dx'$$

where

$$\mathbf{1}(x, x') = \sum_n \psi_n^*(x')\psi_n(x) \quad (1.5)$$

is a function that acts like the identity, i.e. it doesn't change the wave function when integrated over. We will return to this point later.

Consider another arbitrary wave function  $\Phi(x)$  with the expansion

$$\Phi(x) = \sum_{n=0}^{\infty} d_n \psi_n(x) .$$

---

<sup>†</sup>The  $c_n$  are also called the **probability amplitudes** and  $|c_n|^2$  is the **probability** of finding the oscillator in the energy eigenfunction  $\psi_n(x)$ , or equivalently, to find it has energy  $(n + \frac{1}{2})\hbar\omega$ .

The inner product between  $\Phi(x)$  and  $\Psi(x)$  is given by

$$\begin{aligned}
\int \Phi^*(x)\Psi(x)dx &= \int \left[ \sum_{n=0}^{\infty} d_n^* \psi_n^*(x) \right] \left( \sum_{m=0}^{\infty} c_m \psi_m(x) \right) dx \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} d_n^* c_m \left[ \int \psi_n^*(x) \psi_m(x) dx \right] \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} d_n^* c_m (\delta_{n,m}) \\
&= \sum_{n=0}^{\infty} d_n^* c_n .
\end{aligned} \tag{1.6}$$

We will return to this result in a moment.

Let  $\Psi(x)$  be normalised (i.e. have a unit length) then

$$1 = \int \Psi^*(x)\Psi(x)dx .$$

Using the expansion Eq. (1.2) on the right hand side gives

$$\begin{aligned}
1 &= \int \left[ \sum_{n=0}^{\infty} c_n^* \psi_n^*(x) \right] \left[ \sum_{m=0}^{\infty} c_m \psi_m(x) \right] dx \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m \left[ \int \psi_n^*(x) \psi_m(x) dx \right] \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m (\delta_{n,m}) = \sum_{n=0}^{\infty} c_n^* c_n \\
&= \sum_{n=0}^{\infty} |c_n|^2 .
\end{aligned}$$

Thus, the probability distribution  $P_n = |c_n|^2$ , which describes the probability of finding the oscillator in energy eigenfunction  $\psi_n(x)$ , is normalised to unity (as it should be).

### Mapping from wave functions to kets

The kets are abstract objects. We can think of them as complex vectors, since they do have properties of vectors (i.e. direction and length). The **inner product** between two kets  $|a\rangle$  and  $|b\rangle$  is written as

$$\langle a|b\rangle = c$$

where  $c$  is a complex number. The two kets are **orthogonal** if  
 $\langle a|b\rangle = 0$  .

The **norm** (or **length**) of a ket is given by

$$\left\| |a\rangle \right\| \equiv \sqrt{\langle a|a\rangle} .$$

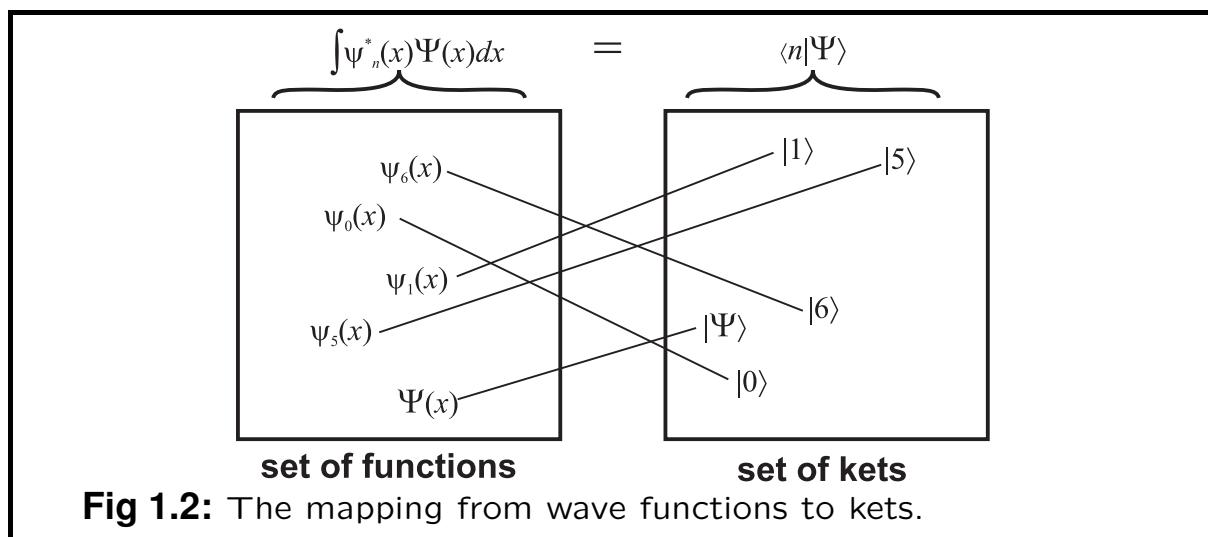
A ket is said to be **normalised** if it has unit norm; for example  $|a\rangle$  is normalised if

$$\left\| |a\rangle \right\|^2 = \langle a|a\rangle = 1 .$$

Two kets are **orthonormal** if, in addition to being orthogonal they are normalised, i.e.

$$\langle a|a\rangle = 1, \langle b|b\rangle = 1, \langle a|b\rangle = 0 .$$

There is a **one-to-one mapping** from each (and every) wave function to a unique ket as illustrated in the figure.



We can represent the mapping from an energy eigenfunction to its corresponding **energy eigenket** using the “maps to” symbol  $\mapsto$  as follows:

$$\psi_n(x) \mapsto |n\rangle .$$

Given that the mapping is one-to-one it can also be done in reverse, i.e.

$$|n\rangle \mapsto \psi_n(x) .$$

### KEY POINT

The value of the inner products on both sets is preserved by the mapping. Structure preserving maps are called isomorphisms (*iso* meaning equal and *morph* meaning shape); here the preserved structure is the inner product.

In fact, we construct the mapping to make this true. Thus the

energy eigenkets  $|n\rangle$  are **orthonormal**:

$$\langle n|m \rangle = \delta_{n,m} .$$

### Inner product

Note that any time you find expressions involving **products** of the form

$$g \cdots \langle W | \cdots f \cdots | Z \rangle \cdots$$

where the  $g$ ,  $f$  and  $\cdots$  represent complex numbers, the bra and the ket form an inner product, i.e.

$$g \cdots \langle W | \cdots f \cdots | Z \rangle \cdots = g \cdots f \cdots \langle W | Z \rangle \cdots .$$

### Outer product

On the other hand expressions involving products of the form

$$g \cdots |Z\rangle \cdots f \cdots \langle W | \cdots = g \cdots f \cdots |Z\rangle \langle W | \cdots$$

cannot be so contracted as it is an **outer product**. The nature of outer products will become apparent later.

Let the general state  $\Psi(x)$  be mapped to the ket  $|\Psi\rangle$ ,

$$\Psi(x) \mapsto |\Psi\rangle = \sum_{n=0}^{\infty} c_n |n\rangle , \quad (1.7)$$

where the coefficients  $c_n$  may be calculated by the inner product

$$\begin{aligned} \langle n | \Psi \rangle &= \langle n | \left( \sum_m c_m |m\rangle \right) = \sum_m c_m \langle n | m \rangle = \sum_m c_m \delta_{n,m} \\ &= c_n , \end{aligned} \quad (1.8)$$

and so Eq. (1.7) can be rewritten as

### Ket expansion

$$|\Psi\rangle = \sum_n \underbrace{\left( \langle n | \Psi \rangle \right)}_{\text{inner product}} |n\rangle . \quad (1.9)$$

Note that Eq. (1.4) and Eq. (1.9) are **equivalent expansions** because we have already stipulated that the **corresponding inner products are equal** (i.e. the mapping between functions and kets preserves the inner products) :

$$\int \psi_n^*(x) \Psi(x) dx = \langle n | \Psi \rangle .$$

Also let the other arbitrary wave function  $\Phi(x)$  be mapped to the

ket  $|\Phi\rangle$ ,

$$\Phi(x) \mapsto |\Phi\rangle = \sum_{n=0}^{\infty} d_n |n\rangle .$$

The inner product between  $|\Phi\rangle$  and  $|\Psi\rangle$  corresponding to Eq. (1.6) is given by

$$\begin{aligned}\langle\Phi|\Psi\rangle &= \langle\Phi| \left( \sum_{n=0}^{\infty} c_n |n\rangle \right) \\ &= \sum_{n=0}^{\infty} c_n (\langle\Phi|n\rangle) .\end{aligned}$$

In order that inner products be preserved by the mapping between functions and kets, we need the right hand side to equal that of Eq. (1.6) for arbitrary coefficients  $c_n$ , and so we find that

$$\langle\Phi|n\rangle = d_n^* .$$

This means that

### Bra version of a ket

$$\langle\Phi| = \sum_{n=0}^{\infty} d_n^* \langle n| \quad (1.10)$$

where

$$|\Phi\rangle = \sum_{n=0}^{\infty} d_n |n\rangle .$$

Using Eq. (1.10) we can now check the normalisation of  $|\Psi\rangle$ :

$$\begin{aligned}1 &= \langle\Psi|\Psi\rangle \\ &= \left( \sum_{n=0}^{\infty} c_n^* \langle n| \right) \left( \sum_{m=0}^{\infty} c_m |m\rangle \right) \\ &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m \langle n|m\rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} c_n^* c_m \delta_{n,m} \\ &= \sum_{n=0}^{\infty} |c_n|^2\end{aligned}$$

which shows that the probability distribution  $P_n = |c_n|^2$  is normalised to unity, as for the wave function case.

Another example of the mapping from wave functions into kets is given in the appendix, in terms of the electronic wave functions of an atom.

### NOTE

We used the harmonic oscillator example to give some meaning to bras and kets in terms of more familiar wave functions. However, in the following we discuss the bras and kets of an **arbitrary system**. The system need not be a harmonic oscillator or an electron in an atom.

## b. Hilbert spaces - in general

The kets form a linear complex vector space. A linear complex vector space is a set of elements  $V = \{|\psi\rangle\}$  for which certain operations are defined, such as vector addition

$$|c\rangle = |a\rangle + |b\rangle$$

and scalar multiplication

$$|d\rangle = x|a\rangle$$

for  $|a\rangle, |b\rangle, |c\rangle, |d\rangle \in V$  and  $x$  is a complex number. These spaces have other properties which we shall not need explicitly here.

### Basis set

A set of vectors such as  $\{|\phi_n\rangle : n = 1, 2, \dots\}$  is said to be a **basis** for  $V$  if any vector  $|\psi\rangle \in V$  can be written as a linear combination of its elements, i.e if  $|\psi\rangle = \sum_n c_n |\phi_n\rangle$ .

In this course we will deal with **Hilbert spaces**. In 1932, the renowned German mathematician John von Neumann formalised the fledgling quantum theory by developing a rigorous framework based on Hilbert spaces. His book, *Mathematical Foundations of Quantum mechanics*, translated by R.T. Beyer (Princeton University Press, Princeton, 1955), is a good source for advanced study.

A Hilbert space,  $H$ , is a special type of linear vector space in that it has the additional properties: an **inner product** (e.g.  $\langle a|b\rangle$ ) is defined on  $H$ , and  $H$  is **closed** with respect to this inner product. The inner product allows the notion of a **norm** of a vector according to

$$\| |a\rangle \| \equiv \sqrt{\langle a|a\rangle}$$

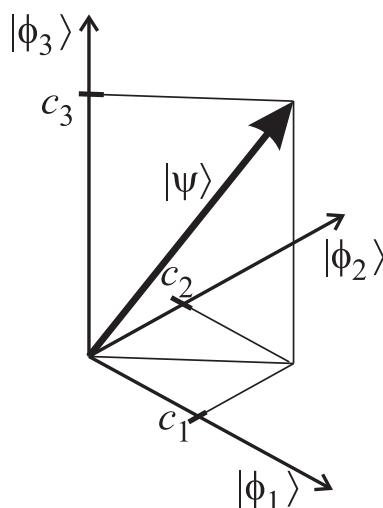
as we have already seen. All vectors in a Hilbert space have a **finite** norm. A norm is necessary in quantum mechanics to allow a probability interpretation. Typically we consider the subset of **normalised** states.

The **closure** with respect to the inner product is a mathematical detail which will not concern us in this course. It requires every convergent sequence of vectors in the space to have a limit point which is also in the space. Thus if  $|a_1\rangle, |a_2\rangle, \dots$  for  $|a_n\rangle \in H$  is a convergent sequence with limit point  $|a\rangle$  such that  $\| |a\rangle - |a_n\rangle \| \rightarrow 0$  as  $n \rightarrow \infty$  then the limit point  $|a\rangle$  must also be in  $H$ .

In this course we will use **orthonormal bases**, that is, a set of normalised, mutually orthogonal kets<sup>¶</sup>

$$\{|\phi_n\rangle : \langle\phi_n|\phi_m\rangle = \delta_{n,m}, n, m = 1, 2, \dots\}$$

which form a basis for the Hilbert space  $H$ . The expansion of a ket in terms of an orthonormal basis, such as  $|\psi\rangle = \sum_n c_n |\phi_n\rangle$ , can be regarded as the decomposition of a vector  $|\psi\rangle$  into components along mutually orthogonal axes whose unit vectors are the basis kets  $|\phi_n\rangle$ , as illustrated in the figure. Care must be taken with the literal interpretation of the figure because the coefficients are complex, in general.



**Fig 1.3:** A graphical illustration of the expansion

$$|\psi\rangle = \sum_{n=1}^3 c_n |\phi_n\rangle .$$

### Dimension of a Hilbert space

The **dimension** of a Hilbert space is the number of elements in any orthonormal basis.

The dimension of the Hilbert space for a spin-1/2 particle is 2 and for a harmonic oscillator and for hydrogen the dimension

<sup>¶</sup>Here the braces  $\{\dots\}$  define a set and the colon “:” is the symbol for “such as”. Thus the expression should be read as *The set of kets  $|\phi_n\rangle$  such that  $\langle\phi_n|\phi_m\rangle = \delta_{n,m}$  for  $n, m = 1, 2, \dots$*

is infinite. The mathematics of finite dimensional Hilbert spaces are rather straightforward, whereas there are many subtleties and special limiting conditions for infinite-dimensional Hilbert spaces.

There are further mathematical subtleties associated with the space of state vectors. The Hilbert space is, in fact, insufficient to contain all vectors of interest. In particular the eigenstates of position and momentum do not belong to a Hilbert space as they are not normalisable (loosely speaking they have infinite norm). A larger structure, known as the **rigged Hilbert space** is needed to accommodate vectors of this kind. The word “rigged” is used in the context of a sailing boat which is rigged and ready for sailing; here it means the space is equipped and ready for use in quantum mechanics. The text by Ballentine gives a brief description of its construction. However, we will not discuss it further in this course.

### KEY POINT

In this course we shall use the terms **state**, **state vector**, **vector**, **ket** and **bra** interchangeably. Indeed, they all refer to the quantum state of a system. Nevertheless, **ket** and **bra** are different mathematical representations of the state as given above and mean different things *mathematically*.

## c. Operators and vectors in the bra-ket formalism

### Operators

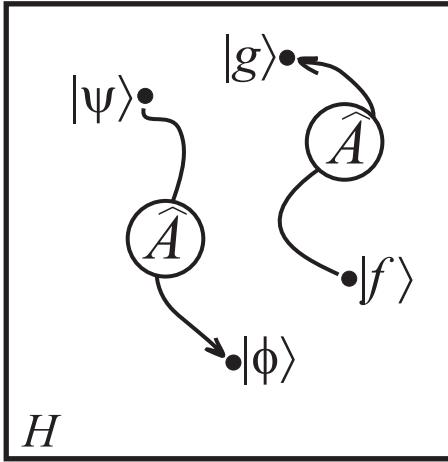
An operator on a Hilbert space is a mapping from one ket to another but staying within the set of kets. In other words it transforms one ket into another, just as the rotation operation transforms one vector into a rotated one in ordinary 3-dimensional space. For example, consider the operator  $\hat{A}$  acting on the ket  $|\psi\rangle$  to produce the ket  $|\phi\rangle$ :

insert  
1

$$\hat{A}|\psi\rangle = |\phi\rangle . \quad (1.11)$$

### NOTE

The “caret”  $\hat{\phantom{x}}$  over a symbol denotes that the symbol is an **operator**.



**Fig 1.4:** Mapping of two kets in the Hilbert space  $H$  by operator  $\hat{A}$ .

The simplest example of an operator in bra-ket notation is given by  $\hat{A} = |a\rangle\langle b|$ . The reason we call such an object an operator is that it maps any state in  $H$ , such as  $|\psi\rangle$ , onto another state in  $H$ , say  $|\phi\rangle$  as follows:

$$\begin{aligned}\hat{A}|\psi\rangle &= (|a\rangle\langle b|)|\psi\rangle \\ &= |a\rangle\langle b|\psi\rangle \\ &= |\phi\rangle\end{aligned}$$

where  $|\phi\rangle = (\langle b|\psi\rangle)|a\rangle$  is a scaled (i.e. multiplied by  $\langle b|\psi\rangle$ ) version of  $|a\rangle$ . The amount of scaling  $\langle b|\psi\rangle$  depends on the original state  $|\psi\rangle$ . This is how  $\hat{A}$  maps  $|\psi\rangle$  onto a scaled version of  $|a\rangle$ .

Notice that  $\hat{A}$  maps **every vector** onto a scaled version of  $|a\rangle$ . The amount of scaling  $\langle b|\psi\rangle$  depends on the original ket  $|\psi\rangle$ . The new ket is essentially in the same “direction” as  $|a\rangle$ . This is like a matrix which maps every vector onto a line.

A more interesting example of an operator is given by  $\hat{B} = \alpha|a\rangle\langle b| + \beta|c\rangle\langle d|$ . This gives the following mapping

$$\begin{aligned}\hat{B}|\psi\rangle &= (\alpha|a\rangle\langle b| + \beta|c\rangle\langle d|)|\psi\rangle \\ &= \alpha|a\rangle\langle b|\psi\rangle + \beta|c\rangle\langle d|\psi\rangle \\ &= (\langle b|\psi\rangle\alpha)|a\rangle + (\langle d|\psi\rangle\beta)|c\rangle \\ &= |\phi'\rangle.\end{aligned}$$

### KEY POINT

Here the “direction” of the new ket  $|\phi'\rangle$  depends on the initial ket  $|\psi\rangle$ , because it is  $\langle b|\psi\rangle\alpha$  along  $|a\rangle$  and  $\langle d|\psi\rangle\beta$  along  $|c\rangle$ .

There is a special operator,  $\hat{1}$ , called the **unit operator** or the **identity operator**, that leaves every ket unchanged, i.e.

insert  
2

$$\hat{\mathbb{1}}|\psi\rangle = |\psi\rangle$$

for every  $|\psi\rangle$  in  $H$ .

### Resolution of the identity

The expression

$$\hat{\mathbb{1}} = \sum_n |a_n\rangle\langle a_n| \quad (1.12)$$

for an orthonormal basis  $\{|a_n\rangle : n = 1, 2, \dots\}$  is called the **resolution of the identity** (or the **closure** or **completeness relation**). The continuous counterpart of this is given by

$$\hat{\mathbb{1}} = \int |\beta\rangle\langle\beta| d\beta.$$

We will tend to use the discrete spectrum case in this course. **Every basis** can be used as a resolution of the identity in this manner.

It is interesting to compare Eq. (1.12) with the analogous expression for wave functions Eq. (1.5).

To see how the identity operator can be used, let  $|\psi\rangle$  have the expansion

$$|\psi\rangle = \sum_n c_n |a_n\rangle \quad (1.13)$$

for coefficients  $c_n$  in the orthonormal basis  $\{|a_n\rangle : n = 1, 2, \dots\}$ . The orthonormality implies

$$\langle a_n | a_m \rangle = \delta_{n,m}.$$

The action of the identity operator on  $|\psi\rangle$  is given by<sup>¶</sup>

$$\begin{aligned} \hat{\mathbb{1}}|\psi\rangle &= \left( \sum_n |a_n\rangle\langle a_n| \right) \left( \sum_m c_m |a_m\rangle \right) \\ &= \sum_{n,m} c_m |a_n\rangle \left( \langle a_n | a_m \rangle \right) = \sum_{n,m} c_m |a_n\rangle \delta_{n,m} = \sum_n c_n |a_n\rangle \\ &= |\psi\rangle \end{aligned}$$

as expected.

<sup>¶</sup>Note that the product of  $\langle a_n |$  with  $|a_m\rangle$  is simply the inner product  $\langle a_n | a_m \rangle = \langle a_n | a_m \rangle$ .

### Vector representation of a ket

The coefficients  $c_n$  form the elements of a column vector

$$\vec{\psi} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix} . \quad (1.14)$$

$\vec{\psi}$  is the vector representation of  $|\psi\rangle$  in the  $\{|a_n\rangle\}$  basis.

An arbitrary operator  $\hat{A}$  can be written, in the basis  $\{|a_n\rangle : n = 1, 2, \dots\}$ , as

insert  
3

$$\hat{A} = \sum_{n,m} A_{n,m} |a_n\rangle \langle a_m| \quad (1.15)$$

where the coefficients  $A_{n,m}$  form the elements of a matrix as follows:

### Matrix elements & matrix representation

$$\underline{A} = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} & \cdots \\ A_{2,1} & A_{2,2} & A_{2,3} & \cdots \\ A_{3,1} & A_{3,2} & A_{3,3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (1.16)$$

The complex numbers  $A_{n,m}$  are called the **matrix elements** of  $\hat{A}$  in the  $\{|a_n\rangle : n = 1, 2, \dots\}$  basis, and the matrix  $\underline{A}$  is the **matrix representation** of  $\hat{A}$  in the same basis.

### Problem

Show that the identity operator in Eq. (1.12) leaves an arbitrary operator, such as  $\hat{A}$  in Eq. (1.15), unchanged. That is, show

$$\hat{1} \times \hat{A} = \hat{A} = \hat{A} \times \hat{1} .$$

We can now see how  $\hat{A}$  transforms vectors. From Eq. (1.13) and Eq. (1.15) we find

$$\begin{aligned} \hat{A}|\psi\rangle &= \left( \sum_{n,m} A_{n,m} |a_n\rangle \langle a_m| \right) \left( \sum_j c_j |a_j\rangle \right) \\ &= \sum_{n,m} \sum_j A_{n,m} c_j |a_n\rangle \left( \langle a_m | a_j \rangle \right) = \sum_{n,m} \sum_j A_{n,m} c_j |a_n\rangle \left( \delta_{m,j} \right) \\ &= \sum_{n,m} A_{n,m} c_m |a_n\rangle . \end{aligned}$$

But from Eq. (1.11) the left hand side is equal to  $|\phi\rangle$ , and so we have

$$|\phi\rangle = \sum_n \left( \sum_m A_{n,m} c_m \right) |a_n\rangle$$

insert  
4

$$= \sum_n d_n |a_n\rangle \quad (1.17)$$

where  $d_n = \sum_m A_{n,m} c_m$  are the coefficients of the transformed vector.

### KEY POINT

The matrix elements  $A_{n,m}$  of  $\hat{A}$  determine how  $\hat{A}$  transforms kets. Indeed, if  $\vec{\phi}$  is the vector representation of  $|\phi\rangle$  then it is straightforward to show that

$$\underline{\mathbf{A}} \vec{\psi} = \vec{\phi},$$

i.e. that the matrix representation of  $\hat{A}$  transforms the vector representation of  $|\psi\rangle$  under ordinary matrix multiplication.

The connection between matrices and vectors and operators and kets is a strong one. In fact we can do all calculations using the matrix and vector representations only. Indeed it is not difficult to show the following.

### Equivalence between operator and matrix algebra

Consider three operators  $\hat{A}$ ,  $\hat{B}$  and  $\hat{C}$  related by

$$\hat{A}\hat{B} = \hat{C},$$

and let the operators have the corresponding matrix representations  $\underline{\mathbf{A}}$ ,  $\underline{\mathbf{B}}$  and  $\underline{\mathbf{C}}$  in some basis. Then

$$\underline{\mathbf{A}} \underline{\mathbf{B}} = \underline{\mathbf{C}}.$$

However, we continue to use the bra-ket notations because it does offer some simplification in certain circumstances.

### Dual space and bras

In defining inner products we have been using the symbol  $\langle\psi|$  for a bra vector. Actually there is a complete Hilbert space of bras  $\langle\psi|$ , one for each corresponding ket  $|\psi\rangle$ . It might seem a trivial (or pointless) thing to regard  $\langle\psi|$  as being different from  $|\psi\rangle$ . However recall that the square of the length of a real row vector, such as  $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ , is given by the inner product with its

vector, such as  $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ , is given by the inner product with its

transpose:<sup>‡</sup>

$$\|\vec{v}\|^2 = (\vec{v})^\top \vec{v} = [x, y] \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + y^2 .$$

The column vector is  $\begin{bmatrix} x \\ y \end{bmatrix}$  is a quite different mathematical object to the row vector  $[x, y]$ , and this is analogous to the difference between bras and kets.

For **complex vectors**, square of the length of a vector involves taking both the complex conjugate and the transpose as follows:

$$\|\vec{v}\|^2 = [(\vec{v})^\top]^* \vec{v} = [x^*, y^*] \begin{bmatrix} x \\ y \end{bmatrix} = x^*x + y^*y = |x|^2 + |y|^2 .$$

### Hermitian conjugation (conjugate transpose, Hermitian transpose)

The combination of the transpose and complex conjugation operations is known as the **Hermitian conjugate** and has the symbol  $\dagger$  (literally a dagger). The operation applies to matrices and vectors as well as operators and kets.

Thus, for example,

$$\begin{aligned} |\psi\rangle^\dagger &= \langle\psi| = \sum_{n=0}^{\infty} c_n^* \langle a_n | & (1.18) \\ \langle\psi|^\dagger &= |\psi\rangle = \sum_{n=0}^{\infty} c_n | a_n \rangle \end{aligned}$$

insert  
5

$$\vec{\psi}^\dagger = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix}^\dagger = [c_1^* \ c_2^* \ \dots]$$

where  $\vec{\psi}$  is the vector representation of  $|\psi\rangle$  in the  $\{|a_n\rangle\}$  basis and we have used Eq. (1.10) for the expansion of the bra  $\langle\Psi|$ .

---

<sup>‡</sup>The symbol  ${}^\top$  denotes the transpose of a vector or matrix, e.g.  $\begin{bmatrix} x \\ y \end{bmatrix} {}^\top = [x, y]$  and  $[x, y] {}^\top = \begin{bmatrix} x \\ y \end{bmatrix}$ .

### Vector representation of a bra

The coefficients  $c_n^*$  in Eq. (1.18) form the elements of the row vector

$$\overrightarrow{\psi}^\dagger = [c_1^* \ c_2^* \ \dots] . \quad (1.19)$$

$\overrightarrow{\psi}^\dagger$  is the vector representation of  $\langle\psi|$ , and  $\overrightarrow{\psi}$  is the vector representation of  $|\psi\rangle$  in the  $\{|n\rangle\}$  basis.

This identification of  $\overrightarrow{\psi}^\dagger$  as the vector representation of  $\langle\phi|$  allows us to evaluate inner product  $\langle\psi|\psi\rangle$  in terms of the vector representation as

$$\langle\psi|\psi\rangle = \overrightarrow{\psi}^\dagger \cdot \overrightarrow{\psi} = \begin{bmatrix} c_1^* & c_2^* & \dots \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix} = \sum_n |c_n|^2 .$$

The Hilbert space comprising the bras is called the **dual Hilbert space** or the **dual of  $H$** .

### Expanding kets and operators

Up to now we have **given** the expansion of arbitrary kets explicitly. We now show how the expansion can be **calculated** for any orthonormal basis  $\{|a_n\rangle : n = 1, 2, \dots\}$ . First we construct the resolution of the identity  $\hat{\mathbb{1}} = \sum_n |a_n\rangle\langle a_n|$  in this basis and then we apply it to  $|\psi\rangle$  as follows:

$$\begin{aligned} |\psi\rangle &= \hat{\mathbb{1}} \times |\psi\rangle = \left( \sum_n |a_n\rangle\langle a_n| \right) |\psi\rangle = \sum_n |a_n\rangle \left( \langle a_n|\psi\rangle \right) \\ &= \sum_n c_n |a_n\rangle \end{aligned} \quad (1.20)$$

where  $c_n = \langle a_n|\psi\rangle$  are the **coefficients** of  $|\psi\rangle$  in the  $\{|a_n\rangle\}$  basis.

Similarly, every operator can be written in terms of bras and kets as follows with respect to a basis. In the  $\{|a_n\rangle\}$  basis the bra-ket

representation of  $\hat{A}$  is

$$\begin{aligned}
\hat{A} &= \hat{\mathbf{1}} \times \hat{A} \times \hat{\mathbf{1}} = \left( \sum_n |a_n\rangle\langle a_n| \right) \hat{A} \left( \sum_m |a_m\rangle\langle a_m| \right) \\
&= \sum_{n,m} |a_n\rangle\langle a_n| \hat{A} |a_m\rangle\langle a_m| \\
&= \sum_{n,m} |a_n\rangle\langle a_m| \left( \langle a_n|\hat{A}|a_m\rangle \right) \\
&= \sum_{n,m} A_{n,m} |a_n\rangle\langle a_m|
\end{aligned} \tag{1.21}$$

where

### Matrix elements

$$A_{n,m} = \langle a_n|\hat{A}|a_m\rangle \tag{1.22}$$

are the **matrix elements** of  $\hat{A}$  in the  $\{|a_n\rangle\}$  basis that we met before in Eq. (1.16).

Note that this representation is not unique; for each orthonormal basis we get a different representation.

Moreover, if  $\hat{A}$  has an orthonormal **eigenbasis**  $\{|\lambda_n\rangle : n = 1, 2, \dots\}$  where  $\hat{A}|\lambda_n\rangle = \lambda_n|\lambda_n\rangle$  then from Eq. (1.21) we find<sup>¶</sup>

$$\begin{aligned}
\hat{A} &= \sum_{n,m} |\lambda_n\rangle\langle\lambda_m| \left( \langle\lambda_n|\hat{A}|\lambda_m\rangle \right) = \sum_{n,m} |\lambda_n\rangle\langle\lambda_m| \left( \lambda_m \langle\lambda_n|\lambda_m\rangle \right) \\
&= \sum_{n,m} |\lambda_n\rangle\langle\lambda_m| \left( \lambda_m \delta_{n,m} \right)
\end{aligned}$$

and so

### Eigenket expansion of an operator

$$\hat{A} = \sum_n \lambda_n |\lambda_n\rangle\langle\lambda_n| \tag{1.23}$$

---

<sup>¶</sup>This is the case for every operator on a *finite* dimensional Hilbert space.

**Problem**

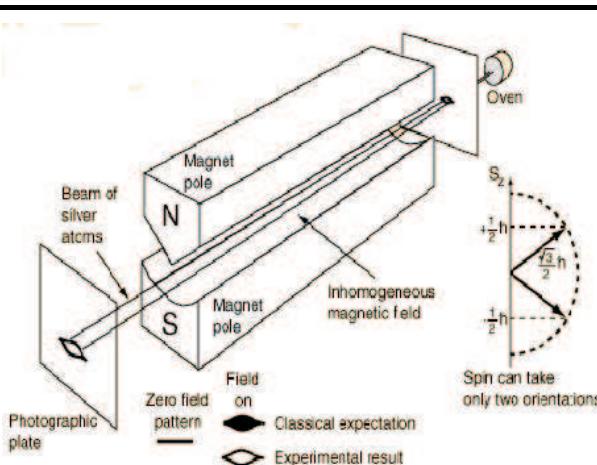
Verify that the representation of  $\hat{A}$  in Eq. (1.23) yields the eigenket equation

$$\hat{A}|\lambda_k\rangle = \lambda_k|\lambda_k\rangle$$

for orthonormal states  $\{|\lambda_n\rangle : n = 1, 2, \dots\}$ .

## d. Spin-1/2 systems

Imagine a beam of silver atoms in a Stern-Gerlach experiment. We are interested in the state  $|\psi\rangle$  of the spin of the unpaired electron in the 4d orbital in an atom that follows the **upper path**.<sup>¶</sup> Let this path represent the spin being **spin up**  $\uparrow$ .



**Fig 1.5:** Stern-Gerlach setup. Image from <http://www.geocities.com/slakerus/fclas/node8.html>

In this case,  $|\psi\rangle$  would be represented as

$$|\psi\rangle = |\frac{1}{2}\rangle_z \quad (1.24)$$

where the  $\frac{1}{2}\hbar$  represents the  $z$ -component of the spin angular momentum of the electron.<sup>§</sup> Correspondingly, the spin state of an electron in the **lower path** would be

$$|\psi\rangle = |-\frac{1}{2}\rangle_z . \quad (1.25)$$

These states form an orthonormal basis for the state of the elec-

<sup>¶</sup>The ground state of Au is  $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^6 5s^2 4d^9$ . There is one unpaired spin “x” in the 4d orbitals  $[\uparrow\downarrow \uparrow\downarrow \uparrow\downarrow x]$  which could be  $\uparrow$  or  $\downarrow$ .

<sup>§</sup>Recall that  $\hbar = h/(2\pi)$ ,  $h$ =Planck's constant= $6.63 \times 10^{-34}$  Js.

tron's spin:

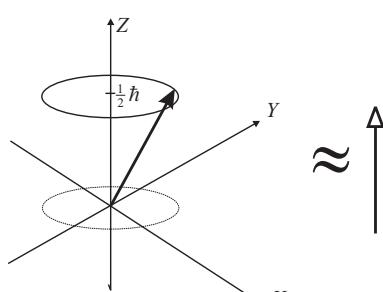
$$\begin{aligned} z\langle \frac{1}{2} | \frac{1}{2} \rangle_z &= 1 \\ z\langle \frac{1}{2} | -\frac{1}{2} \rangle_z &= 0 \\ z\langle -\frac{1}{2} | \frac{1}{2} \rangle_z &= 0 \\ z\langle -\frac{1}{2} | -\frac{1}{2} \rangle_z &= 1 . \end{aligned}$$

They are also eigenkets of the  $z$ -component of spin, which according to Eq. (1.23), is given by

$$\hat{S}_z = \frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle\frac{1}{2}| - \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle-\frac{1}{2}| . \quad (1.26)$$

### Comment

It is common to see the state  $|\frac{1}{2}\rangle_z$  represented as an angular momentum vector of length  $\sqrt{\frac{1}{2}(\frac{1}{2}+1)}\hbar$  randomly precessing about the  $z$  axis as illustrated in the figure below.



**Fig 1.6:** Representation of spin “up”.

The random precessing is introduced to account for  $\langle \hat{S}_x \rangle = \langle \hat{S}_y \rangle = 0$ . This picture is further abbreviated to simply an up-pointing arrow  $\uparrow$ . However you should remember that there is really no precessing at all. This state can be expressed in the  $\hat{S}_x$  and  $\hat{S}_y$  eigenbasis as

$$\begin{aligned} |\frac{1}{2}\rangle_z &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_x + |-\frac{1}{2}\rangle_x \right) \\ &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_y + |-\frac{1}{2}\rangle_y \right) , \end{aligned}$$

and so the reason  $\langle \hat{S}_x \rangle = 0$  is because the spin is in an equal superposition of states with  $\hat{S}_x$  values  $\frac{1}{2}\hbar$  and  $-\frac{1}{2}\hbar$ . You can compare this with an electron in the 1s orbital of hydrogen: the electron is not “randomly orbiting the nucleus” but rather behaves as a wave which is spread around the nucleus. Nevertheless, such conceptual pictures involving random motion are sometimes helpful.

## Ordering of elements in matrix representations

In forming the vector and matrix representation we are implicitly **assuming a specific ordering** of the basis vectors. One way to remember the ordering is to make the associations:

$$|\psi\rangle \iff \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix} \begin{matrix} |a_1\rangle \\ |a_2\rangle \\ \vdots \end{matrix}$$

for  $|\psi\rangle = \sum_n c_n |a_n\rangle$  and

$$\hat{A} \iff \begin{matrix} \langle a_1| & \langle a_2| & \dots \\ |a_1\rangle & \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots \\ A_{2,1} & A_{2,2} & \cdots \\ \vdots & \vdots \end{bmatrix} \\ |a_2\rangle \\ \vdots \end{matrix}$$

for  $\hat{A} = \sum_{n,m} A_{n,m} |a_n\rangle\langle a_m|$ . You must keep this in mind when using the matrix representations.

### The spin-1/2 operators

The three operators corresponding to the components of spin angular momentum along the  $x$ ,  $y$  and  $z$  axes are represented in the  $\{|\frac{1}{2}\rangle_z, |-\frac{1}{2}\rangle_z\}$  basis as

$$\begin{aligned} \hat{S}_x &= \frac{1}{2}\hbar |\frac{1}{2}\rangle_z \langle -\frac{1}{2}| + \frac{1}{2}\hbar |-\frac{1}{2}\rangle_z \langle \frac{1}{2}| \\ \hat{S}_y &= -i\frac{1}{2}\hbar |\frac{1}{2}\rangle_z \langle -\frac{1}{2}| + i\frac{1}{2}\hbar |-\frac{1}{2}\rangle_z \langle \frac{1}{2}| \\ \hat{S}_z &= \frac{1}{2}\hbar |\frac{1}{2}\rangle_z \langle \frac{1}{2}| - \frac{1}{2}\hbar |-\frac{1}{2}\rangle_z \langle -\frac{1}{2}| . \end{aligned} \quad (1.27)$$

The set of 3 operators satisfy the **commutation relation**

insert  
6

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

and cyclic permutations of it. For example,  $[\hat{S}_z, \hat{S}_x] = i\hbar \hat{S}_y$ .

### Spin eigenket notation

We will use the following notation for the eigenkets of a spin operator where, for example,  $|\frac{1}{2}\rangle_x$  is the eigenstate of  $\hat{S}_x$  with eigenvalue  $\frac{1}{2}\hbar$ , and so

$$\begin{aligned} \hat{S}_x |\frac{1}{2}\rangle_x &= \frac{1}{2}\hbar |\frac{1}{2}\rangle_x \\ \hat{S}_x |-\frac{1}{2}\rangle_x &= -\frac{1}{2}\hbar |-\frac{1}{2}\rangle_x . \end{aligned}$$

The subscript  $x$  denotes the kets  $|\pm\frac{1}{2}\rangle_x$  as eigenstates of  $\hat{S}_x$ , and similarly for  $\hat{S}_y$  and  $\hat{S}_z$ . We will however tend to work exclusively in the so-called  **$z$  basis**  $\{|\frac{1}{2}\rangle_z, |-\frac{1}{2}\rangle_z\}$ .

### Eigenkets of $\hat{S}_x$ and $\hat{S}_y$

The eigenkets of  $\hat{S}_x$  in the  $\{|-\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z\}$  basis are given by

$$\begin{aligned} |\frac{1}{2}\rangle_x &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z + |-\frac{1}{2}\rangle_z \right) \\ |-\frac{1}{2}\rangle_x &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z - |-\frac{1}{2}\rangle_z \right), \end{aligned} \quad (1.28)$$

and those of  $\hat{S}_y$  are

$$\begin{aligned} |\frac{1}{2}\rangle_y &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z + i|-\frac{1}{2}\rangle_z \right) \\ |-\frac{1}{2}\rangle_y &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z - i|-\frac{1}{2}\rangle_z \right). \end{aligned} \quad (1.29)$$

### Problem

Verify that Eq. (1.28) and Eq. (1.29) do indeed define the eigenkets of the spin operators  $\hat{S}_x$  and  $\hat{S}_y$ .

### Problem

Use Eq. (1.27) to show that that the matrix representations of  $\hat{S}_x$  and  $\hat{S}_y$  in the  $\{|-\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z\}$  basis are

$$\begin{aligned} \mathbf{S}_x &= \frac{1}{2}\hbar \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \mathbf{S}_y &= \frac{1}{2}\hbar \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \end{aligned}$$

## e. The physical interpretation of operators and states

Thus far, we have primarily dealt with the *mathematics* of quantum mechanics. But this theory claims to be a theory the *physical* world. Given this, how do abstract entities such as state vectors and operators relate to the ‘real’ world? The answer is that, following certain well-established rules, we can use them to calculate the statistical properties of experiments we might do in the laboratory, e.g., spin-polarisation experiments, cooling Helium atoms, experiments with lasers. These rules can be used to manipulate state vectors and operators to produce results relevant to what we observe.

### Observables.

The physical quantities, such as energy, momentum, angular momentum etc., are known as **observables** and are represented by **self-adjoint** operators. A self-adjoint operator  $\hat{A}$  on inner product

space  $V$  is one where

$$\langle a|\hat{A}|b\rangle = \langle Aa|b\rangle = \langle a|Ab\rangle = \langle a|\hat{A}^\dagger|b\rangle , \quad (1.30)$$

where here  $|Ax\rangle = \hat{A}|x\rangle$ , for all states  $|a\rangle, |b\rangle \in V$ .

### Important

Note that  $\langle Aa| = (\langle Aa\rangle)^\dagger = (\hat{A}|a\rangle)^\dagger = \langle a|\hat{A}^\dagger$ .

As Eq. (1.30) is valid for all vectors we can just write that  $\hat{A}$  is self-adjoint if

insert  
7

$$\hat{A} = \hat{A}^\dagger.$$

A **Hermitian** operator on a Hilbert space  $H$  is a bounded self-adjoint operator on  $H$ , that is, a self-adjoint operator for which  $\|\hat{A}|a\rangle\| < \infty$  for all  $|a\rangle \in H$ .

If the Hilbert space has a finite dimension all the operators on it are bounded and so all the self-adjoint operators are also Hermitian.

### Problem

Show that the three spin operators  $\hat{S}_x$ ,  $\hat{S}_y$  and  $\hat{S}_z$  are self-adjoint (or equivalently Hermitian).

**Spectral theorem.** The eigenkets of a self-adjoint operator form an **orthonormal basis for the space**. Thus for the self-adjoint operators

$$\hat{A} = \sum_n a_n |a_n\rangle\langle a_n| , \quad \hat{B} = \int \beta |\beta\rangle\langle\beta| d\beta ,$$

which have a discrete and continuous eigenvalue spectrums, an arbitrary state  $|\psi\rangle$  can be written as

$$|\psi\rangle = \sum_n c_n |a_n\rangle , \quad |\psi\rangle = \int c(\beta) |\beta\rangle d\beta ,$$

respectively. Moreover, the eigenvalues of self-adjoint operators are **real**. This property is the reason why observables are represented by self-adjoint operators, because physical observables, when measured, can only yield **real values**.

### Probability interpretation

Let  $\{|\phi_i\rangle\}$  be an orthonormal basis and consider a system to be

in the superposition state

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle .$$

Let the state be normalised, i.e.  $\langle\psi|\psi\rangle = \sum_i |c_i|^2 = 1$ . The probability  $P_i$  of “finding” the system in the state  $|\phi_i\rangle$  is  $|c_i|^2$ . The **orthogonality** of the states  $\{|\phi_i\rangle\}$  implies that the events “finding the system in  $|\phi_i\rangle$ ” are **mutually exclusive**. (That is, if the system is actually in  $|\phi_k\rangle$  then  $P_i = 0$  unless  $i = k$ .) The probability of finding the system in any of the states  $\{|\phi_i\rangle\}$  is the sum  $\sum_i P_i = \sum_i |c_i|^2$ , which from the normalisation of  $|\psi\rangle$  is 1.

The wording “finding something” here means performing a measurement on the system and getting “something” as the measurement result. Note that  $c_i$  can be complex.

### Probability for finding a system in a state

For a system initially in state  $|\psi\rangle$ , the probability of finding the system in the state  $|\phi\rangle$  after an appropriate measurement is given by

$$P(\phi|\psi) = |\langle\phi|\psi\rangle|^2 .$$

Here  $P(a|b)$  is a conditional probability: the probability of getting event  $a$  given that event  $b$  has occurred.

Consider an observable

insert  
8

$$\hat{A} = \sum_n a_n |a_n\rangle\langle a_n| \quad (1.31)$$

whose eigenkets are  $\{|a_n\rangle\}$ . The probability of finding the system in the eigenket  $|a_k\rangle$  is

$$P_k = |\langle a_k|\psi\rangle|^2 . \quad (1.32)$$

Note that the **value** of the observable for this eigenket is  $a_k$ . Thus there is a probability of  $P_k$  of getting a value of  $a_k$  from a measurement of  $\hat{A}$  on a system which is in the state  $|\psi\rangle$ . This leads to the definition of the expectation value of  $\hat{A}$ .

### Expectation (average) values

Observables seem to contain an element of intrinsic randomness. That is, no matter how carefully we might perform our experiments, these quantities often (but not always) fluctuate randomly when measurements repeated a number of times. From statistics, the expected value or average value for an event which has

a value of  $a_n$  with probability  $P_n$  is given by

$$\text{expected value} = \sum_n a_n P_n .$$

For the quantum case we define the *expectation value* of the observable  $\hat{A}$  for a system in the state  $|\psi\rangle$  as

$$\begin{aligned}\langle \hat{A} \rangle &= \sum_n a_n P_n \\ &= \sum_n a_n |\langle a_n | \psi \rangle|^2 \text{ from Eq. (1.32)} \\ &= \sum_n a_n \langle a_n | \psi \rangle^* \langle a_n | \psi \rangle \\ &= \sum_n a_n \langle \psi | a_n \rangle \langle a_n | \psi \rangle = \langle \psi | \left( \sum_n a_n |a_n \rangle \langle a_n| \right) | \psi \rangle \\ &= \langle \psi | \hat{A} | \psi \rangle \text{ from Eq. (1.31)} .\end{aligned}$$

### Expectation value

The expectation value of an observable  $\hat{A}$  for a system in state  $|\psi\rangle$  is given by

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle .$$

The expectation value is the average value of the observable if a measurement was to be performed on the system to determine which eigenstate of  $\hat{A}$  the system is in.

The definition of the expectation value can be extended to the  $n^{\text{th}}$ -order moment of the observable  $\hat{A}$  as  $\langle \hat{A}^n \rangle = \langle \psi | \hat{A}^n | \psi \rangle$ .

## f. The Schrödinger equation

Quantum systems typically change over time; their change obeys a time-evolution equation. This equation is known as the **Schrödinger equation**. In the bra-ket notation, it evolves a state vector in time according to

### Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi\rangle = \hat{\mathcal{H}} |\psi\rangle ,$$

where  $\hat{\mathcal{H}}$  is the Hamiltonian for the physical system of interest.

The Schrödinger equation only applies to **isolated** systems. For

systems that interact with the things around them (e.g. by losing heat or even being “measured” by a measuring apparatus), the situation is more complicated.

## Appendix

### Relationship between wave functions and kets - atom example

The bra-ket notation is a convenient shorthand for wave mechanics. The solutions of Schrödinger's equation for the hydrogen atom has the form

$$\psi_{n,l,m_l}(r, \theta, \phi) = R_n(r)Y_{l,m_l}(\theta, \phi) .$$

(Note that this neglects the electron spin.) These wavefunctions are the energy eigenfunctions and they form a basis of orthogonal functions, e.g.

$$\int \psi_{n,l,m_l}^*(r, \theta, \phi) \psi_{n',l',m'_l}(r, \theta, \phi) r^2 \sin(\theta) dr d\theta d\phi = \delta_{n,n'} \delta_{l,l'} \delta_{m_l,m'_l} .$$

In general, however, an atom can be in a superposition of such wavefunctions, for example, a general wave function would be

$$\Psi(r, \theta, \phi) = \sum_{n,l,m_l} c_{n,l,m_l} \psi_{n,l,m_l}(r, \theta, \phi) \quad (1.33)$$

where  $c_{n,l,m_l}$  is the probability amplitude for being in the eigenfunction  $\psi_{n,l,m_l}(r, \theta, \phi)$ . In fact these coefficients are obtained by the integral form of the inner product

$$c_{n,l,m_l} = \int \psi_{n,l,m_l}^*(r, \theta, \phi) \Psi(r, \theta, \phi) r^2 \sin(\theta) dr d\theta d\phi . \quad (1.34)$$

This means that the expansion Eq. (1.33) is actually

#### Wave function expansion

$$\begin{aligned} \Psi(r, \theta, \phi) &= \sum_{n,l,m_l} \left[ \underbrace{\int \psi_{n,l,m_l}^*(r, \theta, \phi) \Psi(r, \theta, \phi) r^2 \sin(\theta) dr d\theta d\phi}_{\text{inner product}} \right] \psi_{n,l,m_l}(r, \theta, \phi) \end{aligned} \quad (1.35)$$

Let  $\Phi(r, \theta, \phi) = \sum_{n,l,m_l} d_{n,l,m_l} \psi_{n,l,m_l}(r, \theta, \phi)$  be another wave func-

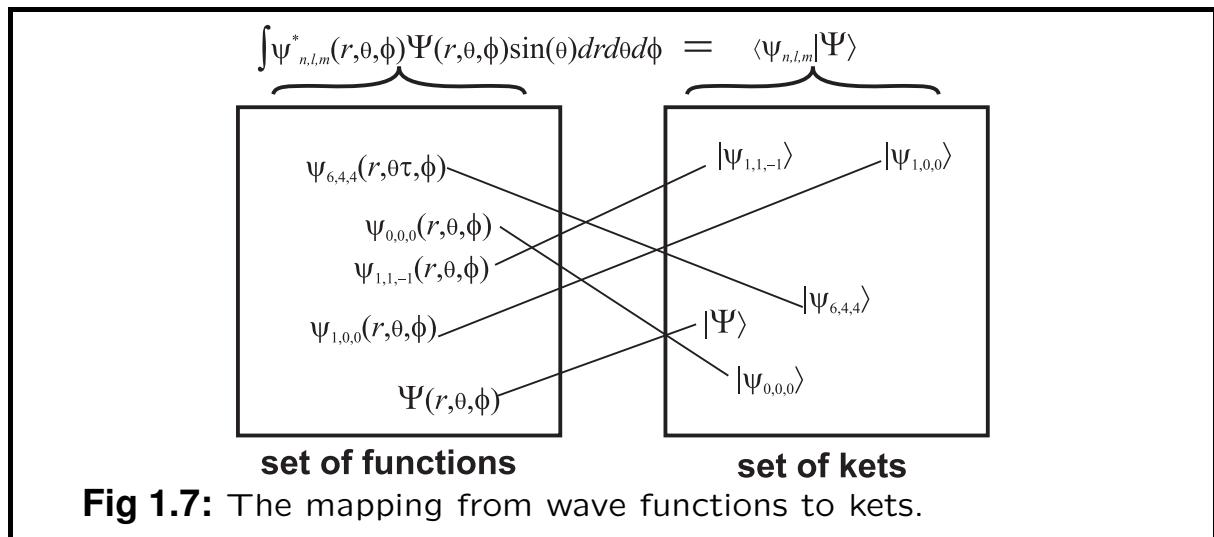
tion. The inner product with  $\Psi(r, \theta, \phi)$  is given by

$$\begin{aligned}
& \int \Psi^*(r, \theta, \phi) \Phi(r, \theta, \phi) r^2 \sin(\theta) dr d\theta d\phi \\
&= \sum_{n,l,m_l} \sum_{n',l',m'_l} \int \psi_{n,l,m_l}^*(r, \theta, \phi) \psi_{n',l',m'_l}(r, \theta, \phi) r^2 \sin(\theta) dr d\theta d\phi \\
&= \sum_{n,l,m_l} c_{n,l,m_l}^* d_{n,l,m_l} .
\end{aligned} \tag{1.36}$$

This notation is rather cumbersome. Consider now the same situation in Dirac's bra-ket form.

### Mapping from wave functions to kets

The kets are abstract objects. We can think of them as complex vectors, since they do have properties of vectors (i.e. direction and length). They form a set. There is a one-to-one mapping from each (and every) wave function to a unique ket as illustrated in the figure.



**Fig 1.7:** The mapping from wave functions to kets.

We can represent the mapping from an energy eigenfunction to its corresponding ket using the “maps to” symbol  $\mapsto$  as follows:

$$|\psi_{n,l,m_l}\rangle \mapsto \psi_{n,l,m_l}(r, \theta, \phi) .$$

Given that the mapping is one-to-one it can also be done in reverse, i.e.

$$\psi_{n,l,m_l}(r, \theta, \phi) \mapsto |\psi_{n,l,m_l}\rangle .$$

The value of the inner products on both sets is preserved by the mapping. In fact, we construct the mapping to make this true. Thus the kets  $|\psi_{n,l,m_l}\rangle$  are the energy **eigenkets** and they are orthonormal

$$\langle \psi_{n,l,m_l} | \psi_{n',l',m'_l} \rangle = \delta_{n,n'} \delta_{l,l'} \delta_{m_l,m'_l} .$$

The general state  $\Psi(r, \theta, \phi)$  is mapped to the ket  $|\Psi\rangle$ ,

$$\Psi(r, \theta, \phi) \mapsto |\Psi\rangle = \sum_{n,l,m_l} c_{n,l,m_l} |\psi_{n,l,m_l}\rangle ,$$

where the coefficients  $c_{n,l,m_l}$  may be calculated by the inner product

$$\begin{aligned} & \langle \psi_{n,l,m_l} | \Psi \rangle \\ &= \langle \psi_{n,l,m_l} | \left( \sum_{n',l',m'_l} c_{n',l',m'_l} |\psi_{n',l',m'_l}\rangle \right) = \sum_{n',l',m'_l} c_{n',l',m'_l} \langle \psi_{n,l,m_l} | \psi_{n',l',m'_l} \rangle \\ &= c_{n,l,m_l} , \end{aligned}$$

and so

### Ket expansion

$$|\Psi\rangle = \sum_{n,l,m_l} \underbrace{[\langle \psi_{n,l,m_l} | \Psi \rangle]}_{\text{inner product}} |\psi_{n,l,m_l}\rangle . \quad (1.37)$$

Note that Eq. (1.35) and Eq. (1.37) are **equivalent expansions** because we have already stipulated that the **corresponding inner products are equal** (i.e. the function-ket mapping preserves the inner products) :

$$\int \psi_{n,l,m_l}^*(r, \theta, \phi) \Psi(r, \theta, \phi) r^2 \sin(\theta) dr d\theta d\phi = \langle \psi_{n,l,m_l} | \Psi \rangle .$$

Let  $|\Phi\rangle = \sum d_{n,l,m_l} |\psi_{n,l,m_l}\rangle$  be the ket associated with the wave function  $\Phi(r, \theta, \phi) = \sum d_{n,l,m_l} \psi_{n,l,m_l}(r, \theta, \phi)$ . The inner product with  $|\Psi\rangle$  is given by simply

$$\begin{aligned} \langle \Psi | \Phi \rangle &= \sum_{n,l,m_l} \sum_{n',l',m'_l} c_{n,l,m_l}^* d_{n',l',m'_l} \langle \psi_{n,l,m_l} | \psi_{n',l',m'_l} \rangle \\ &= \sum_{n,l,m_l} c_{n,l,m_l}^* d_{n,l,m_l} . \end{aligned} \quad (1.38)$$

You should compare the corresponding expressions Eq. (1.36) and Eq. (1.38) for the wave function and bra-ket cases to see how more convenient the bra-ket notation is.

# Problem Sheet

- (1) Using the expansion Eq. (1.2) for  $\Psi(x)$  and the orthonormality relation Eq. (1.1) check that the right hand side of Eq. (1.3) does indeed equal  $c_n$ .

- (2) Consider the kets

$$|a\rangle = \frac{1}{\sqrt{2}}(|\psi\rangle + |\phi\rangle)$$

$$|b\rangle = \frac{1}{\sqrt{3}}|\psi\rangle + i\sqrt{\frac{2}{3}}|\phi\rangle$$

where  $|\psi\rangle$  and  $|\phi\rangle$  are orthonormal kets, i.e.

$$\langle\phi|\phi\rangle = \langle\psi|\psi\rangle = 1 ,$$

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle = 0 .$$

What is the value of the inner product  $\langle a|b\rangle$ ? What is the value of the inner product  $\langle b|a\rangle$ ?

[HINT: You will need to use Eq. (1.10).]

- (3) Show that the identity operator in Eq. (1.12) leaves an arbitrary operator, such as  $\hat{A}$  in Eq. (1.15), unchanged. This is, show

$$\hat{\mathbb{1}} \times \hat{A} = \hat{A} = \hat{A} \times \hat{\mathbb{1}} .$$

- (4) Verify that the representation of  $\hat{A}$  in Eq. (1.23) yields the eigenket equation

$$\hat{A}|\lambda_k\rangle = \lambda_k|\lambda_k\rangle$$

for orthonormal states  $\{|\lambda_n\rangle : n = 1, 2, \dots\}$ .

- (5) Verify that the matrix representation of the identity **operator** in Eq. (1.12) with respect to the  $\{|a_n\rangle\}$  orthonormal basis, is the identity **matrix**

$$\hat{\mathbb{1}} = \begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} .$$

- (6) Calculate the matrix representation of  $\hat{S}_z$  in the  $\{|-\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z\}$  basis.

- (7) Verify that Eq. (1.28) and Eq. (1.29) do indeed define the eigenkets of the spin operators  $\hat{S}_x$  and  $\hat{S}_y$ .

- (8) Use Eq. (1.27) to show that the matrix representations

of  $\hat{S}_x$  and  $\hat{S}_y$  in the  $\{|-\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z\}$  basis are

$$\begin{aligned} \mathbf{S}_x &= \frac{1}{2}\hbar \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \mathbf{S}_y &= \frac{1}{2}\hbar \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \end{aligned}$$

**(9)** Show that the three spin operators  $\hat{S}_x$ ,  $\hat{S}_y$  and  $\hat{S}_z$  are self-adjoint (or equivalently Hermitian).

**(10)** Consider a system in the state

$$|\psi\rangle = \frac{\sqrt{3}}{2}|\phi_1\rangle + \frac{1}{2}|\phi_2\rangle$$

where  $\langle\phi_1|\phi_2\rangle = 0$  and  $\langle\phi_1|\phi_1\rangle = \langle\phi_2|\phi_2\rangle = 1$ . Check that the state  $|\psi\rangle$  is normalised. What is the probability of finding the system in the state  $|\phi_1\rangle$  and, separately, in the state  $|\phi_2\rangle$ ?

**(11)** Let the spin state of an electron be given by

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\frac{1}{2}\rangle_z - i|-\frac{1}{2}\rangle_z\right).$$

**(a)** Show that this state gives the expectation values:

$$\langle\hat{S}_x\rangle = \langle\psi|\hat{S}_x|\psi\rangle = 0$$

$$\langle\hat{S}_y\rangle = \langle\psi|\hat{S}_y|\psi\rangle = -\frac{1}{2}\hbar$$

$$\langle\hat{S}_z\rangle = \langle\psi|\hat{S}_z|\psi\rangle = 0.$$

**(b)** Next show that the squares of the spin operators are all equal to  $\frac{1}{4}\hbar^2\mathbb{1}$ .

**(c)** Then calculate the variance in each of the spin operators where the variance  $\langle\Delta\hat{O}^2\rangle$  in  $\hat{O}$  is given by

$$\langle\Delta\hat{O}^2\rangle = \langle\hat{O}^2\rangle - \langle\hat{O}\rangle^2.$$

**(d)** Finally, identify the operator with a zero variance. What is relationship between  $|\psi\rangle$  and this operator?

# Advanced Quantum Theory

## 1 Dirac's bra-ket notation

### Solutions to Problems

**(1)** Using Eq. (1.2) and Eq. (1.1) we find the right hand side of Eq. (1.3) is

$$\begin{aligned}
 \int \psi_n^*(x)\Psi(x)dx &= \int \psi_n^*(x) \left[ \sum_{m=0}^{\infty} c_m \psi_m(x) \right] dx \\
 &= \sum_{m=0}^{\infty} c_m \left[ \int \psi_n^*(x) \psi_m(x) dx \right] \\
 &= \sum_{m=0}^{\infty} c_m [\delta_{n,m}] \\
 &= c_n .
 \end{aligned}$$

**(2)** The inner products are evaluated as follows:

$$\begin{aligned}
 \langle a|b\rangle &= \left[ \frac{1}{\sqrt{2}}(\langle\psi| + \langle\phi|) \right] \left( \frac{1}{\sqrt{3}}|\psi\rangle + i\sqrt{\frac{2}{3}}|\phi\rangle \right) \\
 &= \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{3}}\langle\psi|\psi\rangle + i\sqrt{\frac{2}{3}}\langle\psi|\phi\rangle + \frac{1}{\sqrt{3}}\langle\phi|\psi\rangle + i\sqrt{\frac{2}{3}}\langle\phi|\phi\rangle \right) \\
 &= \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{3}} + i\sqrt{\frac{2}{3}} \right) , \\
 \langle b|a\rangle &= \left( \frac{1}{\sqrt{3}}\langle\psi| - i\sqrt{\frac{2}{3}}\langle\phi| \right) \left[ \frac{1}{\sqrt{2}}(|\psi\rangle + |\phi\rangle) \right] \\
 &= \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{3}}\langle\psi|\psi\rangle + \frac{1}{\sqrt{3}}\langle\psi|\phi\rangle - i\sqrt{\frac{2}{3}}\langle\phi|\psi\rangle - i\sqrt{\frac{2}{3}}\langle\phi|\phi\rangle \right) \\
 &= \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{3}} - i\sqrt{\frac{2}{3}} \right) .
 \end{aligned}$$

(3) Using the expansions of  $\hat{\mathbf{1}}$  and  $\hat{A}$  we find

$$\begin{aligned}
 \hat{\mathbf{1}} \times \hat{A} &= \left( \sum_n |a_n\rangle\langle a_n| \right) \left( \sum_{i,j} A_{i,j} |a_i\rangle\langle a_j| \right) \\
 &= \sum_n \sum_{i,j} A_{i,j} |a_n\rangle \left( \langle a_n | a_i \rangle \right) \langle a_j | \\
 &= \sum_n \sum_{i,j} A_{i,j} |a_n\rangle \left( \delta_{n,i} \right) \langle a_j | \\
 &= \sum_{i,j} A_{i,j} |a_i\rangle\langle a_j| \\
 &= \hat{A} ,
 \end{aligned}$$

and

$$\begin{aligned}
 \hat{A} \times \hat{\mathbf{1}} &= \left( \sum_{i,j} A_{i,j} |a_i\rangle\langle a_j| \right) \left( \sum_n |a_n\rangle\langle a_n| \right) \\
 &= \sum_{i,j} \sum_n A_{i,j} |a_i\rangle \left( \langle a_j | a_n \rangle \right) \langle a_n | \\
 &= \sum_{i,j} A_{i,j} |a_i\rangle\langle a_j| \\
 &= \hat{A} .
 \end{aligned}$$

(4) Using the matrix representation of  $\hat{A}$  in Eq. (1.23) we find

$$\begin{aligned}
 \hat{A}|\lambda_k\rangle &= \left( \sum_n \lambda_n |\lambda_n\rangle\langle \lambda_n| \right) |\lambda_k\rangle \\
 &= \sum_n \lambda_n |\lambda_n\rangle \left( \langle \lambda_n | \lambda_k \rangle \right) \\
 &= \sum_n \lambda_n |\lambda_n\rangle \left( \delta_{n,k} \right) \\
 &= \lambda_k |\lambda_k\rangle
 \end{aligned}$$

which is the eigenvalue equation.

(5) According to Eq. (1.22) the matrix elements of  $\hat{\mathbf{1}}$  in the  $\{|a_n\rangle\}$  basis are given by

$$\mathbf{1}_{n,m} = \langle a_n | \hat{\mathbf{1}} | a_m \rangle .$$

Noting that  $\hat{\mathbf{1}}|a_n\rangle = |a_n\rangle$  and  $\langle a_n | a_m \rangle = \delta_{n,m}$  we find

$$\mathbf{1}_{n,m} = \delta_{n,m}$$

and hence

$$\mathbb{1} = \begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} .$$

(6) We need to evaluate the matrix elements according to Eq. (1.22) and Eq. (1.16):

$$\begin{aligned} {}_z\langle \frac{1}{2} | \hat{S}_z | \frac{1}{2} \rangle_z &= {}_z\langle \frac{1}{2} | \left( \hat{S}_z | \frac{1}{2} \rangle_z \right) = {}_z\langle \frac{1}{2} | \left( \frac{1}{2}\hbar | \frac{1}{2} \rangle_z \right) \\ &= \frac{1}{2}\hbar {}_z\langle \frac{1}{2} | \frac{1}{2} \rangle_z = \frac{1}{2}\hbar \\ {}_z\langle \frac{1}{2} | \hat{S}_z | -\frac{1}{2} \rangle_z &= {}_z\langle \frac{1}{2} | \left( \hat{S}_z | -\frac{1}{2} \rangle_z \right) = {}_z\langle \frac{1}{2} | \left( -\frac{1}{2}\hbar | -\frac{1}{2} \rangle_z \right) \\ &= -\frac{1}{2}\hbar {}_z\langle \frac{1}{2} | -\frac{1}{2} \rangle_z = 0 \\ {}_z\langle -\frac{1}{2} | \hat{S}_z | \frac{1}{2} \rangle_z &= {}_z\langle -\frac{1}{2} | \left( \hat{S}_z | \frac{1}{2} \rangle_z \right) = {}_z\langle -\frac{1}{2} | \left( \frac{1}{2}\hbar | \frac{1}{2} \rangle_z \right) \\ &= \frac{1}{2}\hbar {}_z\langle -\frac{1}{2} | \frac{1}{2} \rangle_z = 0 \\ {}_z\langle -\frac{1}{2} | \hat{S}_z | -\frac{1}{2} \rangle_z &= {}_z\langle -\frac{1}{2} | \left( \hat{S}_z | -\frac{1}{2} \rangle_z \right) = {}_z\langle -\frac{1}{2} | \left( -\frac{1}{2}\hbar | -\frac{1}{2} \rangle_z \right) \\ &= -\frac{1}{2}\hbar {}_z\langle -\frac{1}{2} | -\frac{1}{2} \rangle_z = -\frac{1}{2}\hbar . \end{aligned}$$

Hence the matrix representation is

$$\mathbf{S}_z = \begin{bmatrix} \frac{1}{2}\hbar & 0 \\ 0 & -\frac{1}{2}\hbar \end{bmatrix} = \frac{1}{2}\hbar \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

(7) From Eq. (1.27) and Eq. (1.28) we have

$$\begin{aligned} \hat{S}_x | \frac{1}{2} \rangle_x &= \left( \frac{1}{2}\hbar | \frac{1}{2} \rangle_z \langle -\frac{1}{2} | + \frac{1}{2}\hbar | -\frac{1}{2} \rangle_z \langle \frac{1}{2} | \right) \frac{1}{\sqrt{2}} \left( | \frac{1}{2} \rangle_z + | -\frac{1}{2} \rangle_z \right) \\ &= \frac{1}{\sqrt{2}} \left( \frac{1}{2}\hbar | \frac{1}{2} \rangle_z + \frac{1}{2}\hbar | -\frac{1}{2} \rangle_z \right) = \frac{1}{2}\hbar | \frac{1}{2} \rangle_x , \\ \hat{S}_x | -\frac{1}{2} \rangle_x &= \left( \frac{1}{2}\hbar | \frac{1}{2} \rangle_z \langle -\frac{1}{2} | + \frac{1}{2}\hbar | -\frac{1}{2} \rangle_z \langle \frac{1}{2} | \right) \frac{1}{\sqrt{2}} \left( | \frac{1}{2} \rangle_z - | -\frac{1}{2} \rangle_z \right) \\ &= \frac{1}{\sqrt{2}} \left( -\frac{1}{2}\hbar | \frac{1}{2} \rangle_z + \frac{1}{2}\hbar | -\frac{1}{2} \rangle_z \right) = -\frac{1}{2}\hbar | -\frac{1}{2} \rangle_x , \end{aligned}$$

and from Eq. (1.29) we have

$$\begin{aligned}\hat{S}_y|\frac{1}{2}\rangle_y &= \left(-i\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| + i\frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)\frac{1}{\sqrt{2}}\left(|\frac{1}{2}\rangle_z + i|-\frac{1}{2}\rangle_z\right) \\ &= \frac{1}{\sqrt{2}}\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_z + i\frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\right) = \frac{1}{2}\hbar|\frac{1}{2}\rangle_y , \\ \hat{S}_y|-\frac{1}{2}\rangle_y &= \left(-i\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| + i\frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)\frac{1}{\sqrt{2}}\left(|\frac{1}{2}\rangle_z - i|-\frac{1}{2}\rangle_z\right) \\ &= \frac{1}{\sqrt{2}}\left(-\frac{1}{2}\hbar|\frac{1}{2}\rangle_z + i\frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\right) = -\frac{1}{2}\hbar|\frac{1}{2}\rangle_y .\end{aligned}$$

(8) From Eq. (1.27) we have

$$\begin{aligned}\langle\frac{1}{2}|\hat{S}_x|\frac{1}{2}\rangle_z &= \langle\frac{1}{2}|\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)|\frac{1}{2}\rangle_z \\ &= \langle\frac{1}{2}|\left(\frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\right) = 0 , \\ \langle\frac{1}{2}|\hat{S}_x|-\frac{1}{2}\rangle_z &= \langle\frac{1}{2}|\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)|-\frac{1}{2}\rangle_z \\ &= \langle\frac{1}{2}|\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\right) = \frac{1}{2}\hbar , \\ \langle-\frac{1}{2}|\hat{S}_x|\frac{1}{2}\rangle_z &= \langle-\frac{1}{2}|\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)|\frac{1}{2}\rangle_z \\ &= \langle-\frac{1}{2}|\left(\frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\right) = \frac{1}{2}\hbar , \\ \langle-\frac{1}{2}|\hat{S}_x|-\frac{1}{2}\rangle_z &= \langle-\frac{1}{2}|\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)|-\frac{1}{2}\rangle_z \\ &= \langle-\frac{1}{2}|\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\right) = 0 .\end{aligned}$$

and so from Eq. (1.22) we have

$$\mathbf{S}_x = \begin{bmatrix} 0 & \frac{1}{2}\hbar \\ \frac{1}{2}\hbar & 0 \end{bmatrix} = \frac{1}{2}\hbar \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} .$$

For  $\mathbf{S}_y$  the calculations are similar, the only difference is the value of the constants,

$$\begin{aligned}_z\langle\frac{1}{2}|\hat{S}_y|\frac{1}{2}\rangle_z &= 0 , \\ _z\langle\frac{1}{2}|\hat{S}_y|-\frac{1}{2}\rangle_z &= -i\frac{1}{2}\hbar , \\ _z\langle-\frac{1}{2}|\hat{S}_y|\frac{1}{2}\rangle_z &= i\frac{1}{2}\hbar , \\ _z\langle-\frac{1}{2}|\hat{S}_y|-\frac{1}{2}\rangle_z &= 0 .\end{aligned}$$

and so

$$\mathbf{S}_y = \begin{bmatrix} 0 & -i\frac{1}{2}\hbar \\ -\frac{1}{2}\hbar & 0 \end{bmatrix} = \frac{1}{2}\hbar \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} .$$

**(9)** We need to show that each operator is equal to its Hermitian conjugate. First note that

$$\begin{aligned}
 i^\dagger &= i^* = -i , \\
 |\pm\frac{1}{2}\rangle_z^\dagger &= \langle\pm\frac{1}{2}| , \\
 \langle\pm\frac{1}{2}|^\dagger &= |\pm\frac{1}{2}\rangle_z , \\
 \left(|\frac{1}{2}\rangle_z\langle-\frac{1}{2}|\right)^\dagger &= |-\frac{1}{2}\rangle_z\langle\frac{1}{2}| , \\
 \left(|-\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)^\dagger &= |\frac{1}{2}\rangle_z\langle-\frac{1}{2}| , \\
 \left(|\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)^\dagger &= |\frac{1}{2}\rangle_z\langle\frac{1}{2}| , \quad (\text{unchanged}) \\
 \left(|-\frac{1}{2}\rangle_z\langle-\frac{1}{2}|\right)^\dagger &= |-\frac{1}{2}\rangle_z\langle-\frac{1}{2}| , \quad (\text{unchanged})
 \end{aligned}$$

and so

$$\begin{aligned}
 \left(\hat{S}_x\right)^\dagger &= \left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)^\dagger \\
 &= \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}| + \frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| \\
 &= \hat{S}_x , \\
 \left(\hat{S}_y\right)^\dagger &= \left(-i\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| + i\frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}|\right)^\dagger \\
 &= i\frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle\frac{1}{2}| - i\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle-\frac{1}{2}| \\
 &= \hat{S}_y , \\
 \left(\hat{S}_z\right)^\dagger &= \left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle\frac{1}{2}| - \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle-\frac{1}{2}|\right)^\dagger \\
 &= \frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle\frac{1}{2}| - \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle-\frac{1}{2}| \\
 &= \hat{S}_z .
 \end{aligned}$$

**(10)** To check the norm we take the inner product

$$\begin{aligned}
 \langle\psi|\psi\rangle &= \left(\frac{\sqrt{3}}{2}\langle\phi_1| + \frac{1}{2}\langle\phi_2|\right)\left(\frac{\sqrt{3}}{2}|\phi_1\rangle + \frac{1}{2}|\phi_2\rangle\right) \\
 &= \left(\frac{\sqrt{3}}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = 1 .
 \end{aligned}$$

Hence the state is normalised.

The probability of finding the system in  $|\phi_1\rangle$  is given by

$$\begin{aligned} P(|\phi_1\rangle) &= |\langle\phi_1|\psi\rangle|^2 = \left| \langle\phi_1| \left( \frac{\sqrt{3}}{2}|\phi_1\rangle + \frac{1}{2}|\phi_2\rangle \right) \right|^2 \\ &= \left| \frac{\sqrt{3}}{2} \right|^2 = \frac{3}{4}. \end{aligned}$$

Similarly, the probability of finding the system in  $|\phi_2\rangle$  is given by

$$\begin{aligned} P(|\phi_2\rangle) &= |\langle\phi_2|\psi\rangle|^2 = \left| \langle\phi_2| \left( \frac{\sqrt{3}}{2}|\phi_1\rangle + \frac{1}{2}|\phi_2\rangle \right) \right|^2 \\ &= \left| \frac{1}{2} \right|^2 = \frac{1}{4}. \end{aligned}$$

**(11) (a)** From Eq. (1.27) we find

$$\begin{aligned} \langle\psi|\hat{S}_x|\psi\rangle &= \frac{1}{\sqrt{2}} \left( z\langle\frac{1}{2}| + i_z\langle-\frac{1}{2}| \right) \left( \frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| \right) \\ &\quad \times \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z - i|-\frac{1}{2}\rangle_z \right) \\ &= \frac{1}{\sqrt{2}} \left( z\langle\frac{1}{2}| + i_z\langle-\frac{1}{2}| \right) \left( -i\frac{1}{2\sqrt{2}}\hbar|\frac{1}{2}\rangle_z + \frac{1}{2\sqrt{2}}\hbar|-\frac{1}{2}\rangle_z \right) \\ &= -i\frac{1}{4} + i\frac{1}{4} = 0, \end{aligned}$$

$$\begin{aligned} \langle\psi|\hat{S}_y|\psi\rangle &= \frac{1}{\sqrt{2}} \left( z\langle\frac{1}{2}| + i_z\langle-\frac{1}{2}| \right) \left( -i\frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + i\frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| \right) \\ &\quad \times \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z - i|-\frac{1}{2}\rangle_z \right) \\ &= \frac{1}{\sqrt{2}} \left( z\langle\frac{1}{2}| + i_z\langle-\frac{1}{2}| \right) \left( -\frac{1}{2\sqrt{2}}\hbar|\frac{1}{2}\rangle_z + i\frac{1}{2\sqrt{2}}\hbar|-\frac{1}{2}\rangle_z \right) \\ &= -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2}, \end{aligned}$$

$$\begin{aligned} \langle\psi|\hat{S}_z|\psi\rangle &= \frac{1}{\sqrt{2}} \left( z\langle\frac{1}{2}| + i_z\langle-\frac{1}{2}| \right) \left( \frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| - \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| \right) \\ &\quad \times \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z - i|-\frac{1}{2}\rangle_z \right) \\ &= \frac{1}{\sqrt{2}} \left( z\langle\frac{1}{2}| + i_z\langle-\frac{1}{2}| \right) \left( \frac{1}{2\sqrt{2}}\hbar|\frac{1}{2}\rangle_z + i\frac{1}{2\sqrt{2}}\hbar|-\frac{1}{2}\rangle_z \right) \\ &= \frac{1}{4} - \frac{1}{4} = 0. \end{aligned}$$

**(b)** Again using Eq. (1.27) we get

$$\begin{aligned}
 (\hat{S}_x)^2 &= \left( \frac{1}{2}\hbar\left|\frac{1}{2}\right\rangle_z\left\langle -\frac{1}{2}\right| + \frac{1}{2}\hbar\left|-\frac{1}{2}\right\rangle_z\left\langle \frac{1}{2}\right| \right) \\
 &\quad \times \left( \frac{1}{2}\hbar\left|\frac{1}{2}\right\rangle_z\left\langle -\frac{1}{2}\right| + \frac{1}{2}\hbar\left|-\frac{1}{2}\right\rangle_z\left\langle \frac{1}{2}\right| \right) \\
 &= \frac{1}{4}\hbar^2\left|\frac{1}{2}\right\rangle_z\left\langle \frac{1}{2}\right| + \frac{1}{4}\hbar^2\left|-\frac{1}{2}\right\rangle_z\left\langle -\frac{1}{2}\right| = \frac{1}{4}\hbar^2\hat{\mathbf{1}} , \\
 (\hat{S}_y)^2 &= \left( -i\frac{1}{2}\hbar\left|\frac{1}{2}\right\rangle_z\left\langle -\frac{1}{2}\right| + i\frac{1}{2}\hbar\left|-\frac{1}{2}\right\rangle_z\left\langle \frac{1}{2}\right| \right)^2 \\
 &= \frac{1}{4}\hbar^2\left|\frac{1}{2}\right\rangle_z\left\langle \frac{1}{2}\right| + \frac{1}{4}\hbar^2\left|-\frac{1}{2}\right\rangle_z\left\langle -\frac{1}{2}\right| = \frac{1}{4}\hbar^2\hat{\mathbf{1}} , \\
 (\hat{S}_z)^2 &= \left( \frac{1}{2}\hbar\left|\frac{1}{2}\right\rangle_z\left\langle \frac{1}{2}\right| - \frac{1}{2}\hbar\left|-\frac{1}{2}\right\rangle_z\left\langle -\frac{1}{2}\right| \right)^2 \\
 &= \frac{1}{4}\hbar^2\left|\frac{1}{2}\right\rangle_z\left\langle \frac{1}{2}\right| + \frac{1}{4}\hbar^2\left|-\frac{1}{2}\right\rangle_z\left\langle -\frac{1}{2}\right| = \frac{1}{4}\hbar^2\hat{\mathbf{1}} .
 \end{aligned}$$

**(c)** Note that  $\langle \hat{\mathbf{1}} \rangle = 1$  for every state and so

$$\begin{aligned}
 \langle \Delta \hat{S}_x^2 \rangle &= \langle (\hat{S}_x)^2 \rangle - \langle \hat{S}_x \rangle^2 = \langle \frac{1}{4}\hbar^2\hat{\mathbf{1}} \rangle - \langle \hat{S}_x \rangle^2 = \frac{1}{4}\hbar^2 - 0 \\
 &= \frac{1}{4}\hbar^2 , \\
 \langle \Delta \hat{S}_y^2 \rangle &= \langle (\hat{S}_y)^2 \rangle - \langle \hat{S}_x \rangle^2 = \langle \frac{1}{4}\hbar^2\hat{\mathbf{1}} \rangle - \langle \hat{S}_x \rangle^2 = \frac{1}{4}\hbar^2 - (-\frac{1}{2}\hbar)^2 \\
 &= 0 \\
 \langle \Delta \hat{S}_z^2 \rangle &= \langle (\hat{S}_z)^2 \rangle - \langle \hat{S}_x \rangle^2 = \langle \frac{1}{4}\hbar^2\hat{\mathbf{1}} \rangle - \langle \hat{S}_x \rangle^2 = \frac{1}{4}\hbar^2 - 0 \\
 &= \frac{1}{4}\hbar^2 .
 \end{aligned}$$

**(d)** The operator that has zero variance is  $\hat{S}_y$ . This means that measuring  $\hat{S}_y$  incurs no uncertainty. The only way for this to happen is if  $|\psi\rangle$  is an eigenstate of  $\hat{S}_y$ .

## 2. Quantum States

### a. Position and Momentum representations

So far we have mainly considered systems for which the orthogonal bases are discrete, i.e. they can be enumerated with an integer index such as  $\{|\varphi_n\rangle : n = 1, 2, \dots\}$ . We now briefly examine a system with a continuous orthogonal basis. The system consists of a simple particle which does not have **internal degrees of freedom** such as spin or rotation (i.e. we assume it to be a spherically-symmetric spin-0 particle). The particle could be free, in which case the Hamiltonian is just the kinetic energy,  $\hat{\mathcal{H}} = \hat{P}^2/(2m)$ , or bound by a potential, in which case the Hamiltonian is  $\hat{\mathcal{H}} = \hat{P}^2/(2m) + \hat{U}$ . An example of the latter is a harmonic oscillator with spring constant  $k$  for which  $\hat{\mathcal{H}} = \hat{P}^2/(2m) + k\hat{X}^2/2$ . In these expressions  $\hat{X}$  represents the position observable and  $\hat{P}$  the momentum observable.

#### Position eigenbasis

The ket representing the eigenstate of the position operator is  $|x\rangle$ . These states are orthonormal in the **Dirac-delta** sense:

insert  
1

$$\langle x|x'\rangle = \delta(x - x')$$

and form a **resolution of the identity operator** for the Hilbert space of states:

$$\int_{-\infty}^{\infty} |x\rangle\langle x| dx = \hat{\mathbb{1}} . \quad (2.1)$$

The Dirac delta “function”  $\delta(x)$  and its properties are discussed in **Appendix A** on page 17.

#### Technical points

- Strictly speaking, we should represent these kets with a **label**, such as  $|x\rangle_X$ , to indicate that they are eigenkets of the position operator  $\hat{X}$ , and to distinguish them from the eigenkets of momentum, say  $|p\rangle_P$ . However this notation can be cumbersome. Instead we will omit the labels and use the **variables**  $x$  and  $p$  in  $|x\rangle$  and  $|p\rangle$  to indicate that these are position and momentum eigenkets, respectively.
- The state space is actually the **rigged Hilbert space** and the kets  $|x\rangle$  represent **distributions**. However we will overlook these mathematical details and treat  $|x\rangle$  as if it were a conventional vector.

The **position eigenvalue equation** is

$$\hat{X}|x\rangle = x|x\rangle \quad (2.2)$$

Thus using the resolution of the identity and the orthogonality relation we find

$$\begin{aligned} \hat{X} &= \hat{X} \times \hat{\mathbf{1}} = \hat{X} \left( \int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) = \int_{-\infty}^{\infty} \hat{X}|x\rangle \langle x| dx \\ &= \int_{-\infty}^{\infty} x|x\rangle \langle x| dx . \end{aligned} \quad (2.3)$$

Notice that  $\hat{X}$  is **self adjoint** (or loosely speaking, Hermitian), as

$$\begin{aligned} \hat{X}^\dagger &= \left( \int_{-\infty}^{\infty} x|x\rangle \langle x| dx \right)^\dagger = \int_{-\infty}^{\infty} \left( x|x\rangle \langle x| \right)^\dagger dx \\ &= \int_{-\infty}^{\infty} x|x\rangle \langle x| dx = \hat{X} . \end{aligned}$$

### Problems

- Verify that the right-hand side of Eq. (2.3) gives the eigenvalue equation  $\hat{X}|x'\rangle = x'|x'\rangle$ .
- Use Eq. (2.3) to show that the  $n$ -th power of  $\hat{X}$  is given by

$$\hat{X}^n = \int_{-\infty}^{\infty} x^n|x\rangle \langle x| dx . \quad (2.4)$$

Let  $|\psi\rangle$  be a general state. The expansion of  $|\psi\rangle$  in terms of the **position basis** is, using Eq. (2.1),

$$\begin{aligned} |\psi\rangle &= \hat{\mathbf{1}}|\psi\rangle = \left( \int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) |\psi\rangle = \int_{-\infty}^{\infty} |x\rangle \langle x|\psi\rangle dx \\ &= \int_{-\infty}^{\infty} \psi(x)|x\rangle dx \end{aligned} \quad (2.5)$$

where

### Position representation

insert  
2

$$\psi(x) = \langle x|\psi\rangle \quad (2.6)$$

is the wave function in the **position representation**.

The moments of  $\hat{X}$  are given by

$$\begin{aligned}\langle \hat{X}^n \rangle &= \langle \psi | \left( \int_{-\infty}^{\infty} x^n |x\rangle \langle x| dx \right) | \psi \rangle \quad \text{from Eq. (2.4)} \\ &= \int_{-\infty}^{\infty} \langle \psi | x \rangle x^n \langle x | \psi \rangle dx = \int_{-\infty}^{\infty} \psi^*(x) x^n \psi(x) dx \quad \text{from Eq. (2.6)} \\ &= \int_{-\infty}^{\infty} P(x) x^n dx\end{aligned}$$

where we have defined the

### Position probability density

$$P(x) = \psi^*(x) \psi(x) = |\psi(x)|^2$$

$P(x)dx$  is the probability of finding the particle between  $x$  and  $x + dx$ , and so the unit of  $P(x)$  itself is 1/length. This is why  $P(x)$  is called a **density**.

The result  $\langle \hat{X}^n \rangle = \int_{-\infty}^{\infty} \psi^*(x) x^n \psi(x) dx$  should be familiar from wave mechanics.

### Momentum operator

In **Appendix B** on page 23 we show that the momentum operator takes the form of a differential operator when acting on position eigenkets, i.e. from Eq. (2.30):

insert  
3

$$\hat{P}|x\rangle = i\hbar \frac{d}{dx}|x\rangle . \quad (2.7)$$

While expressed as a derivative of a ket, this result leads to expressions involving the derivative of a wave function. For example consider

$$\begin{aligned}\langle \psi | \hat{P} &= \langle \psi | \hat{P} \left( \int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) = \langle \psi | \left( \int_{-\infty}^{\infty} \hat{P} |x\rangle \langle x| dx \right) \\ &= \langle \psi | \left[ \int_{-\infty}^{\infty} \left( i\hbar \frac{d}{dx} |x\rangle \right) \langle x| dx \right] = \int_{-\infty}^{\infty} i\hbar \left( \langle \psi | \frac{d}{dx} |x\rangle \right) \langle x| dx \\ &= \int_{-\infty}^{\infty} i\hbar \left( \frac{d}{dx} \langle \psi | x \rangle \right) \langle x| dx \\ &= \int_{-\infty}^{\infty} i\hbar \left( \frac{d}{dx} \psi^*(x) \right) \langle x| dx\end{aligned}$$

and so on. Also, from Eq. (2.31) in the Appendix, the moments

of  $\hat{P}$  are

$$\langle \hat{P}^n \rangle = \int_{-\infty}^{\infty} \psi^*(x) \left[ -i\hbar \frac{d}{dx} \right]^n \psi(x) dx$$

which should be familiar from wave mechanics.

### Momentum eigenbasis

The eigenvalue equation for the momentum operator is

$$\hat{P}|p\rangle = p|p\rangle \quad (2.8)$$

where the eigenkets form an orthonormal basis in the sense that

$$\langle p|p' \rangle = \delta(p - p')$$

and

$$\hat{\mathbf{1}} = \int_{-\infty}^{\infty} |p\rangle \langle p| dp . \quad (2.9)$$

The eigenket expansion of  $\hat{P}$  is

insert  
4

$$\hat{P} = \int_{-\infty}^{\infty} p|p\rangle \langle p| dp$$

which clearly shows that  $\hat{P}$  is self adjoint (i.e.  $\hat{P}^\dagger = \hat{P}$ ).

An arbitrary state  $|\psi\rangle$  can be expanded in the **momentum basis** as

$$\begin{aligned} |\psi\rangle &= \hat{\mathbf{1}}|\psi\rangle \\ &= \left( \int_{-\infty}^{\infty} |p\rangle \langle p| dp \right) |\psi\rangle \quad \text{from Eq. (2.9)} \\ &= \int_{-\infty}^{\infty} \tilde{\psi}(p)|p\rangle dp \end{aligned}$$

where

#### Momentum Representation

$$\tilde{\psi}(p) = \langle p|\psi\rangle$$

is the wave function in the **momentum representation**.

### Relationship between representations

In **Appendix B** we also show, in Eq. (2.35), that the relationship between the representations is given by

$$\langle x|p\rangle = \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} . \quad (2.10)$$

### Problems

- Show that the relationship between the position and momentum representations  $\psi(x) = \langle x|\psi\rangle$  and  $\tilde{\psi}(p) = \langle p|\psi\rangle$  is the Fourier transform

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \psi(x) dx .$$

- Show that the momentum eigenkets can be written in the position basis as

$$|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{ipx/\hbar} |x\rangle dx \quad (2.11)$$

and conversely that

$$|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} |p\rangle dp . \quad (2.12)$$

- Show that the position operator  $\hat{X}$  in the momentum basis is given by

$$\hat{X}|p\rangle = -i\hbar \frac{d}{dp} |p\rangle . \quad (2.13)$$

- **(Harder)** Show that the commutator between  $\hat{X}$  and  $\hat{P}$  is given by

$$[\hat{X}, \hat{P}] = i\hbar . \quad (2.14)$$

## b. Pure states, mixed states and density operators

### Pure states

Up to now we have only treated quantum states in the form of  $|\psi\rangle$ . These are called **pure states**. When a physical system is in a pure state then we know as much as we can about the related physical system. That is, we possess **maximal possible knowledge** of it.

### Comment

One of the fascinating things about quantum mechanics is that we can possess maximal knowledge and yet still not be able to predict everything that a physical system will do. (This is certainly not the case in Newtonian mechanics.) In other words, for instance, we can *know* that an electron is in the pure state  $|\psi\rangle = |\frac{1}{2}\rangle_z$  and yet, when we measure the electron's spin in the  $x$  direction, we will *randomly* get the results  $+\frac{1}{2}\hbar$  and  $-\frac{1}{2}\hbar$  with equal probabilities. This is an example of intrinsic randomness. Though such behaviour is an accepted part of quantum mechanics, Albert Einstein was sceptical about it, as summed up by his well-known

quote “God does not play dice.”

It is useful at this point to write expectation values in a different way. Consider the expectation value of an arbitrary operator  $\hat{A}$  for arbitrary pure state  $|\psi\rangle$ . Let  $\{|\varphi_n\rangle : n = 1, 2, \dots\}$  be an orthonormal basis for the Hilbert space which gives the resolution of the identity  $\hat{\mathbb{1}} = \sum_n |\varphi_n\rangle\langle\varphi_n|$ . We can write the expectation value as

$$\begin{aligned}\langle\psi|\hat{A}|\psi\rangle &= \langle\psi|\hat{\mathbb{1}} \times \hat{A}|\psi\rangle = \langle\psi|\left(\sum_n |\varphi_n\rangle\langle\varphi_n|\right)\hat{A}|\psi\rangle \\ &= \sum_n \langle\psi|\varphi_n\rangle\langle\varphi_n|\hat{A}|\psi\rangle = \sum_n \langle\varphi_n|\hat{A}|\psi\rangle\langle\psi|\varphi_n\rangle \\ &= \sum_n \langle\varphi_n|\left(\hat{A}|\psi\rangle\langle\psi|\right)|\varphi_n\rangle \\ &= \textcolor{red}{\text{Tr}(\hat{A}|\psi\rangle\langle\psi|)} .\end{aligned}$$

insert  
5

### Expectation value as trace

$$\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle = \text{Tr}(\hat{A}|\psi\rangle\langle\psi|) \quad (2.15)$$

In the second line we have made use of the fact that the terms  $\langle\psi|\varphi_n\rangle$  and  $\langle\varphi_n|\hat{A}|\psi\rangle$  are just complex numbers which allows us to interchange their order. We have also introduced the **trace** operation  $\text{Tr}$  which is defined for arbitrary operator  $\hat{O}$  and arbitrary orthonormal basis  $\{|\varphi_n\rangle : n = 1, 2, \dots\}$  as follows

$$\text{Tr}(\hat{O}) \equiv \sum_n \langle\varphi_n|\hat{O}|\varphi_n\rangle .$$

## Properties of the trace operation

- (a) The trace can be evaluated in **any orthonormal basis**.
- (b) The trace of an operator is simply the sum of the **diagonal elements** of the corresponding matrix in any basis.
- (c) For a self-adjoint operator  $\hat{O}$  which has an orthonormal eigenstate expansion  $\hat{O} = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$ , the trace of  $\hat{O}$  is simply  $\text{Tr}(\hat{O}) = \sum_i \lambda_i$ . (To see this use the eigenbasis for the trace).
- (d) The **trace of a sum is the sum of the traces**, i.e.

$$\text{Tr}(\hat{A} + \hat{B}) = \text{Tr}(\hat{A}) + \text{Tr}(\hat{B}) .$$

- (e) An important feature about the trace operation is that we can shift cyclically the operators without changing its value. That is, for any two operators  $\hat{A}$  and  $\hat{B}$

$$\text{Tr}(\hat{A}\hat{B}) = \text{Tr}(\hat{B}\hat{A}) .$$

Furthermore, for any three operators  $\hat{A}$ ,  $\hat{B}$  and  $\hat{C}$ ,

$$\text{Tr}(\hat{A}\hat{B}\hat{C}) = \text{Tr}(\hat{C}\hat{A}\hat{B}) = \text{Tr}(\hat{B}\hat{C}\hat{A}) .$$

Another way of saying this is that the **trace operation is invariant under cyclical permutations** of its operator arguments.

## Mixed states

There are situations where we know less than everything that is possible to know, that is, we have *incomplete knowledge* of the physical system. More formally, we do not know for certain that the system is in a pure state, but rather only that it has a *probability* of being in each of a *set of pure states*. In such situations we call the state of the system a **mixed state**.

For example, consider the Stern-Gerlach setup in Section 1. We know that just before hitting the screen, the atoms in the upper beam are in the state  $|\frac{1}{2}\rangle_z$  and in the lower beam they are in  $|-\frac{1}{2}\rangle_z$ . These are pure states. But in the beam before it enters the magnet we know that there is a 50% chance that an atom is in the state  $|\frac{1}{2}\rangle_z$  and a 50% chance that it is in  $|-\frac{1}{2}\rangle_z$ . It is in a *mixed state* in this part of the apparatus.

## Expectation values for mixed states

Imagine we wish to calculate the expectation value of the operator  $\hat{A}$  for a system which is in the pure state  $|\psi_i\rangle$  with probability  $p_i$  for  $i = 1, 2, \dots, N$ . If the system was actually in the state  $|\psi_i\rangle$  then the expectation value would be simply  $\langle\psi_i|\hat{A}|\psi_i\rangle = \text{Tr}(\hat{A}|\psi_i\rangle\langle\psi_i|)$ . But given we do not know which particular state of the set  $\{|\psi_i\rangle : i = 1, 2, \dots, N\}$  the system is actually in, we need to take the average of  $\langle\psi_i|\hat{A}|\psi_i\rangle$  over the set. Thus our best estimate of the expectation value of  $\hat{A}$  is

$$\begin{aligned}\langle\hat{A}\rangle &= \sum_{i=1}^N p_i \langle\psi_i|\hat{A}|\psi_i\rangle \\ &= \sum_{i=1}^N p_i \text{Tr}\left(\hat{A}|\psi_i\rangle\langle\psi_i|\right) \quad \text{from Eq. (2.15)} \\ &= \text{Tr}\left(\hat{A} \sum_{i=1}^N p_i |\psi_i\rangle\langle\psi_i|\right)\end{aligned}$$

insert  
6

$$= \text{Tr}\left(\hat{A}\hat{\rho}\right)$$

where

$$\hat{\rho} \equiv \sum_i p_i |\psi_i\rangle\langle\psi_i| .$$

### Density operator $\hat{\rho}$

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i| . \quad (2.16)$$

### Expectation value

$$\langle\hat{A}\rangle = \text{Tr}\left(\hat{A}\hat{\rho}\right) . \quad (2.17)$$

Here  $p_i$  is the probability that the system has been prepared in the state  $|\psi_i\rangle$ . That is,  $\hat{\rho}$  represents an **average** of pure states each with the probabilistic weight  $p_i$ . The density operator Eq. (2.16) allows us to represent the general state of a system, be it pure or mixed, and to calculate expectation values according to the simple formula in Eq. (2.17).

## KEY POINTS

- In this course we assume implicitly that all kets and bras are normalised. This is true, in particular, for the pure states  $\{|\psi_i\rangle\}$ .

- Note  $\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$  is the **average state**. Do **not** use

$$\sum_i p_i |\psi_i\rangle \quad (\text{wrong!})$$

which is a pure, superposition state.

## Probability interpretation for mixed states

If the system is in pure state  $|\psi_i\rangle$ , the probability  $P(\phi|\psi_i)$  of finding the system in state  $|\phi\rangle$  is

insert  
7

$$P(\phi|\psi_i) = |\langle\phi|\psi_i\rangle|^2 .$$

However, if the system is prepared in state  $|\psi_i\rangle$  with probability  $p_i$ , then the total (or average) probability of finding  $|\phi\rangle$  irrespective of the preparation is

$$\begin{aligned} P(\phi|\hat{\rho}) &= \sum_i p_i P(\phi|\psi_i) \quad [\text{from Bayes theorem: } p(k) = \sum_i p(i)p(k|i)] \\ &= \sum_i p_i |\langle\phi|\psi_i\rangle|^2 \\ &= \sum_i p_i \langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle \\ &= \langle\phi| \left( \sum_i p_i |\psi_i\rangle\langle\psi_i| \right) |\phi\rangle \\ &= \langle\phi|\hat{\rho}|\phi\rangle \quad \text{from Eq. (2.16)} \end{aligned}$$

where  $\hat{\rho}$  is the density operator for the system.

## Probability interpretation for mixed states

$$P(\phi|\hat{\rho}) = \langle\phi|\hat{\rho}|\phi\rangle \tag{2.18}$$

- Example (1).** Consider a spin-1 particle which we know, with 25% probability, is in each of the states  $|1\rangle_z$  and  $|0\rangle_z$  and, with 50% probability, is in the state  $|-1\rangle_z$ . Find the density operator  $\hat{\rho}$  for the average state of the particle and its corresponding matrix representation  $\rho$  in the  $\{|-1\rangle_z, |0\rangle_z, |1\rangle_z\}$  basis.

The density operator is given by

$$\hat{\rho} = \frac{1}{4}|1\rangle_{zz}\langle 1| + \frac{1}{4}|0\rangle_{zz}\langle 0| + \frac{1}{2}| -1\rangle_{zz}\langle -1| . \quad (2.19)$$

We find the matrix elements of  $\hat{\rho}$  in the  $\hat{S}_z$  basis directly. Let  $\rho$  be the matrix representing  $\hat{\rho}$ , thus

$$\begin{aligned}\rho_{1,1} &= {}_z\langle 1|\hat{\rho}|1\rangle_z = \frac{1}{4}, \\ \rho_{1,0} &= {}_z\langle 1|\hat{\rho}|0\rangle_z = 0, \\ \rho_{1,-1} &= {}_z\langle 1|\hat{\rho}| - 1\rangle_z = 0, \\ \rho_{0,1} &= {}_z\langle 0|\hat{\rho}|1\rangle_z = 0, \\ \rho_{0,0} &= {}_z\langle 0|\hat{\rho}|0\rangle_z = \frac{1}{4}, \\ \rho_{0,-1} &= {}_z\langle 0|\hat{\rho}| - 1\rangle_z = 0, \\ \rho_{-1,1} &= {}_z\langle -1|\hat{\rho}|1\rangle_z = 0, \\ \rho_{-1,0} &= {}_z\langle -1|\hat{\rho}|0\rangle_z = 0, \\ \rho_{-1,-1} &= {}_z\langle -1|\hat{\rho}| - 1\rangle_z = \frac{1}{2},\end{aligned}$$

and so

$$\rho = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} .$$

• • •

### A note on the density matrix

A density matrix is one possible **representation** of a density operator. The representation is with respect to a particular basis, that is, a particular orientation of the axes. For spin (and angular momentum) systems we usually choose a basis in which  $\hat{S}_z$  is diagonal. But this choice is not necessary as we saw in the first lecture; we could choose a basis in which  $\hat{S}_x$  is diagonal, for example. Despite this difference both **density operators** and **density matrices** (regardless of the basis chosen) represent the same quantum states, and so these terms will often be used interchangeably.

### Problem

Calculate the matrix elements of the density operator for the state of the silver atoms in the beam **before** entering the magnet in the Stern-Gerlach experiment discussed in Section 1. Use the density operator to calculate the expectation values  $\langle \hat{S}_x \rangle$ ,  $\langle \hat{S}_y \rangle$  and  $\langle \hat{S}_z \rangle$ . (You should find  $\hat{\rho} = \frac{1}{2}\mathbb{1}$  and that the expectation values are all 0.)

- **Example (2).** Consider a spin-1/2 system consisting of a free neutron. We know the neutron is in the pure state  $|\psi_1\rangle = |\frac{1}{2}\rangle_z$  with probability  $\frac{1}{3}$  and the pure state  $|\psi_2\rangle = \frac{1}{\sqrt{2}}(|\frac{1}{2}\rangle_z + i|-\frac{1}{2}\rangle_z)$  with probability  $\frac{2}{3}$ . Calculate the neutron's density matrix in the  $\{|-\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z\}$  basis.

The density operator is given by

$$\begin{aligned}\hat{\rho} &= \frac{1}{3}|\psi_1\rangle\langle\psi_1| + \frac{2}{3}|\psi_2\rangle\langle\psi_2| \\ &= \frac{1}{3}|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| + \frac{2}{3} \times \frac{1}{2} \left( |\frac{1}{2}\rangle_z + i|-\frac{1}{2}\rangle_z \right) \left( z\langle\frac{1}{2}| - i z\langle-\frac{1}{2}| \right) .\end{aligned}$$

### Note on Hermitian conjugation

Notice the negative sign in the last term  $(z\langle\frac{1}{2}| - i z\langle-\frac{1}{2}|)$ . The Hermitian conjugate operation involves the complex conjugation and so, in general,

$$\left( a|\frac{1}{2}\rangle_z + b|-\frac{1}{2}\rangle_z \right)^\dagger = \left( a^* z\langle\frac{1}{2}| + b^* z\langle-\frac{1}{2}| \right) .$$

Expanding the product in brackets  $(\cdots)(\cdots)$  yields

$$\begin{aligned}\hat{\rho} &= \frac{1}{3}|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| + \frac{1}{3} \left( |\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| - i|\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + i|-\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| + |-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| \right) \\ &= \frac{2}{3}|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| + \frac{1}{3} \left( -i|\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + i|-\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| + |-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| \right)\end{aligned}$$

The matrix elements of  $\hat{\rho}$  can be read off directly:

$$\begin{aligned}\rho_{1,1} &= z\langle\frac{1}{2}|\hat{\rho}|\frac{1}{2}\rangle_z = \frac{2}{3}, \\ \rho_{1,2} &= z\langle\frac{1}{2}|\hat{\rho}|-\frac{1}{2}\rangle_z = -\frac{1}{3}i, \\ \rho_{2,1} &= z\langle-\frac{1}{2}|\hat{\rho}|\frac{1}{2}\rangle_z = \frac{1}{3}i, \\ \rho_{2,2} &= z\langle-\frac{1}{2}|\hat{\rho}|-\frac{1}{2}\rangle_z = \frac{1}{3},\end{aligned}$$

and hence

$$\rho = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3}i \\ \frac{1}{3}i & \frac{1}{3} \end{bmatrix} .$$

• • •

Pure states  $|\psi\rangle$  can also be represented by density operators as  $\hat{\rho} = |\psi\rangle\langle\psi|$ . This is a special case of the equation  $\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ , for which all  $p_i$  are zero except for one which is unity.

### c. What do density operators mean?

Density operators can be thought of as mathematical objects that compactly represent all that we know about some quantum system. In this regard, they are akin to probability distributions we find in classical physics which tell us how much we know about, say, the chance of a coin landing on heads or tails.

- The **diagonal elements** of a density matrix denote the probabilities of finding the system in the corresponding pure states.

For example, suppose that we know that the (spin) state of an electron is

$$\rho = \begin{bmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix} ,$$

in the  $\hat{S}_z$  basis,  $|\frac{1}{2}\rangle_z, |-\frac{1}{2}\rangle_z$ .

The probability of observing the electron's spin in the  $z$  direction to be  $+\hbar/2$  (i.e. spin up  $\uparrow$ ) is

$$P(\frac{1}{2}|\hat{\rho}) = {}_z\langle \frac{1}{2}|\hat{\rho}|\frac{1}{2}\rangle_z = \rho_{1,1} = \frac{1}{4} .$$

Similarly, the probability of observing its spin to be  $-\hbar/2$  (i.e. spin down  $\downarrow$ ) is

$$P(-\frac{1}{2}|\hat{\rho}) = {}_z\langle -\frac{1}{2}|\hat{\rho}|-\frac{1}{2}\rangle_z = \rho_{2,2} = \frac{3}{4} .$$

- The **off-diagonal elements** denote the presence of **coherent superpositions**.

Consider the state  $|\psi\rangle = \left(a\left|\frac{1}{2}\right\rangle_z + b\left|-\frac{1}{2}\right\rangle_z\right)$  where  $|a|^2 + |b|^2 = 1$  for normalisation. This state is a superposition of the eigenstate of  $\hat{S}_z$ . The density matrix representation in the  $\hat{S}_z$  basis is<sup>‡</sup>

insert  
10

$$\rho = \begin{bmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{bmatrix} .$$

If either  $a = 0$  or  $b = 0$  the state  $|\psi\rangle$  would be an eigenstate of  $\hat{S}_z$  (i.e. **not** a superposition) and the off diagonal elements  $\rho_{1,2} = ab^*$  and  $\rho_{1,2} = a^*b$  would be zero. Conversely, if both  $a$  and  $b$  are non-zero, the off-diagonal elements are non-zero and the state  $|\psi\rangle$  represents a **superposition of the basis states**.

## d. Difference between superposition and mixed states

It is important not to confuse a superposition state like

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left( \left|-\frac{1}{2}\right\rangle_z + \left|\frac{1}{2}\right\rangle_z \right) \quad (\text{case I}) \quad (2.20)$$

with a mixed state like

$$\hat{\rho} = \frac{1}{2} \left( \left|-\frac{1}{2}\right\rangle_z \langle -\frac{1}{2}| + \left|\frac{1}{2}\right\rangle_z \langle \frac{1}{2}| \right) . \quad (\text{case II}) \quad (2.21)$$

In both cases there is a 50% chance of the system being found in either the state  $\left|-\frac{1}{2}\right\rangle_z$  or  $\left|\frac{1}{2}\right\rangle_z$  if a measurement of the  $z$  component of spin was made. Nevertheless, Eq. (2.20) and Eq. (2.21) represent different states. This is borne out by a measurement of the  $x$  component of spin which would result in very different outcomes. Indeed we can equivalently represent the two states in Eq. (2.20) and Eq. (2.21) in the basis of the eigenstates of the  $x$  component of spin as:

$$|\psi\rangle = \left|\frac{1}{2}\right\rangle_x , \quad (\text{case I}) \quad (2.22)$$

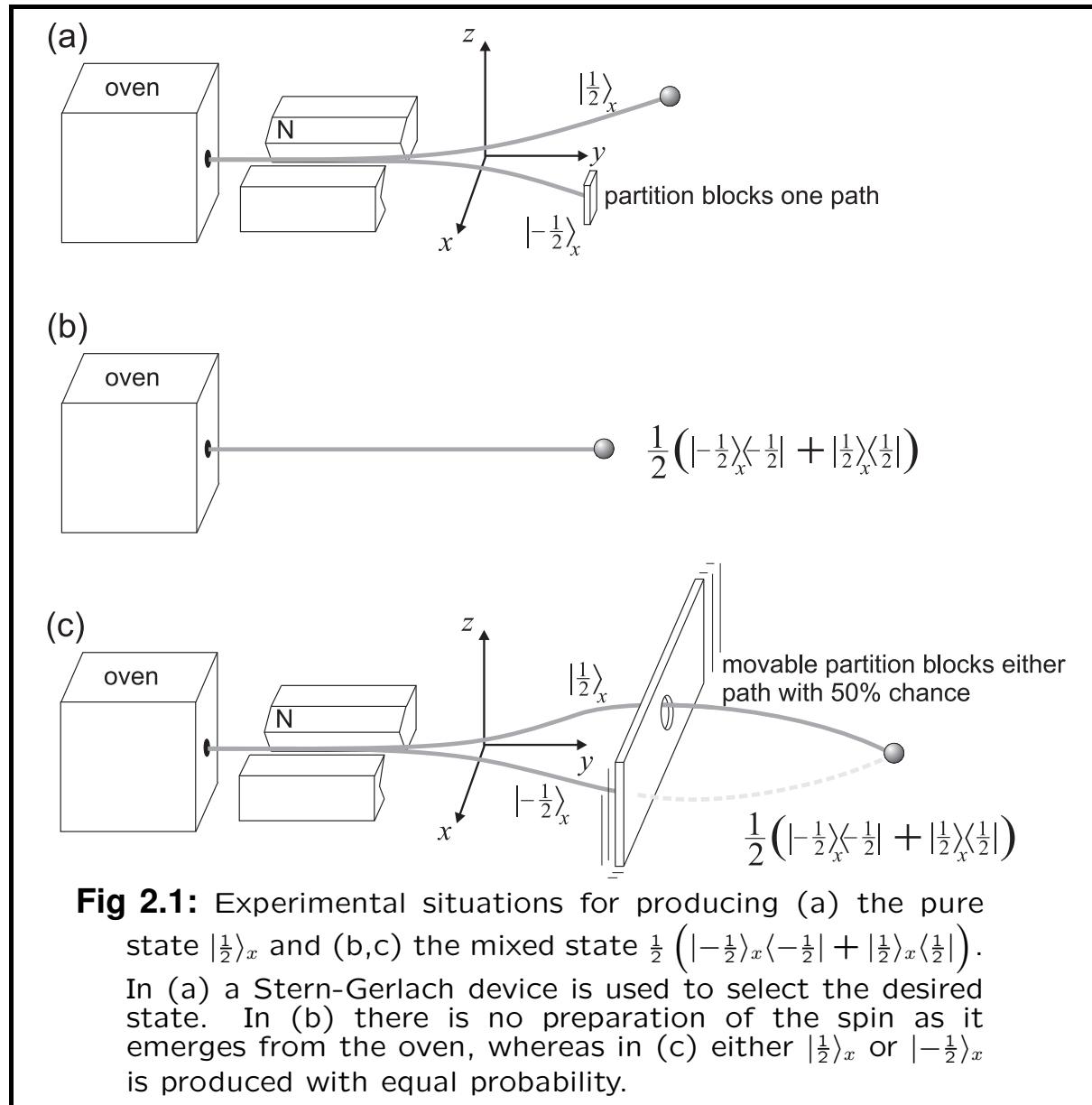
$$\hat{\rho} = \frac{1}{2} \left( \left|-\frac{1}{2}\right\rangle_x \langle -\frac{1}{2}| + \left|\frac{1}{2}\right\rangle_x \langle \frac{1}{2}| \right) . \quad (\text{case II}) \quad (2.23)$$

A measurement of the  $x$  component of spin in the first case will *always* find the system in the state  $\left|\frac{1}{2}\right\rangle_x$ , whereas in the second case there is a 50% chance of the system being found in either the state  $\left|-\frac{1}{2}\right\rangle_x$  or  $\left|\frac{1}{2}\right\rangle_x$ .

<sup>‡</sup>The density operator is given by

$$|\psi\rangle\langle\psi| = |a|^2\left|\frac{1}{2}\right\rangle_{zz}\langle\frac{1}{2}| + ab^*\left|\frac{1}{2}\right\rangle_{zz}\langle-\frac{1}{2}| + a^*b\left|-\frac{1}{2}\right\rangle_{zz}\langle\frac{1}{2}| + |b|^2\left|-\frac{1}{2}\right\rangle_{zz}\langle-\frac{1}{2}| .$$

It is also important to see how the two cases correspond to different experimental situations. Case I represents the outcome of an experiment that produces a spin in the state  $|\frac{1}{2}\rangle_x$ , as illustrated in Fig. 2.1(a). In contrast, case II represents the outcome of an experiment that just produces a spin without preparing its state, as illustrated in Fig. 2.1(b), or in which the state is prepared with 50% chance to be in  $|-\frac{1}{2}\rangle_x$  or  $|\frac{1}{2}\rangle_x$ , as illustrated in Fig. 2.1(c).



**Fig 2.1:** Experimental situations for producing (a) the pure state  $|\frac{1}{2}\rangle_x$  and (b,c) the mixed state  $\frac{1}{2}(|-\frac{1}{2}\rangle_x\langle -\frac{1}{2}| + |\frac{1}{2}\rangle_x\langle \frac{1}{2}|)$ .

In (a) a Stern-Gerlach device is used to select the desired state. In (b) there is no preparation of the spin as it emerges from the oven, whereas in (c) either  $|\frac{1}{2}\rangle_x$  or  $|-\frac{1}{2}\rangle_x$  is produced with equal probability.

## e. Properties of $\hat{\rho}$

We derive the essential properties of general density operators in **Appendix C** on page 26. We simply quote these properties here. (Note that you should know these properties, but it is not necessary to know the proofs as given in the Appendix.)

### Properties of $\hat{\rho}$

- (a)  $\text{Tr}(\hat{\rho}) = 1$  (trace property).
- (b)  $\hat{\rho}$  is Hermitian:  $\hat{\rho} = \hat{\rho}^\dagger$ .
- (c)  $\hat{\rho}$  is a **nonnegative definite** operator, that is, its eigenvalues are non-negative:

$$\hat{\rho} = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|, \quad \lambda_i \geq 0.$$

- (d) The **purity**  $\text{Tr}(\hat{\rho}^2)$  has the following values:

$$\begin{aligned}\text{Tr}(\hat{\rho}^2) &= 1, && \text{if } \hat{\rho} = |\psi\rangle\langle\psi| \text{ is a pure state} \\ \text{Tr}(\hat{\rho}^2) &< 1, && \text{if } \hat{\rho} \text{ is a mixed state.}\end{aligned}$$

## f. Time evolution of the density operator $\hat{\rho}$

The density operator represents the (mixed) state of a physical system. As the system evolves in time the state may change. We now derive the equation of motion for the density operator. Consider the general density operator:

$$\hat{\rho} = \sum_n p_n |\psi_n\rangle\langle\psi_n| \quad (2.24)$$

where  $\sum_n p_n = 1$ ,  $p_n \geq 0$  and where the pure states  $\{|\psi_n\rangle\}$  are normalised but not necessarily orthogonal. We assume the time dependance of  $\hat{\rho}$  is due to the time dependance of the pure states  $\{|\psi_n\rangle\}$  only and that the probabilities  $p_n$  are constant in time. Thus

$$\begin{aligned}\frac{d}{dt}\hat{\rho} &= \sum_n p_n \frac{d}{dt} \left( |\psi_n\rangle\langle\psi_n| \right) \\ &= \sum_n p_n \left[ \left( \frac{d}{dt} |\psi_n\rangle \right) \langle\psi_n| + |\psi_n\rangle \left( \frac{d}{dt} \langle\psi_n| \right) \right] \quad (2.25)\end{aligned}$$

by the product rule. Recall Schrödinger's equation:

insert  
11

$$i\hbar \frac{d}{dt} |\psi\rangle = \hat{\mathcal{H}} |\psi\rangle$$

where  $\hat{\mathcal{H}}$  is the Hamiltonian. Taking the Hermitian conjugate yields

$$\begin{aligned}\left( i\hbar \frac{d}{dt} |\psi\rangle \right)^\dagger &= \left( \hat{\mathcal{H}} |\psi\rangle \right)^\dagger \\ \text{i.e. } -i\hbar \frac{d}{dt} \langle\psi| &= \langle\psi| \hat{\mathcal{H}}^\dagger \\ \text{or } -i\hbar \frac{d}{dt} \langle\psi| &= \langle\psi| \hat{\mathcal{H}} \quad (\text{because } \hat{\mathcal{H}} \text{ is Hermitian}).\end{aligned}$$

Using these results in Eq. (2.25) gives

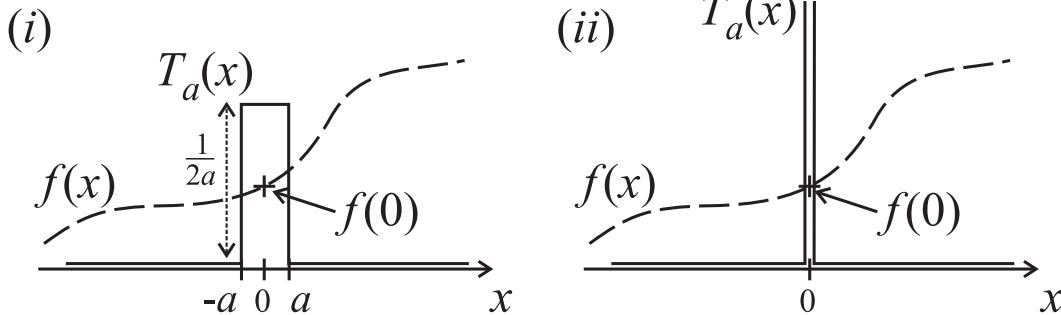
$$\begin{aligned}
 \frac{d}{dt}\hat{\rho} &= \sum_n p_n \left[ \left( -\frac{i}{\hbar} \hat{\mathcal{H}} |\psi_n\rangle \right) \langle \psi_n | + |\psi_n\rangle \left( \frac{i}{\hbar} \langle \psi_n | \hat{\mathcal{H}} \right) \right] \\
 &= -\frac{i}{\hbar} \hat{\mathcal{H}} \left( \sum_n p_n |\psi_n\rangle \langle \psi_n| \right) + \frac{i}{\hbar} \left( \sum_n p_n |\psi_n\rangle \langle \psi_n| \right) \hat{\mathcal{H}} \\
 &= -\frac{i}{\hbar} \left( \hat{\mathcal{H}} \hat{\rho} - \hat{\rho} \hat{\mathcal{H}} \right) \quad [\text{from Eq. (2.24)}] \\
 &= -\frac{i}{\hbar} \left[ \hat{\mathcal{H}}, \hat{\rho} \right].
 \end{aligned}$$

**Schrödinger equation for  $\hat{\rho}$**

$$\frac{d}{dt}\hat{\rho} = -\frac{i}{\hbar} \left[ \hat{\mathcal{H}}, \hat{\rho} \right].$$

## Appendix A - The Dirac $\delta$ distribution

Although  $\delta(x)$  is often called the Dirac delta “function” it is not a function at all, but rather belongs to a more general class of objects called **distributions** (or generalised functions). A brief review of distribution theory is given in Appendix A of Messiah (1961). Here we discuss only those properties of the Dirac delta distribution of direct relevance to this course. The formal theory of distribution, however, is an interesting and inspiring branch of mathematics.



**Fig 2.2:** The top hat function  $T_a(x)$  (solid curve) and an arbitrary “smooth” function  $f(x)$  (dashed curve). The value of  $a$  is considerably smaller in (ii) compared to (i) and illustrates how the value  $f(0)$  is selected in the limit  $a \rightarrow 0$ .

Our approach to the Dirac  $\delta$  distribution will be rather “hands-on”. Imagine a top hat function  $T_a(x)$ ,

$$T_a(x) = \begin{cases} \frac{1}{2a}, & \text{for } -a \leq x \leq a \\ 0, & \text{otherwise.} \end{cases}$$

as illustrated in the figure. Notice that  $T_a(x)$  has been constructed so

$$\int_{-\infty}^{\infty} T_a(x) dx = 1 , \quad (2.26)$$

i.e. the area under the curve of  $T_a(x)$  is unity, irrespective of the value of  $a$ . Consider the integral

$$I_a = \int_{-\infty}^{\infty} f(x) T_a(x) dx$$

of the product of  $T_a(x)$  with an arbitrary function  $f(x)$ . Because the value of  $T_a(x)$  is zero outside the interval  $-a \leq x \leq a$ , and  $\frac{1}{2a}$  inside the interval, the integral can be replaced with

$$I_a = \int_{-a}^a f(x) \frac{1}{2a} dx = \frac{1}{2a} \int_{-a}^a f(x) dx .$$

Imagine a very small value of  $a \ll 1$ . If the function  $f(x)$  is

sufficiently “smooth” its value over the interval  $-a \leq x \leq a$  will approach its value at the origin  $f(0)$ , i.e.

$$f(x) \rightarrow f(0) \text{ as } a \rightarrow 0 \text{ for } x \text{ in the interval } -a \leq x \leq a .$$

We now approximate the integral using Euler's method which is correct to first order in  $a$ , i.e.

$$\begin{aligned} I_a &= \frac{1}{2a} f(0) 2a + \mathcal{O}(a^2) \\ &= f(0) + \mathcal{O}(a^2) . \end{aligned}$$

In the limit  $a \rightarrow 0$  we have

$$I \equiv \lim_{a \rightarrow 0} I_a = f(0)$$

and so collecting results,

$$I = \lim_{a \rightarrow 0} \int_{-\infty}^{\infty} f(x) T_a(x) dx = f(0) . \quad (2.27)$$

### The crux

The real issue is whether we can take the “limit” as  $a \rightarrow 0$  of  $T_a(x)$  itself. Well the answer is no, because  $T_a(0)$  diverges as  $a \rightarrow 0$  and a limit requires a convergent sequence. It **incorrect** (although it tempting) to say that in the limit  $a \rightarrow 0$

- $T_a(x)$  is zero everywhere except at the origin where it is infinite,
- the area under the curve of  $T_a(x)$  is unity, and so
- $\int_{-\infty}^{\infty} f(x) T_a(x) dx = f(0)$  as a result!

Actually it is the **limit of the integral** that is well defined, not the limit of the top hat function. Let's see how the limit works. Imagine the sequence of values of  $a = 0.1, 0.001, 0.0001, 0.00001, \dots$ . There is a corresponding sequence of integrals:

$$I_{0.1} = \frac{1}{0.2} \int_{-0.1}^{0.1} f(x) dx = f(0) + \mathcal{O}(10^{-2}) ,$$

$$I_{0.01} = \frac{1}{0.02} \int_{-0.01}^{0.01} f(x) dx = f(0) + \mathcal{O}(10^{-4}) ,$$

$$I_{0.001} = \frac{1}{0.002} \int_{-0.001}^{0.001} f(x) dx = f(0) + \mathcal{O}(10^{-6}) ,$$

$$I_{0.0001} = \frac{1}{0.0002} \int_{-0.0001}^{0.0001} f(x) dx = f(0) + \mathcal{O}(10^{-8}) ,$$

:

The limit point of this sequence is exactly

$$f(0) .$$

Notice that  $a > 0$  in each of the terms above. At no point is there a term involving  $T_a(x)$  for  $a = 0$ . Taking the limit of the integral neatly side steps the need for such an ill-defined term.

Instead of having to write down the full expression in Eq. (2.27) each time we want to perform an integral involving  $T_a(x)$  a **short-hand notation** has been developed. We simply write

$$\text{“} \int f(x)\delta(x)dx \text{”} \quad \text{for} \quad \lim_{a \rightarrow 0} \int_{-\infty}^{\infty} f(x)T_a(x)dx . \quad (2.28)$$

### Dirac $\delta$ distribution

It is **understood** that any integral involving  $\delta(x)$  such as

$$\int f(x)\delta(x)dx = f(0)$$

is shorthand for the limiting procedure in Eq. (2.27).

### Why “distribution”?

The reason  $\delta(x)$  is called a **distribution** is that  $\delta(x)$  only makes sense inside an integral. Indeed the shorthand notation in Eq. (2.28) involves the **whole integral expression** in quotes, and not  $\delta(x)$  on its own. These kinds of integrals are just like the integrals for probability distributions where, for example, the expectation values of  $x$  for distribution  $P(x)$  is given by

$$\bar{x} = \int xP(x)dx$$

and, more generally,

$$\overline{f(x)} = \int f(x)P(x)dx .$$

Interestingly, there are many equivalent<sup>¶</sup> representations of the  $\delta$  distribution. One is generated by the smooth normal distributions

$$N_a(x) = \frac{1}{a\sqrt{\pi}} e^{-x^2/a^2}$$

which are **bell-shaped curves** of unit area and with a width (standard deviation) of  $a$  with the properties that (quoted without

---

<sup>¶</sup>By “equivalent” we mean having the same end effect.

proof)

$$\int_{-\infty}^{\infty} N_a(x) dx = 1 \quad \& \quad \lim_{a \rightarrow 0} \int_{-\infty}^{\infty} f(x) N_a(x) dx = f(0) .$$

Yet another is

$$S_a(x) = \frac{\sin(x/a)}{\pi x}$$

for which (quoted without proof)

$$\int_{-\infty}^{\infty} S_a(x) dx = 1 \quad \& \quad \lim_{a \rightarrow 0} \int_{-\infty}^{\infty} f(x) S_a(x) dx = f(0) .$$

Indeed, this representation is related to the *Fourier transform* of unity (1). Consider

$$\frac{1}{2\pi} \int_{-1/a}^{1/a} 1 \cdot e^{ixy} dy = \frac{\sin(x/a)}{\pi x} = S_a(x) .$$

and so the Fourier transform of 1 is, symbolically<sup>§</sup>

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 1 \cdot e^{ixy} dy &= \lim_{a \rightarrow 0} \frac{1}{\sqrt{2\pi}} \int_{-1/a}^{1/a} 1 \cdot e^{ixy} dy \\ &= \sqrt{2\pi} \lim_{a \rightarrow 0} S_a(x) \\ &= \sqrt{2\pi} \delta(x) . \end{aligned}$$

Thus, symbolically,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixy} dy = \delta(x) . \quad (2.29)$$

• **Example (3).** Using the method above, show that

$$\int_{-\infty}^{\infty} T_a(x - x_0) dx = 1$$

and

$$\lim_{a \rightarrow 0} \int_{-\infty}^{\infty} f(x) T_a(x - x_0) dx = f(x_0)$$

where  $x_0$  is a constant real number.

---

<sup>§</sup>The limit  $a \rightarrow 0$  does not actually exist here. Rather the limit should be delayed until after another integral over  $x$  is taken, e.g.  $\lim_{a \rightarrow 0} \int f(x) \left[ \frac{1}{\sqrt{2\pi}} \int_{-1/a}^{1/a} 1 \cdot e^{ixy} dy \right] dx$ . We imagine that this extra integral is **implicit** in the expressions. This is what is meant by “symbolically”.

### Solution

Using the change of variable,  $x' = x - x_0$  (for which  $dx' = dx$ ) and Eq. (2.26) we find

$$\begin{aligned}\int_{-\infty}^{\infty} T_a(x - x_0) dx &= \int_{-\infty}^{\infty} T_a(x') dx' \\ &= 1\end{aligned}$$

as required.

Again using the change of variable,  $x' = x - x_0$ , and the substitution  $g(x) = f(x + x_0)$  we find

$$\begin{aligned}\lim_{a \rightarrow 0} \int_{-\infty}^{\infty} f(x) T_a(x - x_0) dx &= \lim_{a \rightarrow 0} \int_{-\infty}^{\infty} f(x + x_0) T_a(x') dx' \\ &= \lim_{a \rightarrow 0} \int_{-\infty}^{\infty} g(x) T_a(x') dx' \\ &= g(0)\end{aligned}$$

by Eq. (2.27). As  $g(0) = f(x_0)$  from the substitution, we finally have

$$\lim_{a \rightarrow 0} \int_{-\infty}^{\infty} f(x) T_a(x - x_0) dx = f(x_0)$$

as required.

• • •

### Some properties of the Dirac $\delta$ distribution

(a)  $\int \delta(x - x_0) dx = 1$

(b)  $\int f(x) \delta(x - x_0) dx = f(x_0)$

(c)  $\int f(x) \frac{d}{dx} \delta(x - x_0) dx = \left. \frac{d}{dx} f(x) \right|_{x=x_0}$

(d)  $\delta(ax) = \frac{1}{|a|} \delta(x)$

(e)  $\delta(x^2 - a^2) = \frac{1}{2|a|} \left[ \delta(x + a) + \delta(x - a) \right]$

Finally recall that the symbol

$$\text{“} \int f(x)dx \text{”}$$

is actually shorthand for the limit

$$\lim_{\Delta x \rightarrow 0} \sum_n f(x_n) \Delta x \quad \text{where } x_m = m\Delta x + x_0 \text{ etc. .}$$

So the idea of using shorthand symbols is not new at all.

## Appendix B - The momentum operator

### Background

- The Taylor series for differentiable function  $f(x)$  can be written as

$$\begin{aligned} f(x + \epsilon) &= \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} \frac{d^n}{dx^n} f(x) \\ &= \exp\left(\epsilon \frac{d}{dx}\right) f(x) \end{aligned}$$

where we have used the exponential series  $\exp(y) = \sum_n y^n/n!$ . The differential operator  $d/dx$  can be thought of as a generator of translations in  $f(x)$ .

- The textbook by Sakurai (p. 46) gives an interesting account of why the momentum operator should be related to the generator of translations in position. Fundamentally, defining momentum in this way ensures that the *conservation of momentum* is identical to the *lack of a preferred origin* in space (and vice versa).

We use **translations in position** to define the **momentum operator**  $\hat{X}$ . The generator of translations is defined<sup>¶</sup> to be  $-i\hat{P}/\hbar$  and so

$$|x + \epsilon\rangle = \exp\left(-i\epsilon\hat{P}/\hbar\right)|x\rangle .$$

Expanding the right-hand side to first order in  $\epsilon$  and rearranging yields

$$\begin{aligned} |x + \epsilon\rangle &= \left(1 - i\epsilon\hat{P}/\hbar\right)|x\rangle + \mathcal{O}(\epsilon^2) \\ &= |x\rangle - i\epsilon\hat{P}|x\rangle/\hbar + \mathcal{O}(\epsilon^2) \\ \hat{P}|x\rangle &= i\hbar \frac{|x + \epsilon\rangle - |x\rangle}{\epsilon} + \mathcal{O}(\epsilon) . \end{aligned}$$

In the limit  $\epsilon \rightarrow 0$  we get

$$\hat{P}|x\rangle = i\hbar \frac{d}{dx}|x\rangle . \quad (2.30)$$

---

<sup>¶</sup>This definition is consistent with  $\hat{X}$  and  $\hat{P}$  satisfying the **canonical commutation** relation  $[\hat{X}, \hat{P}] = i\hbar$ , as Eq. (2.14) shows.

Thus the moments of  $\hat{P}$  are

$$\begin{aligned}
\langle \hat{P}^n \rangle &= \langle \psi | \hat{P}^n | \psi \rangle = \langle \psi | \hat{P}^n \times \hat{\mathbf{1}} | \psi \rangle \\
&= \langle \psi | \hat{P}^n \left( \int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) | \psi \rangle \quad \text{from Eq. (2.1)} \\
&= \int_{-\infty}^{\infty} \langle \psi | \hat{P}^n | x \rangle \langle x | \psi \rangle dx \\
&= \int_{-\infty}^{\infty} \langle \psi | \left[ (i\hbar)^n \frac{d^n}{dx^n} |x\rangle \right] \langle x | \psi \rangle dx \quad \text{from Eq. (2.7)} \\
&= \int_{-\infty}^{\infty} (i\hbar)^n \left[ \frac{d^n}{dx^n} \langle \psi | x \rangle \right] \langle x | \psi \rangle dx \\
&= \int_{-\infty}^{\infty} (i\hbar)^n \left[ \frac{d^n}{dx^n} \psi^*(x) \right] \psi(x) dx \quad \text{from Eq. (2.6)} \\
&= \int_{-\infty}^{\infty} \psi^*(x) \left[ -i\hbar \frac{d}{dx} \right]^n \psi(x) dx
\end{aligned} \tag{2.31}$$

where we have used integration by parts  $n$  times and the assumption that  $[\frac{d^n}{dx^n} \psi^*(x) \psi(x)]_{-\infty}^{\infty} = 0$  etc. to arrive at the last line.

### Relationship between representations

To find the relationship between the position and momentum bases we examine the expression  $\langle x | \hat{P} | p \rangle$  in two different ways. First from Eq. (2.8) we get

$$\langle x | \hat{P} | p \rangle = p \langle x | p \rangle . \tag{2.32}$$

Next noting that  $\langle A | B \rangle = (\langle B | A \rangle)^*$  and so  $\langle x | \hat{P} | p \rangle = \langle x | Pp \rangle = (\langle Pp | x \rangle)^* = (\langle p | \hat{P}^\dagger | x \rangle)^* = (\langle p | \hat{P} | x \rangle)^*$ , as  $\hat{P}$  is self-adjoint, we get

$$\begin{aligned}
\langle x | \hat{P} | p \rangle &= \left( \langle p | \hat{P} | x \rangle \right)^* \\
&= \left( i\hbar \frac{d}{dx} \langle p | x \rangle \right)^* \quad \text{from Eq. (2.7)} \\
&= -i\hbar \frac{d}{dx} \langle x | p \rangle .
\end{aligned} \tag{2.33}$$

Equating the right-hand sides of Eqs. (2.32) and (2.33) gives the differential equation

$$-i\hbar \frac{d}{dx} \langle x | p \rangle = p \langle x | p \rangle$$

whose solution is

$$\langle x | p \rangle = N e^{ipx/\hbar} \tag{2.34}$$

where  $N$  is a constant whose value is determined as follows.

Consider the orthogonality of the position eigenkets:

$$\begin{aligned}
\delta(x - x') &= \langle x|x'\rangle = \langle x|\hat{\mathbf{1}}|x'\rangle \\
&= \langle x|\left(\int_{-\infty}^{\infty}|p\rangle\langle p|dp\right)|x'\rangle \quad \text{from Eq. (2.9)} \\
&= \int_{-\infty}^{\infty}\langle x|p\rangle\langle p|x'\rangle dp \\
&= \int_{-\infty}^{\infty}Ne^{ipx/\hbar}\left(Ne^{ipx'/\hbar}\right)^*dp \quad \text{from Eq. (2.34)} .
\end{aligned}$$

We can choose  $N$  to be real valued without loss of generality, and so we find

$$\begin{aligned}
\delta(x - x') &= N^2 \int_{-\infty}^{\infty} e^{ip(x-x')/\hbar} dp \\
&= N^2 2\pi\hbar\delta(x - x')
\end{aligned}$$

where we have used the representation of the delta function Eq. (2.29)

$$\int_{-\infty}^{\infty} e^{izy} dz = 2\pi\hbar\delta(y) .$$

Thus  $N = 1/\sqrt{2\pi\hbar}$  and so

$$\langle x|p\rangle = \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} . \quad (2.35)$$

## Appendix C - Properties of the density operator

We derive the essential properties of general density operators. Consider an arbitrary density operator formed from a statistical mixture of pure states:

$$\hat{\rho} = \sum_n p_n |\psi_n\rangle\langle\psi_n| \quad (2.36)$$

where  $\sum_n p_n = 1$  for the probabilities  $p_n \geq 0$ . Note that the pure states  $\{|\psi_n\rangle : n = 1, 2, \dots\}$  need not be orthogonal, but we assume that they are normalised.

- **Self-adjoint**

The density operator is **self-adjoint** because

$$\hat{\rho}^\dagger = \sum_n p_n^* \left( |\psi_n\rangle\langle\psi_n| \right)^\dagger = \sum_n p_n |\psi_n\rangle\langle\psi_n| = \hat{\rho} .$$

- **Non-negative**

From the spectral theorem we can expand  $\hat{\rho}$  in its eigenket basis

$$\hat{\rho} = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i| \quad (2.37)$$

where  $\{\lambda_i\}$  are the (real) eigenvalues and the eigenkets  $\{|\lambda_i\rangle\}$  form an orthonormal basis such that  $\hat{\mathbb{1}} = \sum_i |\lambda_i\rangle\langle\lambda_i|$ . Taking the trace of Eq. (2.36) in the eigenbasis gives

$$\begin{aligned} \text{Tr}(\hat{\rho}) &= \sum_i \langle\lambda_i| \left( \sum_n p_n |\psi_n\rangle\langle\psi_n| \right) |\lambda_i\rangle = \sum_i \sum_n p_n \langle\lambda_i|\psi_n\rangle\langle\psi_n|\lambda_i\rangle \\ &= \sum_i \sum_n p_n \langle\psi_n|\lambda_i\rangle\langle\lambda_i|\psi_n\rangle = \sum_n p_n \langle\psi_n| \left( \sum_i |\lambda_i\rangle\langle\lambda_i| \right) |\psi_n\rangle \\ &= \sum_n p_n \langle\psi_n|\hat{\mathbb{1}}|\psi_n\rangle = \sum_n p_n \quad (\text{from } \langle\psi_n|\psi_n\rangle = 1) \\ &= 1 . \end{aligned} \quad (2.38)$$

whereas taking the trace of Eq. (2.37) in the eigenbasis gives

$$\text{Tr}(\hat{\rho}) = \sum_i \lambda_i , \quad (2.39)$$

and so we find from Eq. (2.38) and Eq. (2.39)

$$\sum_i \lambda_i = 1 . \quad (2.40)$$

The probability of finding the system in the eigenket  $|\lambda_i\rangle$  is, using

Eq. (2.36) for  $\hat{\rho}$ ,

$$\begin{aligned}
 P(\lambda_i|\hat{\rho}) &= \langle \lambda_i | \hat{\rho} | \lambda_i \rangle = \langle \lambda_i | \left( \sum_n p_n |\psi_n\rangle \langle \psi_n| \right) | \lambda_i \rangle \\
 &= \sum_n p_n \langle \lambda_i | \psi_n \rangle \langle \psi_n | \lambda_i \rangle = \sum_n p_n |\langle \lambda_i | \psi_n \rangle|^2 \\
 &\geq 0 \quad (\text{because } |\langle \lambda_i | \psi_n \rangle|^2 \geq 0)
 \end{aligned} \tag{2.41}$$

or, using Eq. (2.37) for  $\hat{\rho}$ ,

$$\begin{aligned}
 P(\lambda_i|\hat{\rho}) &= \langle \lambda_i | \hat{\rho} | \lambda_i \rangle = \langle \lambda_i | \left( \sum_j \lambda_j |\lambda_j\rangle \langle \lambda_j| \right) | \lambda_i \rangle \\
 &= \lambda_i
 \end{aligned} \tag{2.42}$$

and so, comparing Eq. (2.41) with Eq. (2.42), we find that

$$\lambda_i \geq 0 \tag{2.43}$$

and so the density operator is **non-negative**.

- **Bounded & Hermitian**

First consider the probability of finding the system in the arbitrary state  $|\phi\rangle$ :

$$\begin{aligned}
 P(\phi|\hat{\rho}) &= \langle \phi | \hat{\rho} | \phi \rangle = \langle \phi | \left( \sum_i \lambda_i |\lambda_i\rangle \langle \lambda_i| \right) | \phi \rangle = \sum_i \lambda_i \langle \phi | \lambda_i \rangle \langle \lambda_i | \phi \rangle \\
 &= \sum_i \lambda_i |\langle \phi | \lambda_i \rangle|^2 .
 \end{aligned} \tag{2.44}$$

The left-hand side is a probability and so  $P(\phi|\hat{\rho}) \leq 1$ ; thus Eq. (2.44) implies

$$\sum_i \lambda_i |\langle \phi | \lambda_i \rangle|^2 \leq 1 . \tag{2.45}$$

Next consider the value of

$$\begin{aligned}
 \langle \phi | \hat{\rho}^2 | \phi \rangle &= \langle \phi | \left( \sum_i \lambda_i^2 |\lambda_i\rangle \langle \lambda_i| \right) | \phi \rangle \\
 &= \sum_i \lambda_i^2 |\langle \phi | \lambda_i \rangle|^2 \\
 &\leq \sum_i \lambda_i |\langle \phi | \lambda_i \rangle|^2 \quad (\text{because } \lambda_i^2 \leq \lambda_i \leq 1 \text{ and } \lambda_i \geq 0)
 \end{aligned}$$

and so from Eq. (2.45) we find

$$\langle \phi | \hat{\rho}^2 | \phi \rangle \leq 1 . \tag{2.46}$$

The norm of  $\hat{\rho}|\phi\rangle$  for any  $|\phi\rangle$  in the Hilbert space is thus

$$\begin{aligned}\left\| \hat{\rho}|\phi\rangle \right\| &= \sqrt{\langle \phi | \hat{\rho}^\dagger \hat{\rho} | \phi \rangle} = \sqrt{\langle \phi | \hat{\rho}^2 | \phi \rangle} \quad (\text{because } \hat{\rho}^\dagger = \hat{\rho}) \\ &\leq 1 \quad [\text{from Eq. (2.46)}]\end{aligned}$$

and so  $\hat{\rho}$  is a **bounded operator**. Hence  $\hat{\rho}$  is a bounded self-adjoint operator, and so it is also **Hermitian**.

### • Purity

Finally consider the trace of  $\hat{\rho}^2$  in the eigenket basis Eq. (2.37):

$$\text{Tr}(\hat{\rho}^2) = \text{Tr}\left(\sum_i \lambda_i^2 |\lambda_i\rangle\langle\lambda_i|\right) = \sum_i \lambda_i^2 .$$

If only one of the eigenvalues is unity, say  $\lambda_k = 1$ , and the remainder are zero then  $\text{Tr}(\hat{\rho}^2) = \sum_i \lambda_i^2 = \lambda_k^2 = 1$ ; this is a pure state case  $\hat{\rho} = |\lambda_k\rangle\langle\lambda_k|$ . In all other cases  $\lambda_i < 1$ , so  $\text{Tr}(\hat{\rho}^2) = \sum_i \lambda_i^2 < \sum_i \lambda_i$  and from Eq. (2.40) we find  $\text{Tr}(\hat{\rho}^2) < 1$ ; these represent mixed states. Hence  $\text{Tr}(\hat{\rho}^2)$  can be used to distinguish pure states ( $\text{Tr}(\hat{\rho}^2) = 1$ ) from mixed ones ( $\text{Tr}(\hat{\rho}^2) < 1$ ). Technically,  $\text{Tr}(\hat{\rho}^2)$  is called the **purity** of  $\hat{\rho}$ .

# Problem Sheet

- (1) Verify that the right-hand side of Eq. (2.3) gives the eigenvalue equation  $\hat{X}|x'\rangle = x'|x'\rangle$ .

- (2) Use Eq. (2.3) to show that the  $n$ -th power of  $\hat{X}$  is given by

$$\hat{X}^n = \int_{-\infty}^{\infty} x^n |x\rangle\langle x| dx .$$

- (3) Show that the relationship between the position and momentum representations  $\psi(x) = \langle x|\psi\rangle$  and  $\tilde{\psi}(p) = \langle p|\psi\rangle$  is the Fourier transform

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \psi(x) dx .$$

[**HINTS:** Begin with the expression  $\tilde{\psi}(p) = \langle p|\psi\rangle$ , write the right-hand side as  $\langle p|\hat{1}\!\!1|\psi\rangle$  and use the resolution of the identity in Eq. (2.1) to replace  $\hat{1}\!\!1$ , replace  $\langle p|x\rangle = \langle x|p\rangle^*$  using Eq. (2.10) and replace  $\langle x|\psi\rangle$  using Eq. (2.6).]

- (4) Show that the momentum eigenkets can be written in the position basis as

$$|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{ipx/\hbar} |x\rangle dx$$

and conversely that

$$|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} |p\rangle dp .$$

[**HINTS:** Begin with  $|p\rangle = \hat{1}\!\!1|p\rangle$  and replace  $\hat{1}\!\!1$  using Eq. (2.1) and  $\langle x|p\rangle$  using Eq. (2.10).]

- (5) Show that the position operator  $\hat{X}$  in the momentum basis is given by

$$\hat{X}|p\rangle = -i\hbar \frac{d}{dp} |p\rangle .$$

[**HINTS:** Begin with  $\hat{X}|p\rangle$ , replace  $\hat{X}$  using Eq. (2.3), replace  $\langle x|p\rangle$  using Eq. (2.10), recognise that  $xe^{ipx/\hbar} = -i\hbar \frac{d}{dp} e^{ipx/\hbar}$ , take the derivative outside the integral, and convert into an expression involving  $|p\rangle$ .]

- (6) (**Harder**) Show that the commutator between  $\hat{X}$  and  $\hat{P}$  is given by

$$[\hat{X}, \hat{P}] = i\hbar .$$

[**HINT:** write the commutator as  $[\hat{X}, \hat{P}] \times \hat{\mathbf{1}}$ , use the resolution of the identity given by Eq. (2.1) or Eq. (2.9) and one of Eq. (2.7) and Eq. (2.13). Also note that

- $\hat{X}\hat{P}|x\rangle = \hat{X} i\hbar \frac{d}{dx}|x\rangle = i\hbar \frac{d}{dx}(\hat{X}|x\rangle) = i\hbar \frac{d}{dx}(x|x\rangle)$ ,
- $\hat{P}\hat{X}|x\rangle = \hat{P} x|x\rangle = x\hat{P}|x\rangle = xi\hbar \frac{d}{dx}|x\rangle$ , and
- $\frac{d}{dx}(x|x\rangle) = |x\rangle + x\frac{d}{dx}|x\rangle$  by the product rule.

]

- (7) Show that the spin operators  $\hat{S}_x$ ,  $\hat{S}_y$  and  $\hat{S}_z$  are **traceless**, that is, that they all have a zero value for the trace.
- (8) Calculate the matrix elements of the density operator for the state of the silver atoms in the beam **before** entering the magnet in the Stern-Gerlach experiment discussed in Section 1. Use the density operator to calculate the expectation values  $\langle \hat{S}_x \rangle$ ,  $\langle \hat{S}_y \rangle$  and  $\langle \hat{S}_z \rangle$ . (You should find  $\hat{\rho} = \frac{1}{2}\hat{\mathbf{1}}$  and that the expectation values are all 0.)
- (9) (a) Calculate the density operator  $\hat{\rho}$  for a system that is in the pure state  $|-\frac{1}{2}\rangle_z$  with probability  $\frac{1}{4}$  and the state  $\left(\frac{1}{\sqrt{3}}|\frac{1}{2}\rangle_z + i\frac{\sqrt{2}}{\sqrt{3}}|-\frac{1}{2}\rangle_z\right)$  with probability  $\frac{3}{4}$ .
- (b) What is the probability of measuring  $\hat{S}_z$  and finding the value of  $-\frac{1}{2}\hbar$ ?
- (c) Show that  $\hat{\rho}$  is nonnegative definite (i.e. that its eigenvalues are greater than or equal to zero).
- (d) Show that  $\text{Tr}(\hat{\rho}^2) = \frac{7}{8}$ .

# Advanced Quantum Theory

## 2 Quantum States

### Solutions to Problems

- (1)** Substituting Eq. (2.3) into the eigenvalue equation Eq. (2.2) yields

$$\begin{aligned}
 \hat{X}|x'\rangle &= \left( \int_{-\infty}^{\infty} x|x\rangle\langle x|dx \right)|x'\rangle = \int_{-\infty}^{\infty} x|x\rangle \left( \langle x|x' \rangle \right) dx \\
 &= \int_{-\infty}^{\infty} x|x\rangle \delta(x - x') dx \\
 &= x'|x'\rangle .
 \end{aligned}$$

- (2)** First extend the eigenvalue equation Eq. (2.2) to

$$\begin{aligned}
 \hat{X}^n|x\rangle &= \hat{X}^{n-1}\hat{X}|x\rangle = X^{n-1}x|x\rangle = x\hat{X}^{n-2}\hat{X}|x\rangle \\
 &= x\hat{X}^{n-2}x|x\rangle = x^2\hat{X}^{n-3}\hat{X}|x\rangle = x^3\hat{X}^{n-3}|x\rangle = \dots \\
 &= x^n|x\rangle ,
 \end{aligned}$$

then

$$\begin{aligned}
 \hat{X}^n &= \hat{X}^n \times \hat{1} = \hat{X}^n \times \left( \int_{-\infty}^{\infty} |x\rangle\langle x|dx \right) = \int_{-\infty}^{\infty} \hat{X}^n|x\rangle\langle x|dx \\
 &= \int_{-\infty}^{\infty} x^n|x\rangle\langle x|dx
 \end{aligned}$$

as required.

- (3)** Following the Hints we have

$$\begin{aligned}
 \tilde{\psi}(p) &= \langle p|\psi \rangle = \langle p|\hat{1}|\psi \rangle \\
 &= \langle p| \left( \int_{-\infty}^{\infty} |x\rangle\langle x|dx \right) |\psi \rangle \quad \text{from Eq. (2.1)} \\
 &= \int_{-\infty}^{\infty} \langle p|x\rangle\langle x|\psi \rangle dx = \int_{-\infty}^{\infty} \left( \langle x|p \rangle \right)^* \langle x|\psi \rangle dx \\
 &= \int_{-\infty}^{\infty} \left( \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} \right)^* \langle x|\psi \rangle dx \quad \text{from Eq. (2.10)} \\
 &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \psi(x) dx \quad \text{from Eq. (2.6)}
 \end{aligned}$$

which is the Fourier transform of  $\psi(x)$ .

**(4)** We begin with  $|p\rangle = \hat{\mathbb{1}}|p\rangle$ , that is

$$\begin{aligned} |p\rangle &= \hat{\mathbb{1}}|p\rangle = \left( \int_{-\infty}^{\infty} |x\rangle\langle x| dx \right) |p\rangle \quad \text{from Eq. (2.1)} \\ &= \int_{-\infty}^{\infty} |x\rangle\langle x| p \rangle dx = \int_{-\infty}^{\infty} |x\rangle \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} dx \quad \text{from Eq. (2.10)} \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{ipx/\hbar} |x\rangle dx . \end{aligned}$$

Conversely,

$$\begin{aligned} |x\rangle &= \hat{\mathbb{1}}|x\rangle = \left( \int_{-\infty}^{\infty} |p\rangle\langle p| dp \right) |x\rangle \quad \text{from Eq. (2.9)} \\ &= \int_{-\infty}^{\infty} |p\rangle\langle p| x \rangle dp = \int_{-\infty}^{\infty} |p\rangle \left( \langle x|p \rangle \right)^* dp \\ &= \int_{-\infty}^{\infty} |p\rangle \frac{e^{-ipx/\hbar}}{\sqrt{2\pi\hbar}} dp \quad \text{from Eq. (2.10)} \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} |p\rangle dp . \end{aligned}$$

**(5)** Replacing  $\hat{X}$  using Eq. (2.3) gives

$$\begin{aligned} \hat{X}|p\rangle &= \left( \int_{-\infty}^{\infty} x|x\rangle\langle x| dx \right) |p\rangle \\ &= \int_{-\infty}^{\infty} x|x\rangle\langle x| p \rangle dx \\ &= \int_{-\infty}^{\infty} x|x\rangle \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} dx \quad \text{from Eq. (2.10)} \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} |x\rangle x e^{ipx/\hbar} dx \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} |x\rangle \left( -i\hbar \frac{d}{dp} e^{ipx/\hbar} \right) dx \\ &= -i\hbar \frac{d}{dp} \int_{-\infty}^{\infty} |x\rangle \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} dx \\ &= -i\hbar \frac{d}{dp} \int_{-\infty}^{\infty} |x\rangle\langle x| p \rangle dx \quad \text{from Eq. (2.10)} \\ &= -i\hbar \frac{d}{dp} |p\rangle \quad \text{from Eq. (2.1)} . \end{aligned}$$

(6) Using the resolution of the identity in the  $|x\rangle$  basis gives

$$\begin{aligned}
[\hat{X}, \hat{P}] &= [\hat{X}, \hat{P}] \times \hat{\mathbb{1}} = [\hat{X}, \hat{P}] \left( \int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) \\
&= \hat{X} \hat{P} \left( \int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) - \hat{P} \hat{X} \left( \int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) \\
&= \hat{X} \int_{-\infty}^{\infty} \hat{P} |x\rangle \langle x| dx - \hat{P} \int_{-\infty}^{\infty} \hat{X} |x\rangle \langle x| dx \\
&= \hat{X} \int_{-\infty}^{\infty} \left( i\hbar \frac{d}{dx} |x\rangle \right) \langle x| dx - \hat{P} \int_{-\infty}^{\infty} x |x\rangle \langle x| dx \\
&\quad \text{from Eq. (2.7)} \\
&= \int_{-\infty}^{\infty} \left[ i\hbar \frac{d}{dx} \left( \hat{X} |x\rangle \right) \right] \langle x| dx - \int_{-\infty}^{\infty} x \hat{P} |x\rangle \langle x| dx \\
&= \int_{-\infty}^{\infty} \left[ i\hbar \frac{d}{dx} \left( x |x\rangle \right) \right] \langle x| dx - \int_{-\infty}^{\infty} x \left( i\hbar \frac{d}{dx} |x\rangle \right) \langle x| dx \\
&\quad \text{from Eq. (2.7)} \\
&= i\hbar \int_{-\infty}^{\infty} \left( |x\rangle + x \frac{d}{dx} |x\rangle \right) \langle x| dx - i\hbar \int_{-\infty}^{\infty} x \left( \frac{d}{dx} |x\rangle \right) \langle x| dx \\
&\quad \text{by the product rule} \\
&= i\hbar \int_{-\infty}^{\infty} |x\rangle \langle x| dx \quad \text{and using Eq. (2.1) we get} \\
&= i\hbar .
\end{aligned}$$

(7) The trace of operator  $\hat{O}$  is evaluated in the  $\{|-\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z\}$  basis as

$$\begin{aligned}
\text{Tr}(\hat{O}) &= {}_z\langle -\frac{1}{2} | \hat{O} | -\frac{1}{2} \rangle_z + {}_z\langle \frac{1}{2} | \hat{O} | \frac{1}{2} \rangle_z \\
&= O_{1,1} + O_{2,2}
\end{aligned}$$

where  $O_{n,m}$  are the matrix elements of  $\hat{O}$  in the same basis. From the first section, **Dirac's bra-ket notation**, we found that the matrix representations of the spin operators in the

$\{|-\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z\}$  basis are

$$\begin{aligned} \mathbf{S}_z &= \frac{1}{2}\hbar \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ \mathbf{S}_x &= \frac{1}{2}\hbar \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \mathbf{S}_y &= \frac{1}{2}\hbar \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \end{aligned}$$

Each of these matrices has a zero trace. Hence the spin operators  $\hat{S}_x$ ,  $\hat{S}_y$  and  $\hat{S}_z$  are **traceless**.

- (8) It was mentioned above that before the atom reaches the magnet there is a 50% chance that it is in the state  $|\frac{1}{2}\rangle_z$  and a 50% chance that it is in  $|-\frac{1}{2}\rangle_z$ . According to Eq. (2.16) this means that the density operator is

$$\begin{aligned} \hat{\rho} &= \frac{1}{2}|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + \frac{1}{2}|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| \\ &= \frac{1}{2}\hat{\mathbb{1}}. \end{aligned}$$

Recall from Section 1 that the spin operators are given by

$$\begin{aligned} \hat{S}_x &= \frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| \\ \hat{S}_y &= -i\frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + i\frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| \\ \hat{S}_z &= \frac{1}{2}\hbar|\frac{1}{2}\rangle_z\langle\frac{1}{2}| - \frac{1}{2}\hbar|-\frac{1}{2}\rangle_z\langle-\frac{1}{2}|. \end{aligned}$$

Their expectation values are found using Eq. (2.17) and so, evaluating the trace in the  $\{|-\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z\}$  basis we find

$$\begin{aligned} \langle\hat{S}_x\rangle &= \text{Tr}\left(\hat{\rho}\hat{S}_x\right) \\ &= \text{Tr}\left[\hat{\rho}\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right)\right] \\ &= {}_z\langle-\frac{1}{2}| \left[\hat{\rho}\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right)\right] |-\frac{1}{2}\rangle_z \\ &\quad + {}_z\langle\frac{1}{2}| \left[\hat{\rho}\left(\frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right)\right] |\frac{1}{2}\rangle_z \\ &= \frac{1}{2}\hbar\left({}_z\langle-\frac{1}{2}|\hat{\rho}|\frac{1}{2}\rangle_z + {}_z\langle\frac{1}{2}|\hat{\rho}|-\frac{1}{2}\rangle_z\right) \\ &= \frac{1}{2}\hbar\left({}_z\langle-\frac{1}{2}|\frac{1}{2}\hat{\mathbb{1}}|\frac{1}{2}\rangle_z + {}_z\langle\frac{1}{2}|\frac{1}{2}\hat{\mathbb{1}}|-\frac{1}{2}\rangle_z\right) \\ &= 0. \end{aligned}$$

An alternate and more efficient method is to relate the ex-

pectation value to the trace of  $\hat{S}_x$  itself:

$$\begin{aligned}\langle \hat{S}_x \rangle &= \text{Tr}(\hat{\rho} \hat{S}_x) = \text{Tr}\left(\frac{1}{2} \mathbb{1} \hat{S}_x\right) = \frac{1}{2} \text{Tr}(\hat{S}_x) \\ &= 0\end{aligned}$$

because  $\hat{S}_x$  is traceless. For the same reason

$$\begin{aligned}\langle \hat{S}_y \rangle &= \text{Tr}(\hat{\rho} \hat{S}_y) = \text{Tr}\left(\frac{1}{2} \mathbb{1} \hat{S}_y\right) = \frac{1}{2} \text{Tr}(\hat{S}_y) \\ &= 0,\end{aligned}$$

and

$$\begin{aligned}\langle \hat{S}_z \rangle &= \text{Tr}(\hat{\rho} \hat{S}_z) = \text{Tr}\left(\frac{1}{2} \mathbb{1} \hat{S}_z\right) = \frac{1}{2} \text{Tr}(\hat{S}_z) \\ &= 0.\end{aligned}$$

**(9) (a)** From Eq. (2.16) we have

$$\begin{aligned}\hat{\rho} &= \frac{1}{4} \left| -\frac{1}{2} \right\rangle_{zz} \left\langle -\frac{1}{2} \right| \\ &\quad + \frac{3}{4} \left( \frac{1}{\sqrt{3}} \left| \frac{1}{2} \right\rangle_z + i \frac{\sqrt{2}}{\sqrt{3}} \left| -\frac{1}{2} \right\rangle_z \right) \left( \frac{1}{\sqrt{3}} z \left\langle \frac{1}{2} \right| - i \frac{\sqrt{2}}{\sqrt{3}} z \left\langle -\frac{1}{2} \right| \right)\end{aligned}$$

**(b)** The measurement of  $\hat{S}_z$  with outcome  $-\frac{1}{2}\hbar$  corresponds to finding the system in the eigenstate  $\left| -\frac{1}{2} \right\rangle_z$ . From Eq. (2.18) this means the probability for this event is

$$\begin{aligned}P\left(-\frac{1}{2} | \hat{\rho}\right) &= {}_z \langle -\frac{1}{2} | \hat{\rho} | -\frac{1}{2} \rangle_z \\ &= {}_z \langle -\frac{1}{2} | \left[ \frac{1}{4} \left| -\frac{1}{2} \right\rangle_{zz} \left\langle -\frac{1}{2} \right| \right. \\ &\quad \left. + \frac{3}{4} \left( \frac{1}{\sqrt{3}} \left| \frac{1}{2} \right\rangle_z + i \frac{\sqrt{2}}{\sqrt{3}} \left| -\frac{1}{2} \right\rangle_z \right) \left( \frac{1}{\sqrt{3}} z \left\langle \frac{1}{2} \right| - i \frac{\sqrt{2}}{\sqrt{3}} z \left\langle -\frac{1}{2} \right| \right) \right] | -\frac{1}{2} \rangle_z \\ &= {}_z \langle -\frac{1}{2} | \left[ \frac{1}{4} \left| -\frac{1}{2} \right\rangle_{zz} \left\langle -\frac{1}{2} \right| \right] | -\frac{1}{2} \rangle_z \\ &\quad + {}_z \langle -\frac{1}{2} | \left[ \frac{3}{4} \left( \frac{1}{\sqrt{3}} \left| \frac{1}{2} \right\rangle_z + i \frac{\sqrt{2}}{\sqrt{3}} \left| -\frac{1}{2} \right\rangle_z \right) \left( \frac{1}{\sqrt{3}} z \left\langle \frac{1}{2} \right| - i \frac{\sqrt{2}}{\sqrt{3}} z \left\langle -\frac{1}{2} \right| \right) \right] | -\frac{1}{2} \rangle_z \\ &= \frac{1}{4} + \frac{3}{4} \left( i \frac{\sqrt{2}}{\sqrt{3}} \right) \left( -i \frac{\sqrt{2}}{\sqrt{3}} \right) = \frac{1}{4} + \frac{3}{4} \frac{2}{3} \\ &= \frac{3}{4}.\end{aligned}$$

**(c)** For this part we need to find the matrix elements of  $\hat{\rho}$ ,

which we shall do in the  $\{|-\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z\}$  basis.

$$\begin{aligned} {}_z\langle \frac{1}{2} | \hat{\rho} | \frac{1}{2} \rangle_z &= {}_z\langle \frac{1}{2} | \left[ \frac{1}{4} | -\frac{1}{2} \rangle_z z \langle -\frac{1}{2} | \right. \\ &\quad \left. + \frac{3}{4} \left( \frac{1}{\sqrt{3}} | \frac{1}{2} \rangle_z + i \frac{\sqrt{2}}{\sqrt{3}} | -\frac{1}{2} \rangle_z \right) \left( \frac{1}{\sqrt{3}} z \langle \frac{1}{2} | - i \frac{\sqrt{2}}{\sqrt{3}} z \langle -\frac{1}{2} | \right) \right] | \frac{1}{2} \rangle_z \\ &= \frac{3}{4} \left( \frac{1}{\sqrt{3}} \right) \left( \frac{1}{\sqrt{3}} \right) \\ &= \frac{1}{4}, \end{aligned}$$

$$\begin{aligned} {}_z\langle \frac{1}{2} | \hat{\rho} | -\frac{1}{2} \rangle_z &= {}_z\langle \frac{1}{2} | \left[ \frac{1}{4} | -\frac{1}{2} \rangle_z z \langle -\frac{1}{2} | \right. \\ &\quad \left. + \frac{3}{4} \left( \frac{1}{\sqrt{3}} | \frac{1}{2} \rangle_z + i \frac{\sqrt{2}}{\sqrt{3}} | -\frac{1}{2} \rangle_z \right) \left( \frac{1}{\sqrt{3}} z \langle \frac{1}{2} | - i \frac{\sqrt{2}}{\sqrt{3}} z \langle -\frac{1}{2} | \right) \right] | -\frac{1}{2} \rangle_z \\ &= \frac{3}{4} \left( \frac{1}{\sqrt{3}} \right) \left( -i \frac{\sqrt{2}}{\sqrt{3}} \right) \\ &= -i \frac{\sqrt{2}}{4}, \end{aligned}$$

$$\begin{aligned} {}_z\langle -\frac{1}{2} | \hat{\rho} | \frac{1}{2} \rangle_z &= \left( {}_z\langle -\frac{1}{2} | \hat{\rho} | \frac{1}{2} \rangle_z \right)^* = \left( -i \frac{\sqrt{2}}{4} \right)^* \\ &= i \frac{\sqrt{2}}{4}, \end{aligned}$$

$${}_z\langle -\frac{1}{2} | \hat{\rho} | -\frac{1}{2} \rangle_z = \frac{3}{4} \quad \text{from part (b)}$$

and so the density matrix is

$$\rho = \begin{bmatrix} \frac{3}{4} & i \frac{\sqrt{2}}{4} \\ -i \frac{\sqrt{2}}{4} & \frac{1}{4} \end{bmatrix}.$$

The eigenvalues of this matrix are found using the usual method:

$$\begin{aligned} 0 = |\rho - \lambda \mathbf{1}| &= \begin{vmatrix} \frac{3}{4} - \lambda & i \frac{\sqrt{2}}{4} \\ -i \frac{\sqrt{2}}{4} & \frac{1}{4} - \lambda \end{vmatrix} \\ &= \left( \frac{3}{4} - \lambda \right) \left( \frac{1}{4} - \lambda \right) - \left( i \frac{\sqrt{2}}{4} \right) \left( -i \frac{\sqrt{2}}{4} \right) \\ &= \frac{3}{16} - \frac{1}{8} - \left( \frac{3}{4} + \frac{1}{4} \right) \lambda + \lambda^2 = \frac{1}{16} - \lambda + \lambda^2 \end{aligned}$$

and so the eigenvalues are

$$\lambda = \frac{1}{2} \pm \frac{\sqrt{3}}{4} = \frac{2 \pm \sqrt{3}}{4} \quad (\text{note } \sqrt{4} > \sqrt{3}, \therefore 2 > \sqrt{3})$$

which are both positive. Hence the matrix is **positive definite**.

**(d)** We need only calculate the diagonal elements of the square of the density matrix  $\rho^2$ :

$$\begin{aligned}\rho^2 &= \begin{bmatrix} \frac{3}{4} & i\frac{\sqrt{2}}{4} \\ -i\frac{\sqrt{2}}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{3}{4} & i\frac{\sqrt{2}}{4} \\ -i\frac{\sqrt{2}}{4} & \frac{1}{4} \end{bmatrix} \\ &= \begin{bmatrix} \frac{11}{16} & \dots \\ \dots & \frac{3}{16} \end{bmatrix}\end{aligned}$$

and so

$$\text{Tr}(\hat{\rho}^2) = \text{Tr}(\rho^2) = \frac{11}{16} + \frac{3}{16} = \frac{7}{8}$$

which is less than unity (and so  $\hat{\rho}$  is **mixed**).

# 3. Measurement Theory

## a. von Neumann Projective Measurements

### Measurements so far

We have already discussed measurement briefly in the sense that

insert  
1

$$P(\psi|\hat{\rho}) = \langle\psi|\hat{\rho}|\psi\rangle$$

is the probability of finding the system in the state  $|\psi\rangle$  after a measurement given that the system was initially in the state  $\hat{\rho}$ . Also the expectation value of an observable  $\hat{A}$  is given by

$$\langle\hat{A}\rangle = \text{Tr}(\hat{A}\hat{\rho})$$

for a system prepared in  $\hat{\rho}$ . The observable has a spectrum of values given by its eigenvalues  $\lambda_i$  where

$$\hat{A} = \sum_i a_i |a_i\rangle\langle a_i| ,$$

and so

$$\begin{aligned} \langle\hat{A}\rangle &= \text{Tr}\left[\left(\sum_i a_i |a_i\rangle\langle a_i|\right)\hat{\rho}\right] = \sum_i a_i \text{Tr}\left[|a_i\rangle\langle a_i|\hat{\rho}\right] \\ &= \sum_i a_i \langle a_i|\hat{\rho}|a_i\rangle = \sum_i a_i P(a_i|\hat{\rho}) . \end{aligned}$$

Thus  $\langle\hat{A}\rangle$  represents the expected (or average) value of the eigenvalues on the measurement of  $\hat{A}$ .

The expectation value is the **average**; the average is taken over many measurements. In just one instance of the measurement, the value of the observable would be any **one (and only one)** of the eigenvalues  $a_i$  and the system would be in the corresponding eigenket  $|a_i\rangle$ .

This is an example of the simplest kind of measurement – the **projective measurement**. These measurements project the state of the system onto an eigenket of the observable being measured.

- **Example (1).** The Stern-Gerlach experiment separates a beam of silver atoms, each of which is in the mixed state

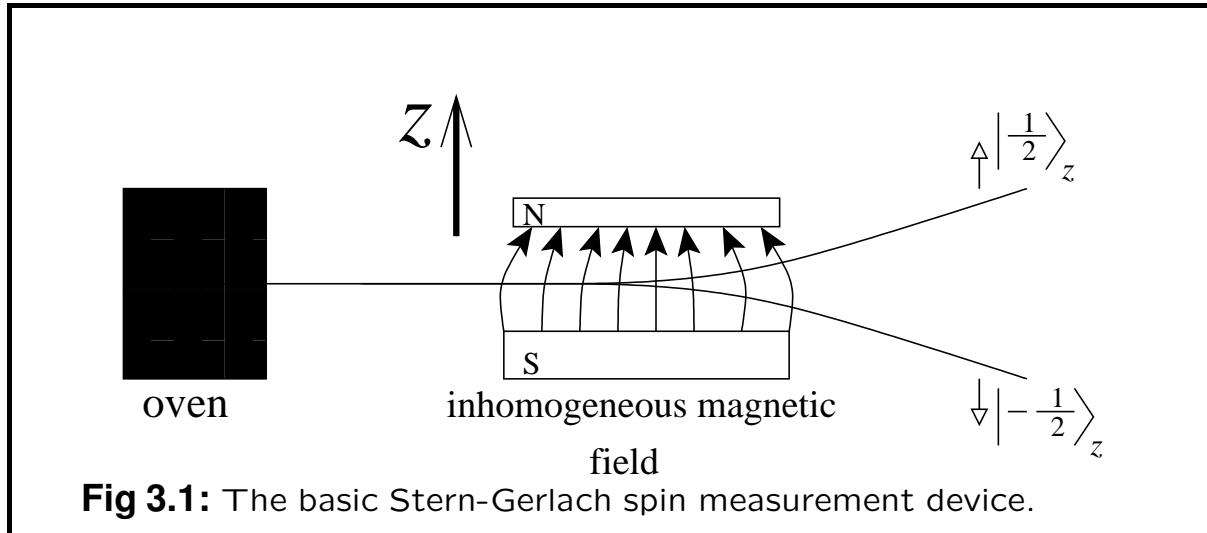
insert  
2

$$\hat{\rho} = \frac{1}{2}|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| + \frac{1}{2}|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}|$$

into two separated beams: the upper beam containing only atoms in the pure state  $|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|$  and the lower beam containing only atoms in the pure state  $|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}|$ . For any particular atom in

the beam, there is a 50% chance that it will be found in the upper beam (i.e. in the state  $|\frac{1}{2}\rangle_z$ ) and a 50% chance that it will be found in the lower beam (i.e. in the state  $|-\frac{1}{2}\rangle_z$ ). These are the **results of the measurement**. The measurement has projected the state onto the eigenbasis of the operator  $\hat{S}_z$ .

• • •



**Fig 3.1:** The basic Stern-Gerlach spin measurement device.

**Simple projection measurements.** The Stern-Gerlach experiment, as illustrated in Fig. 3.1 above, simply separates a mixed state into its pure state components. But the same device can be used to make measurements on any initial state including a pure state. This type of measurement is called a **von Neumann** measurement and is described as follows:

### **von Neumann projection measurement**

#### Pure initial state case

Initial state of the system :  $|\psi\rangle$

Measurement operator:  $\hat{A} = \sum_i a_i |a_i\rangle\langle a_i|$

Outcome  $i$  probability:  $P(a_i|\psi) = |\langle a_i|\psi\rangle|^2 = \langle\psi|a_i\rangle\langle a_i|\psi\rangle$   
final state:  $|a_i\rangle$

Averaged final state: 
$$\begin{aligned} \sum_i P(a_i|\psi)|a_i\rangle\langle a_i| \\ = \sum_i |\langle a_i|\psi\rangle|^2 |a_i\rangle\langle a_i| . \end{aligned}$$

#### General case

Initial state of the system :  $\hat{\rho}$

Measurement operator:  $\hat{A} = \sum_i a_i |a_i\rangle\langle a_i|$

Outcome  $i$  probability:  $P(a_i|\hat{\rho}) = \langle a_i|\hat{\rho}|a_i\rangle = \text{Tr}(\hat{\rho}|a_i\rangle\langle a_i|)$   
final state:  $|a_i\rangle$

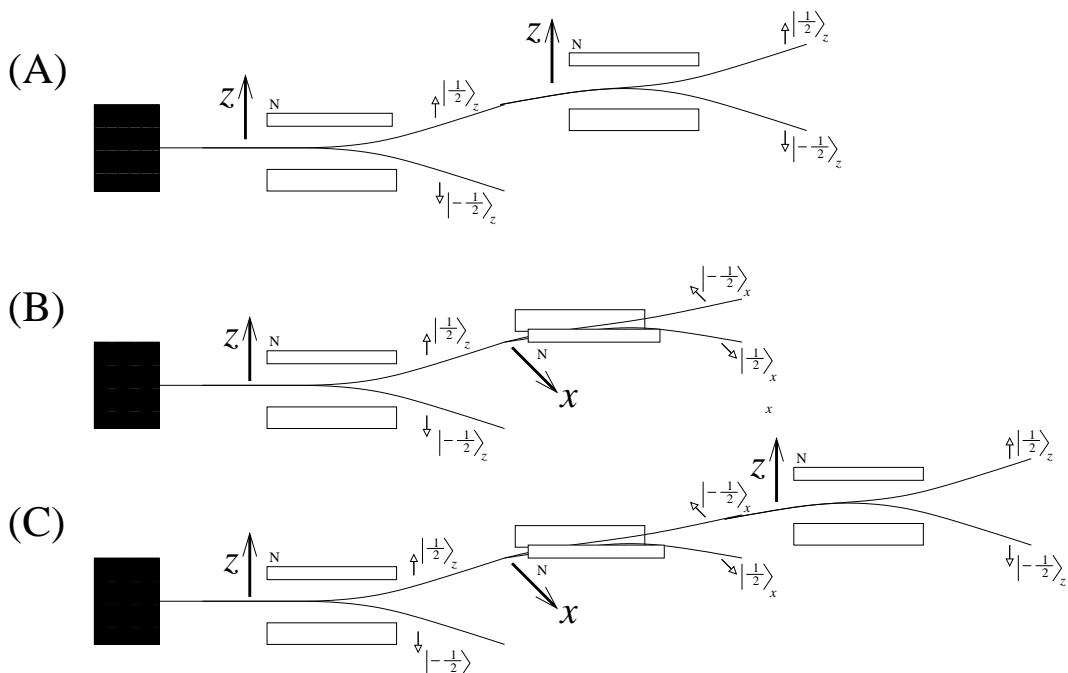
Averaged final state: 
$$\begin{aligned} \sum_i P(a_i|\hat{\rho})|a_i\rangle\langle a_i| \\ = \sum_i \langle a_i|\hat{\rho}|a_i\rangle |a_i\rangle\langle a_i| . \end{aligned}$$

### Averaged final state

You should think of the averaged final state as follows. Imagine that a measurement of  $\hat{A}$  has been performed on the system but you haven't been told the outcome. The **averaged final state** is the expected state you assign to the system given what you know about it, that is, the final state is  $|a_i\rangle$  with probability  $P(a_i|\hat{\rho})$ . Once you know of the outcome however, say for example the outcome is  $i = 4$ , you would assign the final state as being the corresponding ket  $|a_4\rangle$ . Hence, the state we assign to a system corresponds to **our knowledge** of it.

This is just the usual case in probability theory. Imagine a box contains 2 balls, one red and the other green. A person picks out a ball at random from the box but hides it from you. Your knowledge of the outcome is that the colour of the selected ball is red or green with equal probability. This is the averaged final state. However, once you are told that the colour of the selected ball is green, you **know** it is green.

Consider diagram (A) in Fig. 3.2 where a second device is placed in the upper beam. The magnet of the second device is aligned parallel to first. The measurement operator is, again,  $\hat{S}_z$ . What are the possible outcomes and their probabilities from the second device?



**Fig 3.2:** Cascading Stern-Gerlach devices.

The second device is in the upper path, which we know has atoms in the state  $|\frac{1}{2}\rangle_z$ . The measurement of  $\hat{S}_z$  will therefore **yield the result  $|\frac{1}{2}\rangle_z$  with certainty**. This is borne out by following

the procedure on page 2:

### Measurement description

Initial state of the system :

$$|\frac{1}{2}\rangle_z$$

Measurement operator:

$$\hat{S}_z = \frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| - \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}|$$

Outcome  $\frac{1}{2}\hbar$  probability:  
final state:

$$\left|z\langle\frac{1}{2}|\frac{1}{2}\rangle_z\right|^2 = 1$$

$$|\frac{1}{2}\rangle_z$$

Outcome  $-\frac{1}{2}\hbar$  probability:  
final state:

$$\left|z\langle-\frac{1}{2}|\frac{1}{2}\rangle_z\right|^2 = 0$$

$$|-\frac{1}{2}\rangle_z$$

Averaged final state:

$$1|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| + 0|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| = |\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|.$$

Next consider diagram (B) of Fig. 3.2. The magnet of the second device is rotated by  $90^\circ$  with respect to the first so that the north-south poles are aligned with the  $x$  axis. In this case the measurement operator is  $\hat{S}_x$ . The state of the silver atoms entering the second device is  $|\frac{1}{2}\rangle_z$ . Thus we have

### Measurement description

Initial state of the system :

$$|\frac{1}{2}\rangle_z$$

Measurement operator:

$$\hat{S}_x = \frac{1}{2}\hbar|\frac{1}{2}\rangle_{xx}\langle\frac{1}{2}| - \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{xx}\langle-\frac{1}{2}|$$

Outcome  $\frac{1}{2}\hbar$  probability:  
final state:

$$\left|x\langle\frac{1}{2}|\frac{1}{2}\rangle_z\right|^2 = \frac{1}{2}$$

$$|\frac{1}{2}\rangle_x$$

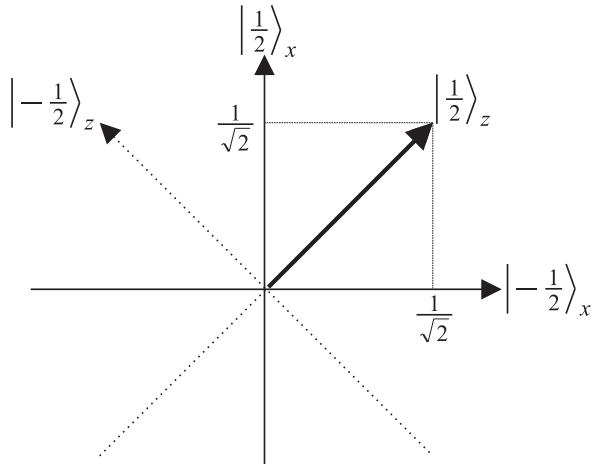
Outcome  $-\frac{1}{2}\hbar$  probability:  
final state:

$$\left|x\langle-\frac{1}{2}|\frac{1}{2}\rangle_z\right|^2 = \frac{1}{2}$$

$$|-\frac{1}{2}\rangle_x$$

Averaged final state:

$$\frac{1}{2}|\frac{1}{2}\rangle_{xx}\langle\frac{1}{2}| + \frac{1}{2}|-\frac{1}{2}\rangle_{xx}\langle-\frac{1}{2}| = \frac{1}{2}\hat{\mathbf{1}}.$$



**Fig 3.3:** The measurement of  $\hat{S}_x$  for system in the state  $| \frac{1}{2} \rangle_z = \frac{1}{\sqrt{2}} (| \frac{1}{2} \rangle_x + | -\frac{1}{2} \rangle_x)$ .

What are the possible outcomes and their probabilities if a third device is placed in the beam that is deflected along the negative  $x$  axis, as shown in diagram (C) of Fig. 3.2? The magnet of the third device is aligned with the magnet of the first device.

The orientation of the magnet (along the  $z$  axis) implies that the device again measures the  $\hat{S}_z$  operator. The magnet is in the path where atoms are known to be in the state  $| -\frac{1}{2} \rangle_x$ . Thus we have here:

#### Measurement description

Initial state of the system :	$  -\frac{1}{2} \rangle_x$
Measurement operator:	$\hat{S}_z = \frac{1}{2}\hbar  \frac{1}{2} \rangle_{zz} \langle \frac{1}{2}   - \frac{1}{2}\hbar  -\frac{1}{2} \rangle_{zz} \langle -\frac{1}{2}  $
Outcome $\frac{1}{2}\hbar$	probability: $\left  z \langle \frac{1}{2}   - \frac{1}{2} \rangle_x \right ^2 = \frac{1}{2}$
	final state: $  \frac{1}{2} \rangle_z$
Outcome $-\frac{1}{2}\hbar$	probability: $\left  z \langle -\frac{1}{2}   - \frac{1}{2} \rangle_x \right ^2 = \frac{1}{2}$
	final state: $  -\frac{1}{2} \rangle_z$
Averaged final state:	$\frac{1}{2}  \frac{1}{2} \rangle_{zz} \langle \frac{1}{2}   + \frac{1}{2}  -\frac{1}{2} \rangle_{zz} \langle -\frac{1}{2}   = \frac{1}{2} \hat{\mathbf{1}}$

The possible outcomes are  $| \frac{1}{2} \rangle_z$  and  $| -\frac{1}{2} \rangle_z$  with equal probability.

Notice that even though atoms enter the first Stern-Gerlach device in the spin state  $| \frac{1}{2} \rangle_z$  the measurement by the third Stern-Gerlach device yields outcomes  $| \frac{1}{2} \rangle_z$  and  $| -\frac{1}{2} \rangle_z$  with equal likelihood. Clearly the second Stern-Gerlach device has disturbed the

state of the atom passing through it. This effect is easily seen with polarising films and light.

## b. General measurements and POVMs

We now generalise these concepts in the following steps:

- First we rewrite the measurements we have been considering in terms of **projection operators**,
- then we consider measurements with **non-orthogonal projection operators**,
- and finally, we treat the most general measurements in terms of **POVMs**.

### Projection Operators

A projection operator is a Hermitian operator of the form

insert  
3

$$\hat{P}_S = \sum_{n \in S} |\varphi_n\rangle\langle\varphi_n|$$

where  $S$  is a set of positive integers, and  $\{|\varphi_n\rangle : n = 1, 2, \dots\}$  is an orthonormal basis. The set  $S$  selects a subset of the kets in  $\{|\varphi_n\rangle : n = 1, 2, \dots\}$ . Thus for example,

$$\begin{aligned}\hat{P}_1 &= |\varphi_1\rangle\langle\varphi_1|, & \hat{P}_2 &= |\varphi_1\rangle\langle\varphi_1| + |\varphi_5\rangle\langle\varphi_5|, \\ \hat{P}_3 &= |\varphi_2\rangle\langle\varphi_2|, & \hat{P}_4 &= |\varphi_1\rangle\langle\varphi_1| + |\varphi_3\rangle\langle\varphi_3| + |\varphi_8\rangle\langle\varphi_8| \\ \text{and} & & \hat{\mathbb{1}} &= \sum_n |\varphi_n\rangle\langle\varphi_n|\end{aligned}$$

are all projection operators. An operator  $\hat{P}$  is a projection operator if, and only if, it satisfies the following definition.

#### Definition of a projection operator

$$(\hat{P})^2 = \hat{P}. \quad (3.1)$$

From the eigenvalue equation  $\hat{P}|\lambda_i\rangle = \lambda_i|\lambda_i\rangle$  we find

$$\begin{aligned}(\hat{P})^2|\lambda_i\rangle &= \hat{P}|\lambda_i\rangle, \quad \text{and so} \\ \lambda_i^2|\lambda_i\rangle &= \lambda_i|\lambda_i\rangle.\end{aligned}$$

[Note here that  $(\hat{P})^2|\lambda_i\rangle = \hat{P}\lambda_i|\lambda_i\rangle = \lambda_i^2|\lambda_i\rangle$ .] Thus  $\lambda_i^2 - \lambda_i = \lambda_i(\lambda_i - 1) = 0$  and so the eigenvalues are either  $\lambda_i = 0$  or  $\lambda_i = 1$  and so  $\hat{P}$  is self-adjoint. Moreover  $\|\hat{P}|\psi\rangle\|^2 = \langle\psi|\hat{P}^2|\psi\rangle = \langle\psi|\hat{P}|\psi\rangle = \sum_i \lambda_i |\langle\lambda_i|\psi\rangle|^2 \leq \sum_i |\langle\lambda_i|\psi\rangle|^2 = 1$  for all  $|\psi\rangle$  in the Hilbert space. Hence projection operators are Hermitian (i.e. self-adjoint and bounded):

## Hermitian property of a projection operator

$$\hat{P}^\dagger = \hat{P} . \quad (3.2)$$

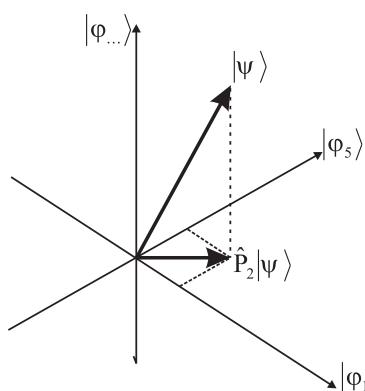
### Problem

Check that the projection operators  $\hat{P}_1$ , and  $\hat{P}_2$  satisfy Eqs. (3.1) and (3.2).

The action of a projection operator is to project states onto a **subspace**. For example, the projection operator  $\hat{P}_2$  projects onto the subspace spanned by the kets  $|\varphi_1\rangle$  and  $|\varphi_5\rangle$ . Consider the normalised state  $|\psi\rangle = \sum_n c_n |\varphi_n\rangle$ ; the action of  $\hat{P}_2$  on this state is

$$\hat{P}_2|\psi\rangle = \hat{P}_2\left(\sum_n c_n |\varphi_n\rangle\right) = \sum_n c_n \hat{P}_2|\varphi_n\rangle = c_1|\varphi_1\rangle + c_5|\varphi_5\rangle .$$

Note that the result  $(c_1|\varphi_1\rangle + c_5|\varphi_5\rangle)$  is a non-normalised state if  $(|c_1|^2 + |c_5|^2) \neq 1$ . Fig. 3.4 below illustrates the projection operation.



**Fig 3.4:** The effect of  $\hat{P}_2$  as the projection onto the  $|\varphi_1\rangle - |\varphi_5\rangle$  plane.

The action of a projection operator on a **density operator** is to transform the density operator into a mixture of projected states. That is, consider the general form of a density operator,

insert  
4

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle \langle \psi_i|$$

and an arbitrary projection operator  $\hat{P}$ .  $\hat{P}$  projects each state  $|\psi_i\rangle$  to  $\hat{P}|\psi_i\rangle$ . Thus the transformed density operator is

$$\hat{\rho}_{\text{projected}} = \sum_i p_i (\hat{P}|\psi_i\rangle) (\langle \psi_i| \hat{P}^\dagger) = \hat{P} \left( \sum_i p_i |\psi_i\rangle \langle \psi_i| \right) \hat{P}^\dagger = \hat{P} \hat{\rho} \hat{P}^\dagger .$$

The **rank** of a projection operator is the dimension of the subspace it projects onto. For example, the rank of  $\hat{P}_1$  and  $\hat{P}_3$  are both 1, the rank of  $\hat{P}_2$  is 2, the rank of  $\hat{P}_4$  is 3 and the rank of  $\hat{\mathbf{1}}$  is the dimension of the Hilbert space on which it acts. The rank is given simply by the trace of the projection operator:

$$\text{Rank}(\hat{P}) = \text{Tr}(\hat{P}) .$$

The measurement operator  $\hat{A}$  we have been considering comprises a weighted sum of rank 1 projection operators, e.g.

$$\hat{A} = \sum_i a_i \hat{P}_i, \quad \text{where} \quad \hat{P}_i = |a_i\rangle\langle a_i| .$$

In many situations, particularly in quantum information processing and quantum computing, we are not concerned with the eigenvalues of the observable. Often the eigenvalues merely serve to index the measurement outcome, and any such indexing is sufficient. For example experiments to measure the operators

$$\begin{aligned} \hat{A} &= \hat{P}_1 + 2\hat{P}_2 && (\text{eigenvalues } 1, 2) \\ \hat{A}' &= 100\hat{P}_1 + 200\hat{P}_2 && (\text{eigenvalues } 100, 200) \end{aligned}$$

leave the system in exactly the same final states. It is the **measurement projection operators** that determine the effect of the measurement on the system and **not the eigenvalues** of the measurement operators. Hence we need a formalism for quantum measurement that is in terms of projection operators only.

In place of the measurement observable, we use just the set of measurement projection operators  $\{\hat{P}_i : i = 1, 2, \dots\}$  with the condition that  $\sum_i \hat{P}_i = \hat{\mathbf{1}}$ . The state after outcome  $i$  is  $\hat{P}_i|\psi\rangle/\sqrt{p_i}$  where  $p_i = \|\hat{P}_i|\psi\rangle\|^2$  is the probability that outcome  $i$  occurs. We need the sum of the probabilities of all outcomes to be equal to 1 and so

$$1 = \sum_i p_i = \sum_i \|\hat{P}_i|\psi\rangle\|^2 = \sum_i \langle\psi|\hat{P}_i^\dagger \hat{P}_i|\psi\rangle = \langle\psi|\left(\sum_i \hat{P}_i^\dagger \hat{P}_i\right)|\psi\rangle . \quad (3.3)$$

This is satisfied because  $\sum_i \hat{P}_i^\dagger \hat{P}_i = \sum_i \hat{P}_i = \hat{\mathbf{1}}$  and we assume  $\langle\psi|\psi\rangle = 1$ .

A discussion of measurement operators involving projection operators of rank greater than 1 is given in **Appendix A** on page 20. Here, however we simply state the following.

## Measurement of projection operators

Initial state of the system :	$ \psi\rangle$
Set of projector operators:	$\{\hat{P}_i\}$ , $\sum_i \hat{P}_i = \mathbf{1}$
Outcome $i$	probability: $p_i = \ \hat{P}_i \psi\rangle\ ^2$
	final state: $ \phi_i\rangle = \hat{P}_i \psi\rangle/\sqrt{p_i}$
Averaged final state:	$\sum_i p_i  \phi_i\rangle\langle\phi_i  = \sum_i \hat{P}_i \psi\rangle\langle\psi \hat{P}_i$ .

## Non-orthogonal projection operators

There are other possibilities we should consider. The measurement projection operators we have considered so far are orthogonal in the sense that

insert  
5

$$\text{Tr}(\hat{P}_n \hat{P}_m) = 0 \quad \text{for } n \neq m .$$

This need not be the case, however. For example, consider the physical system consisting of a single photon. The photon can be polarised, for example, vertically or horizontally. The Hilbert space is spanned by the orthonormal kets  $|H\rangle$  and  $|V\rangle$  which represent horizontal and vertical polarisations states, respectively, and so  $\mathbf{1} = |H\rangle\langle H| + |V\rangle\langle V|$ . Consider the projection operators  $\hat{Q}_1$  and  $\hat{Q}_2$  where

$$\begin{aligned}\hat{Q}_1 &= |H\rangle\langle H| \\ \hat{Q}_2 &= \frac{1}{R^2+1} \left( R|H\rangle + |V\rangle \right) \left( R\langle H| + \langle V| \right)\end{aligned}\quad (3.4)$$

for  $0 < R < 1$ . It is easy to check that  $(\hat{Q}_i)^2 = \hat{Q}_i$  and so these operators are projection operators according to the definition Eq. (3.1). But  $\text{Tr}(\hat{Q}_n \hat{Q}_m) \neq 0$  for  $n \neq m$  and so they are not orthogonal.

## Unitary operators

A unitary operator  $\hat{U}$  has the following property

$$\hat{U}^\dagger = \hat{U}^{-1} \quad \text{and so } \hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \mathbf{1} .$$

Unitary operators are important because their action doesn't change the **norm of a ket**:

$$\|\hat{U}|\psi\rangle\|^2 = \langle\psi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\psi|\psi\rangle = \||\psi\rangle\|^2$$

nor the **trace of a density operator**.<sup>¶</sup>

### Problem

Show that a unitary operator does not change the trace of a density operator, i.e. show that

$$\text{Tr}(\hat{U}\hat{\rho}\hat{U}^\dagger) = \text{Tr}(\hat{\rho}) .$$

### Additional unitary operation

Finally the apparatus performing the measurement could apply a unitary operation to the state at the output. (For example, extra half-wave plates could be placed in each of the output beams of the experiment depicted in Fig. 3.8 on page 22.) Thus instead of the final (un-normalised) state for outcome  $i$  of the measurement projection operators in Eq. (3.4) being

$$\hat{P}_i|\psi\rangle$$

the measuring apparatus could additionally apply the unitary operator  $\hat{U}_i$  and so the actual final (un-normalised) state of the system becomes:

insert  
6

$$\hat{U}_i\hat{P}_i|\psi\rangle .$$

The effective overall measurement operators would therefore be  $\hat{M}_1 = \hat{U}_1\hat{P}_1$  and  $\hat{M}_2 = \hat{U}_2\hat{P}_2$ , etc. This means that the measurement operators can be quite general.

### General measurement in terms of POVMs

The full analysis of the most general type of measurement is beyond the scope of this course. We will just quote the results.

Let

$$\{\hat{M}_i : i = 1, 2, \dots\} \quad (3.5)$$

be a set of **measurement operators** and let the system be in the pure state  $|\psi\rangle$ . As before, the probability that outcome  $i = 1$  of the measurement occurs is given by  $P(i|\psi) = \langle\psi|\hat{M}_i^\dagger\hat{M}_i|\psi\rangle$  and the corresponding final state is  $\hat{M}_i|\psi\rangle/\sqrt{P(i|\psi)}$ .

<sup>¶</sup>The action on a density operator is derived from the action on the component kets. For example, consider  $\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ . The action of  $\hat{U}$  on  $|\psi_i\rangle$  produces  $|\psi'_i\rangle = \hat{U}|\psi_i\rangle$ , and so the transformed density operator is

$$\hat{\rho}' = \sum_i p_i |\psi'_i\rangle\langle\psi'_i| = \sum_i p_i \hat{U}|\psi_i\rangle\langle\psi_i|\hat{U}^\dagger = \hat{U}\hat{\rho}\hat{U}^\dagger ,$$

where we have used  $\langle\psi'_i| = |\psi'_i\rangle^\dagger = (\hat{U}|\psi_i\rangle)^\dagger = \langle\psi_i|\hat{U}^\dagger$ .

It is useful to define the set of operators:

$$\hat{E}_i = \hat{M}_i^\dagger \hat{M}_i . \quad (3.6)$$

All the eigenvalues of  $\hat{E}_i$  are positive or zero;  $\hat{E}_i$  is therefore a **positive** (or, more correctly, a **non-negative**) operator. Moreover, the normalisation of the probabilities  $P(i|\psi)$  requires that

insert

7

$$\sum_i \hat{E}_i = \hat{\mathbb{1}} . \quad (3.7)$$

The **set** of positive operators  $\{\hat{E}_i : i = 1, 2, \dots\}$  is called a **positive operator-valued measure** or **POVM**, and the operators  $\hat{E}_i$  are **POVM elements**.

### Measurement operators and POVMs

It is important to distinguish the following terms:

**measurement operator**  $\hat{M}_i$

**POVM element**  $\hat{E}_i = \hat{M}_i^\dagger \hat{M}_i$

**POVM**  $\{\hat{E}_i : i = 1, 2, \dots\}$

The only restriction on the set of measurement operators Eq. (3.5) is that they factor the POVM elements according to Eq. (3.6) and that the POVM satisfies Eq. (3.7). Such measurement operators can be quite general and need not be related at all to projection operators etc.

## General measurement

For pure states

Initial state of the system :	$ \psi\rangle$
Set of measurement operators:	$\{\hat{M}_i\}$
POVM:	$\{\hat{E}_i = \hat{M}_i^\dagger \hat{M}_i\}, \sum_i \hat{E}_i = \hat{\mathbb{1}}$
Outcome $i$	probability: $P(i \psi) = \langle \hat{E}_i \rangle$
	final state: $\frac{\hat{M}_i  \psi\rangle}{\sqrt{P(i \psi)}}$
Averaged final state:	$\sum_i \hat{M}_i  \psi\rangle \langle \psi   \hat{M}_i^\dagger .$

## General states

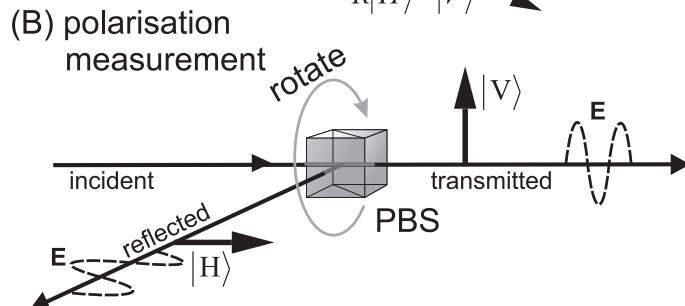
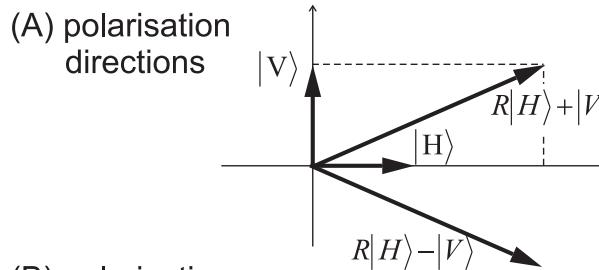
Initial state of the system :	$\hat{\rho}$
Set of measurement operators:	$\{\hat{M}_i\}$
POVM:	$\{\hat{E}_i = \hat{M}_i^\dagger \hat{M}_i\}, \sum_i \hat{E}_i = \hat{\mathbb{1}}$
Outcome $i$	probability: $P(i \hat{\rho}) = \langle \hat{E}_i \rangle$
	final state: $\frac{\hat{M}_i \hat{\rho} \hat{M}_i^\dagger}{P(i \hat{\rho})}$
Averaged final state:	$\sum_i \hat{M}_i \hat{\rho} \hat{M}_i^\dagger .$

You should compare the form of general measurements with that of the simpler von Neumann projective measurements on page 2.

Consider the following set of measurement operators

$$\begin{aligned}\hat{M}_1 &= T|H\rangle\langle H| , \\ \hat{M}_2 &= \sqrt{\frac{1}{2(R^2+1)}}(R|H\rangle + |V\rangle)(R\langle H| + \langle V|) , \\ \hat{M}_3 &= \sqrt{\frac{1}{2(R^2+1)}}(R|H\rangle - |V\rangle)(R\langle H| - \langle V|) ,\end{aligned}$$

where  $T^2 + R^2 = 1$ ,  $|H\rangle$  and  $|V\rangle$  are orthonormal kets which represent horizontal and vertical polarisations states of a single photon, respectively. They form an orthonormal basis and so  $\hat{\mathbb{1}} = |H\rangle\langle H| + |V\rangle\langle V|$ . Fig. 3.5 illustrates the horizontal and vertical polarisation directions as well as the directions corresponding to  $R|H\rangle \pm |V\rangle$ . The POVM  $\{\hat{M}_1^\dagger \hat{M}_1, \hat{M}_2^\dagger \hat{M}_2, \hat{M}_3^\dagger \hat{M}_3\}$  was the **first non-orthogonal POVM to be measured experimentally** and is described in **Appendix B** on page 22.



**Fig 3.5:** (A) Polarisation directions corresponding to  $|H\rangle$ ,  $|V\rangle$ , and  $R|H\rangle \pm |V\rangle$ ; (B) The use of a polarisation beam splitter (PBS) to separate orthogonal polarisations. By rotating the PBS as shown the polarisation of the transmitted light can be made to correspond to the  $R|H\rangle \pm |V\rangle$  directions.

### Problem

Check that the POVM  $\{\hat{E}_i = \hat{M}_i^\dagger \hat{M}_i : i = 1, 2, 3\}$  comprising the measurement operators  $\hat{M}_i$  satisfies  $\sum_i \hat{E}_i = \hat{\mathbb{1}}$ .

### Problem

Consider a system whose Hilbert space is spanned by the 3 orthonormal states  $\{|\varphi_1\rangle, |\varphi_2\rangle, |\varphi_3\rangle\}$ . A measurement is to be performed on the system, however the set of measurement operators is incomplete. The following elements are known,

$$\begin{aligned}\hat{M}_1 &= \frac{1}{2}|\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| , \\ \hat{M}_2 &= \frac{1}{2}|\varphi_1\rangle\langle\varphi_1| + \frac{1}{\sqrt{2}}|\varphi_3\rangle\langle\varphi_3| , \\ \hat{M}_3 &= \frac{1}{\sqrt{8}}\left(|\varphi_1\rangle + |\varphi_3\rangle\right)\left(\langle\varphi_1| + \langle\varphi_3|\right) ,\end{aligned}\quad (3.8)$$

but the fourth element  $\hat{M}_4$  is yet to be determined.

- Calculate the POVM associated with the measurement.
- Hence find a possible solution for  $\hat{M}_4$ .
- Write down the format for the general solution for  $\hat{M}_4$ .

[Answers:  $\text{POVM} = \left\{ \frac{1}{4}|\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2|, \frac{1}{4}|\varphi_1\rangle\langle\varphi_1| + \frac{1}{2}|\varphi_3\rangle\langle\varphi_3|, \frac{1}{4}\left(|\varphi_1\rangle + |\varphi_3\rangle\right)\left(\langle\varphi_1| + \langle\varphi_3|\right), \frac{1}{4}\left(|\varphi_1\rangle - |\varphi_3\rangle\right)\left(\langle\varphi_1| - \langle\varphi_3|\right) \right\}$ , one possible solution is  $\hat{M}_4 = \frac{1}{\sqrt{8}}\left(|\varphi_1\rangle - |\varphi_3\rangle\right)\left(\langle\varphi_1| - \langle\varphi_3|\right)$ , the general solution is  $\hat{U}\hat{M}_4$  for arbitrary unitary operator  $\hat{U}$ .]

## c. Quantum Cryptography

We now review an application of quantum measurement in quantum cryptography. This area encompasses the use of quantum physics to ensure the authenticity and security of information. The first application was in *quantum key distribution* in 1984 by Bennett and Brassard. Their protocol is now known as the BB84 protocol.

**One-time pad.** The most secure way to secretly send a message is by using a **one-time pad**. A one-time pad is a list of random numbers, say  $\{R_i : i = 1, 2, \dots\}$  where  $0 \leq R_i \leq 27$ . The sender, Alice, and receiver, Bob, arrange to each have an identical copy of the pad, and ensure that no one else has a copy. Alice uses a numerical representation of the letters of her message, such as 'A'=0, 'B'=1, ..., 'Z'=25, ' '=26, '.'=27. Thus her message would consist of the sequence of numbers  $L_1, L_2, \dots$ . She adds one random number from the pad to each letter to generate the **encrypted** message  $E_1, E_2, \dots$  where  $E_i = (L_i + R_i) \bmod 28$ . The modulo operation is given by

$$x \bmod y = \begin{cases} x + y & \text{if } x < 0 \\ x & \text{if } 0 \leq x < y \\ x - y & \text{if } x \geq y \end{cases}$$

and so

$x$	$x \bmod 28$
0	0
1	1
$\vdots$	$\vdots$
27	27
28	0
29	1
$\vdots$	$\vdots$

The modulo 28 operation ensures that the value of  $E_i$  is between 0 and 27. The encrypted message is sent to Bob by any means, e.g. broadcast over public radio, published in a newspaper, etc. Without having a copy of the one-time pad the message cannot be decrypted. Indeed the encrypted message appears to be a set of **random numbers**. Bob, however, has a copy of the pad. He performs the subtraction for each encrypted letter  $(E_i - R_i) \bmod 28$ . The result is the original message. Alice has sent her message to Bob with absolute security.

As an example, consider the message “NO” for which  $L_1 = 13$  and  $L_2 = 14$ . Let the one-time pad be  $R_1 = 5$ ,  $R_2 = 20$ . Alice’s encryption is

$$\begin{aligned} E_1 &= L_1 + R_1 = (13 + 5) \bmod 28 = 18 \\ E_2 &= L_2 + R_2 = (14 + 20) \bmod 28 = 34 - 28 = 6 , \end{aligned}$$

and Bob’s decryption is

$$\begin{aligned} L_1 &= E_1 - R_1 = (18 - 5) \bmod 28 = 13 \\ L_2 &= E_2 - R_2 = (6 - 20) \bmod 28 = -14 + 28 = 14 . \end{aligned}$$

However, in order for this protocol to work, Alice and Bob need to have identical copies of the one-time pad, and no one else must have access to the pad. The problem of distributing the pad securely (i.e. secretly), which is also called a **key**, is called **key distribution**. Bennett and Brassard’s protocol from 1984 is a quantum protocol for key distribution.

**BB84.** An optical channel for sending polarised photons between Alice and Bob is set up. Two bases are chosen:

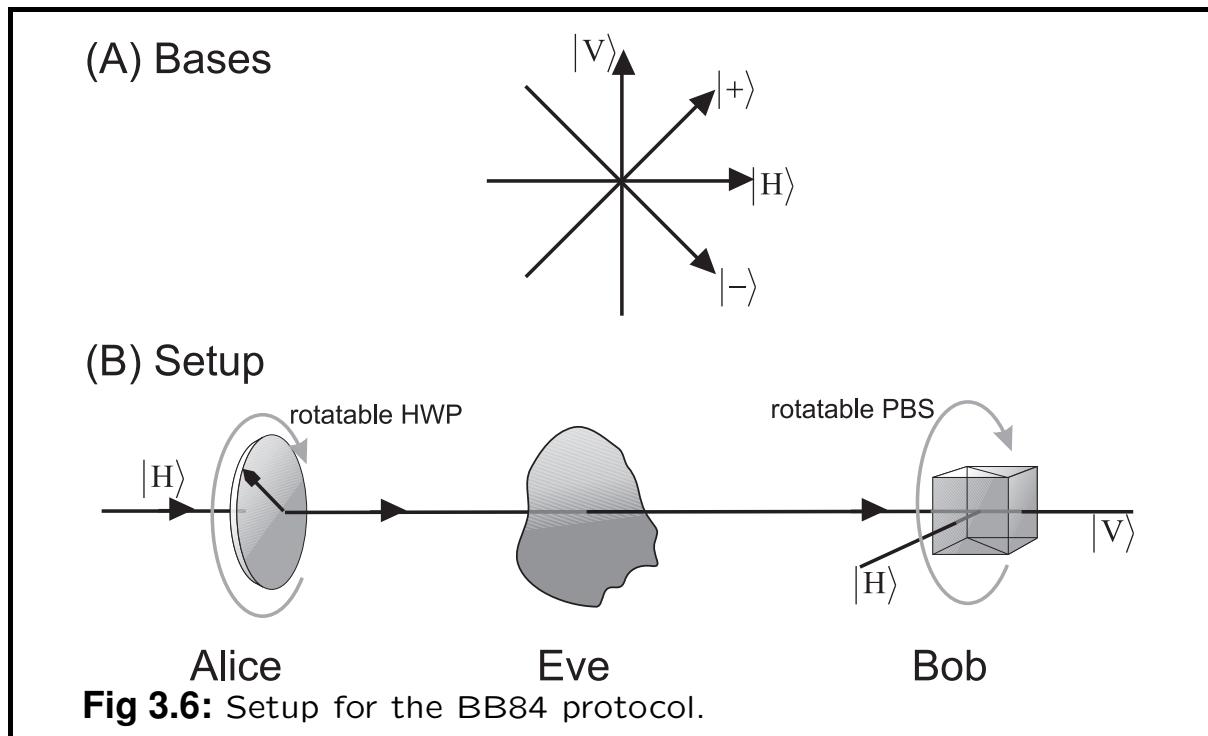
$$\begin{aligned} B_{\text{HV}} &= \{|H\rangle, |V\rangle\} \\ B_{\pm} &= \{|+\rangle, |-\rangle\} \end{aligned}$$

where

$$\begin{aligned} |\text{H}\rangle &= \text{horizontal polarisation ,} \\ |\text{V}\rangle &= \text{vertical polarisation ,} \\ |\text{+}\rangle &= \frac{1}{\sqrt{2}}(|\text{H}\rangle + |\text{V}\rangle) , \\ |-\rangle &= \frac{1}{\sqrt{2}}(|\text{H}\rangle - |\text{V}\rangle) . \end{aligned}$$

The figure below shows the relation between the two bases. Two observables are chosen:

$$\begin{aligned} \hat{A}_{\text{HV}} &= 0|\text{H}\rangle\langle\text{H}| + 1|\text{V}\rangle\langle\text{V}| = |\text{V}\rangle\langle\text{V}| , \\ \hat{A}_{\pm} &= 0|-\rangle\langle-| + 1|+\rangle\langle+| = |+\rangle\langle+| . \end{aligned}$$



Alice has a source of horizontally polarised photons. She orients a half-wave plate (HWP) to rotate the polarisation and generate one of the 4 states:  $\{|\text{H}\rangle, |\text{V}\rangle, |+\rangle, |-\rangle\}$ . Bob makes measurements using a polarisation beam splitter (PBS). A PBS transmits vertically polarised light and reflects horizontally polarised light. The orientation of the PBS in the figure will allow him to measure  $\hat{A}_{\text{HV}}$ . He rotates the PBS by 45% to measure  $\hat{A}_{\pm}$ . Eve is a potential eavesdropper. She would make her own measurements and retransmit photons to Bob.

The protocol proceeds as follows:

- (1) Alice sends  $N$  photons to Bob. For the  $i$ -th photon, Alice randomly chooses one state  $|\psi_i\rangle$  from the set  $\{|\text{H}\rangle, |\text{V}\rangle, |+\rangle, |-\rangle\}$ . She prepares the photon to be in state  $|\psi_i\rangle$ , and then she sends the photon to Bob. She records the preparation by

setting  $P_i$  as follows:

$$P_i = \begin{cases} 0 & \text{if } |\psi_i\rangle = |\text{H}\rangle \text{ or } |\psi_i\rangle = |-\rangle \\ 1 & \text{if } |\psi_i\rangle = |\text{V}\rangle \text{ or } |\psi_i\rangle = |+\rangle \end{cases} .$$

- (2) On receiving each photon, Bob randomly chooses an observable from  $\hat{A}_{\text{HV}}$  or  $\hat{A}_{\pm}$  and measures the photon accordingly. He records the outcome of the measurement, i.e. the eigenvalue of the observable (0 or 1), as  $O_i$  for each measurement.
- (3) Alice announces publicly which basis she used for the state of each photon, either  $B_{\text{HV}}$  if she sent  $|\text{H}\rangle$  or  $|\text{V}\rangle$ , or  $B_{\pm}$  if she sent  $|+\rangle$  or  $|-\rangle$ .
- (4) Bob publicly announces which photons he measured in the same basis as used by Alice, that is, when he measured  $B_{\text{HV}}$  and Alice used  $\hat{A}_{\text{HV}}$  or when he measured  $B_{\pm}$  and Alice used  $\hat{A}_{\pm}$ . Alice and Bob disregard all other photon events. They have about  $N/2$  photon events left. If no eavesdropping occurred, the result of Bob's measurements for these photons,  $O_i$ , would match exactly the state prepared by Alice,  $P_i$ , because they used the same basis.
- (5) Alice and Bob perform a **security check for eavesdropping**. Alice selects a random subset of  $M$  of the remaining photon events, and announces publicly the preparation  $P_i$  for the photon. Bob announces publicly the corresponding outcome  $O_i$  of his measurement for the same photon. If these values do not agree then they know an eavesdropper has interfered with the channel and **they abort the procedure**.
- (6) The cryptographic key (i.e. for the one-time pad) is the corresponding values of  $P_i$  and  $O_i$  that Alice and Bob have for the remaining photon events. If the security check is successful, there is a high probability that Alice and Bob have distributed a secure key. This probability can be made to approach 1 by increasing the size of the security check  $M$ .

### Detecting an attack

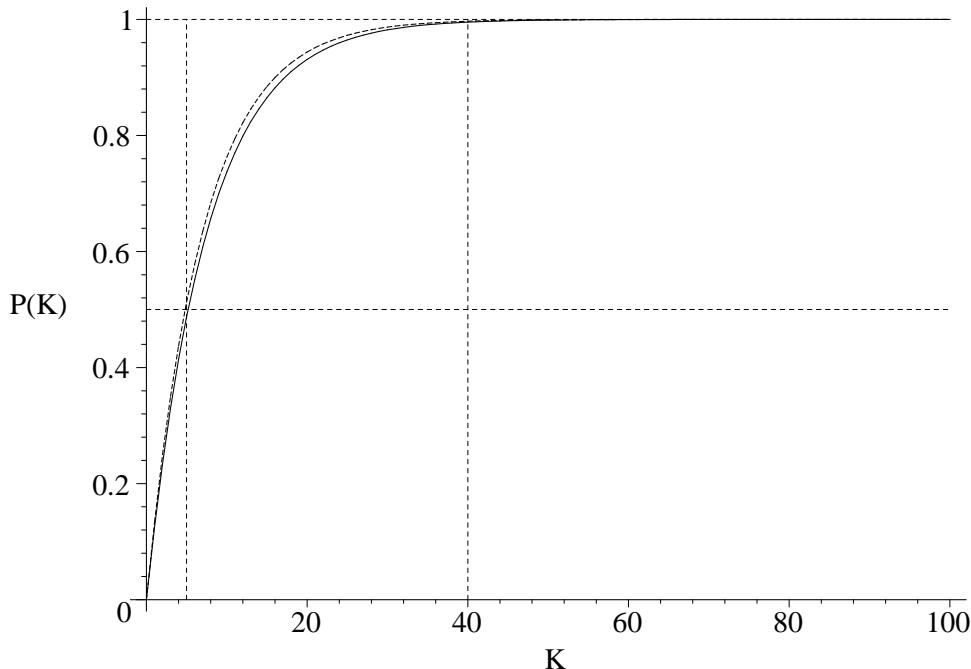
If an eavesdropper, Eve, tries to measure the state of a photon in the channel she cannot avoid disturbing the polarisation state at least some of the time. For example, as Eve doesn't know what basis was used in the preparation, she will choose the wrong observable to measure 50% of the time. This means that when she re-sends the photon its state will be different to what Alice sent. This disturbance is very likely to be detected in the security check, provided the number of photon events checked  $M$  is large enough.

There have been further refinements to the protocol. Even if an eavesdropper is present, there are protocols called **privacy amplification** that allow Alice and Bob to sacrifice more photon events to reduce the amount of information the eavesdropper has about the remaining events.

Quantum key distribution devices are now on the market. See, for example, the US firm MagiQ (<http://www.magiqtech.com/>) and the Swiss firm id Quantique (<http://www.idquantique.com/>)

The probability  $P(K)$  of detecting Eve's eavesdropping if Eve makes measurements on  $K$  photons,  $N'$  photon events remain after step **(4)** and  $M$  photons are chosen for the security check in step **(5)** is derived in Appendix C on page 25. The final result is given in Eq. (3.10). The figure below (Fig. 3.7) shows the probability of detection  $P(K)$  as a function of  $K$  for  $N' = 1000$  and  $M = 500$ . If the protocol is successful it will yield a secure key of 500 bits (i.e. key length is  $N' - M = 500$ ).

A rough estimate of  $P(K)$  can also be derived as follows. Not all of the  $K$  photons that Eve measured will be selected for the security check. The average number  $\bar{K}$  that will be selected is given by  $\bar{K} = K \frac{M}{N'}$ . The probability that her activity will escape detection is  $(\frac{3}{4})^{\bar{K}}$ . Thus an estimate of the probability of being detected is  $P(K) \approx 1 - (\frac{3}{4})^{\bar{K}} = 1 - (\frac{3}{4})^{KM/N'}$ . This approximation is shown as a dashed curve in the figure.



**Fig 3.7:** Probability of detecting an eavesdropper.

If Eve measures more than about 40 photons, it is very likely that her malicious activities will be detected. However if she only makes 5 measurements, there is about an even chance she will not. Of these 5 measurements, on average 2.5 will be used by the security check (as  $M = 500$  and  $N' = 1000$ ) leaving her with about 2.5 photons in which to extract information. However, in half of these she will have used the wrong bases. That leaves her with, on average, about 1 bit of information about the 500 bit key. If she tries to gain more information than this she risks being detected. Moreover, privacy amplification can reduce the information she has, at the expense of reducing the key length.

## Appendix A – Degenerate projective measurements

We consider here situations where the measurement comprises projection operators of rank higher than 1. These occur when the spectrum of eigenvalues of the measurement operator are degenerate. For example consider the measurement operator  $\hat{B}$ :

$$\begin{aligned}\hat{B} &= a|a_1\rangle\langle a_1| + b|a_2\rangle\langle a_2| + b|a_3\rangle\langle a_3| \\ &= a|a_1\rangle\langle a_1| + b(|a_2\rangle\langle a_2| + |a_3\rangle\langle a_3|) \\ &= a\hat{P}_1 + b\hat{P}_2\end{aligned}$$

where  $\hat{P}_1 = |a_1\rangle\langle a_1|$  is a rank 1 projection operator and  $\hat{P}_2 = (|a_2\rangle\langle a_2| + |a_3\rangle\langle a_3|)$  is a rank 2 projection operator. Getting observable value  $a$  indicates that the final state is projected onto  $|a_1\rangle$ . However, the observable value  $b$  does not distinguish between the states  $|a_2\rangle$  or  $|a_3\rangle$ . In fact, on getting this outcome the state of the system is projected by  $\hat{P}_2$  onto the the subspace spanned by  $|a_2\rangle$  or  $|a_3\rangle$ .

For example, let the initial state of the system be given by

$$|\psi\rangle = \alpha|a_1\rangle + \beta|a_2\rangle + \gamma|a_3\rangle$$

with  $|\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1$  for normalisation. Assume that the Hilbert space is 3 dimensional. Thus,

$$\hat{\mathbf{1}} = |a_1\rangle\langle a_1| + |a_2\rangle\langle a_2| + |a_3\rangle\langle a_3| = \hat{P}_1 + \hat{P}_2 ,$$

and so, for later reference,

$$\hat{P}_2 = \hat{\mathbf{1}} - \hat{P}_1 . \quad (3.9)$$

Consider making a measurement of  $\hat{B}$ . The outcome for observable value  $a$  is straightforward: the probability of it occurring is

$P(a|\psi) = |\langle a_1|\psi\rangle|^2 = \langle\psi|a_1\rangle\langle a_1|\psi\rangle = \text{Tr}(\hat{P}_1|\psi\rangle\langle\psi|) = \langle\hat{P}_1\rangle = |\alpha|^2$  and the final state is  $|a_1\rangle$ . Now for the outcome for observable value  $b$ . The probability of it occurring is

$$\begin{aligned}P(b|\psi) &= 1 - P(a|\psi) = \langle\hat{\mathbf{1}}\rangle - \langle\hat{P}_1\rangle = \langle\hat{\mathbf{1}} - \hat{P}_1\rangle \\ &= \langle\hat{P}_2\rangle \quad [\text{from Eq. (3.9)}] \\ &= \text{Tr}(\hat{P}_2|\psi\rangle\langle\psi|) = \langle\psi|\hat{P}_2|\psi\rangle = |\beta|^2 + |\gamma|^2\end{aligned}$$

and the final state is given by **projecting**  $|\psi\rangle$  onto the subspace spanned by  $\{|a_2\rangle, |a_3\rangle\}$ , i.e.

$$|\phi'\rangle = \hat{P}_2|\psi\rangle = \beta|a_2\rangle + \gamma|a_3\rangle ,$$

and **normalising** the result:

$$|\phi''\rangle = \frac{\hat{P}_2|\psi\rangle}{N}$$

where  $N = \|\phi'\| = \sqrt{\langle\phi'|\phi'\rangle} = \sqrt{\langle\psi|\hat{P}_2^\dagger\hat{P}_2|\psi\rangle} = \sqrt{\langle\psi|\hat{P}_2|\psi\rangle} = \sqrt{P(b|\psi)}$ , i.e., the final state is

$$\frac{\hat{P}_2|\psi\rangle}{\sqrt{P(b|\psi)}} = \frac{\beta|a_2\rangle + \gamma|a_3\rangle}{\sqrt{|\beta|^2 + |\gamma|^2}}.$$

Thus the situation for measuring operators involving projection operators of arbitrary rank is:

### For pure states

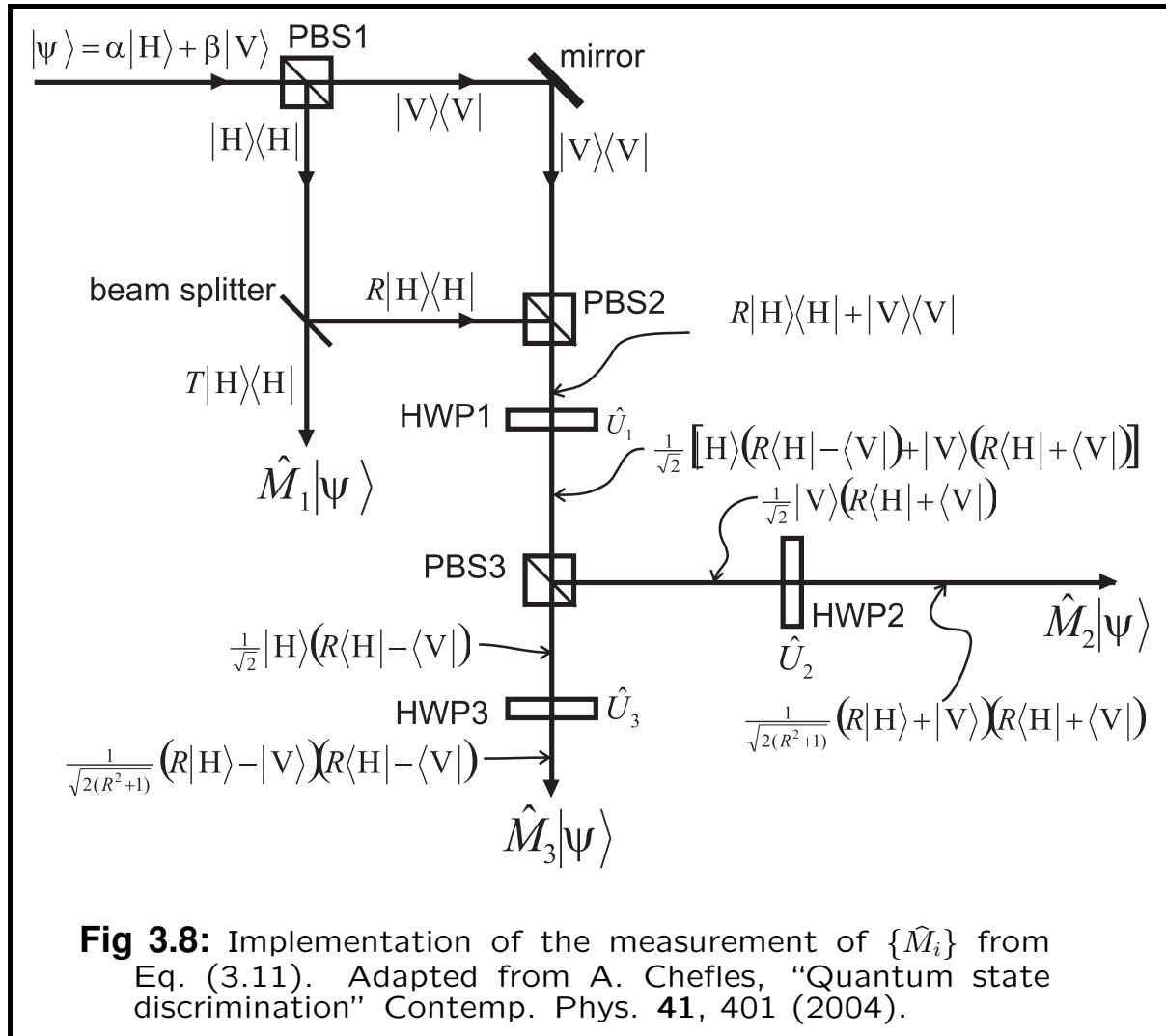
Initial state of the system :	$ \psi\rangle$
Measurement operator:	$\hat{A} = \sum_i a_i \hat{P}_i$
Outcome $i$	probability: $P(a_i \psi) = \langle\hat{P}_i\rangle$
	final state: $\frac{\hat{P}_i \psi\rangle}{\sqrt{P(a_i \psi)}}$
Averaged final state:	$\sum_i \hat{P}_i \psi\rangle\langle\psi \hat{P}_i^\dagger$ .

### General states

Initial state of the system :	$\hat{\rho}$
Measurement operator:	$\hat{A} = \sum_i a_i \hat{P}_i$
Outcome $i$	probability: $P(a_i \hat{\rho}) = \langle\hat{P}_i\rangle$
	final state: $\frac{\hat{P}_i\hat{\rho}\hat{P}_i^\dagger}{P(a_i \hat{\rho})}$
Averaged final state:	$\sum_i \hat{P}_i\hat{\rho}\hat{P}_i^\dagger$ .

## Appendix B – Realisation of a non-orthogonal projective measurement

The figure below shows an implementation of the set of measurement operators  $\{\hat{M}_1, \hat{M}_2, \hat{M}_3\}$  given in Eq. (3.11). A detailed understanding of the figure is not required for this course – it is included as an example of how non-orthogonal projective measurement can be implemented experimentally.



The various elements in the figure are:

- *beam splitter* – a semi-silvered mirror with transmission and reflection probabilities  $T^2$  and  $R^2$ , respectively. Note  $T^2 + R^2 = 1$ .
- *PBS $n$*  – polarisation beam splitters. They separate the two polarisation modes: vertical polarisation ( $|V\rangle$ ) is transmitted and horizontal polarisation ( $|H\rangle$ ) is reflected.
- *mirror*. A 100% reflecting mirror.
- *HWP $n$*  – half-wave plates. Their function here is to rotate the polarisation of the light.

A photon in the polarisation state  $|\psi\rangle$  enters the apparatus from the upper left. PBS1 separates the horizontal and vertical polarisation components. The horizontal component is directed downwards to the beam splitter. The state of the transmitted light is given by  $T|H\rangle\langle H||\psi\rangle = \hat{M}_1|\psi\rangle$ . A photon detected here corresponds to outcome  $i = 1$ .

The reflected light from the beam splitter is recombined with the vertically polarised component at PBS2. The state of the light at this point is

$$\left( R|H\rangle\langle H| + |V\rangle\langle V| \right) |\psi\rangle .$$

HWP1 applies the unitary operation  $\hat{U}_1$  such that

$$\begin{aligned} \hat{U}_1|H\rangle &= \frac{1}{\sqrt{2}} \left( |H\rangle + |V\rangle \right) \\ \hat{U}_1|V\rangle &= -\frac{1}{\sqrt{2}} \left( |H\rangle - |V\rangle \right) . \end{aligned}$$

and so the state of the light after HWP1 becomes

$$\begin{aligned} &\hat{U}_1 \left( R|H\rangle\langle H| + |V\rangle\langle V| \right) |\psi\rangle \\ &= \left( R\hat{U}_1|H\rangle\langle H| + \hat{U}_1|V\rangle\langle V| \right) |\psi\rangle \\ &= \left[ R\frac{1}{\sqrt{2}} \left( |H\rangle + |V\rangle \right) \langle H| - \frac{1}{\sqrt{2}} \left( |H\rangle - |V\rangle \right) \langle V| \right] |\psi\rangle \\ &= \frac{1}{\sqrt{2}} \left[ |H\rangle \left( R\langle H| - \langle V| \right) + |V\rangle \left( R\langle H| + \langle V| \right) \right] |\psi\rangle . \end{aligned}$$

PBS3 separates the horizontal and vertical polarisation components so the state of the light in the two output ports is

$$\begin{aligned} &\frac{1}{\sqrt{2}} \left[ |H\rangle \left( R\langle H| - \langle V| \right) \right] |\psi\rangle \\ &\frac{1}{\sqrt{2}} \left[ |V\rangle \left( R\langle H| + \langle V| \right) \right] |\psi\rangle , \end{aligned}$$

respectively. Finally HWP2 and HWP3 apply the unitary operations  $\hat{U}_2$  and  $\hat{U}_3$ , respectively, where

$$\begin{aligned} \hat{U}_2|H\rangle &= \frac{1}{R^2+1} \left( R|H\rangle - |V\rangle \right) \\ \hat{U}_3|V\rangle &= \frac{1}{R^2+1} \left( R|H\rangle + |V\rangle \right) . \end{aligned}$$

Thus the state of the light after HWP2 and HWP3 is

$$\begin{aligned}\hat{U}_2 \frac{1}{\sqrt{2}} & \left[ |\text{H}\rangle \left( R\langle \text{H}| - \langle \text{V}| \right) \right] |\psi\rangle \\ &= \frac{1}{2(R^2+1)} \left[ \left( R|\text{H}\rangle - |\text{V}\rangle \right) \left( R\langle \text{H}| - \langle \text{V}| \right) \right] |\psi\rangle \\ &= \hat{M}_2 |\psi\rangle ,\end{aligned}$$

and

$$\begin{aligned}\hat{U}_3 \frac{1}{\sqrt{2}} & \left[ |\text{H}\rangle \left( R\langle \text{H}| + \langle \text{V}| \right) \right] |\psi\rangle \\ &= \frac{1}{2(R^2+1)} \left[ \left( R|\text{H}\rangle + |\text{V}\rangle \right) \left( R\langle \text{H}| + \langle \text{V}| \right) \right] |\psi\rangle \\ &= \hat{M}_3 |\psi\rangle ,\end{aligned}$$

respectively. Detecting the photon in these positions corresponds to outcomes  $i = 2$  and  $i = 3$  respectively.

## Appendix C – Detecting an eavesdropper

Imagine Eve has made a measurement (using  $\hat{A}_{\text{HV}}$  or  $\hat{A}_{\pm}$ ) on the 10th photon. Let's calculate the probability that Eve's measurement will **not** be detected if this photon is chosen for the security check.

Eve avoids detection in two ways. The first way is that she could (A) be using the **same measurement basis** as the security check. There is a  $P(A) = \frac{1}{2}$  (50%) chance that this happens. The second way is that Eve could (B) use a different measurement basis and (C) by chance not be detected. There is a  $P(B) = \frac{1}{2}$  chance she uses a different basis. Imagine Alice has sent  $|V\rangle$  but Eve resent  $|+\rangle$ . There is a  $P(C|B) = \frac{1}{2}$  (50%) chance that Bob's measurement will yield  $|V\rangle$ , in which case Eve escapes detection. The probability she escapes detection this way is the probability that she chooses (B) a different basis **and** (C) Bob still measures the correct state, i.e.

$$P(\text{B and C}) = P(\text{B})P(\text{C}|B)$$

The total probability for escaping detection is therefore

$$\begin{aligned} P(\text{total}) &= P(A) + P(\text{B and C}) = P(A) + P(\text{B})P(\text{C}|B) \\ &= \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \\ &= \frac{3}{4}. \end{aligned}$$

We now calculate the probability that Eve's disturbance will **not** be detected if Eve makes measurements on  $k$  of the photons in the security check. Let  $E_i$  represent the event that she escapes detection for the  $i$ th photon. The probability that she escapes detection in all  $k$  events is

$$P(E_1 \text{ and } E_2 \text{ and } E_3 \dots) = P(E_1)P(E_2)P(E_3)\dots$$

because each of the events  $E_1, E_2, \dots, E_k$  are independent. They each have the same probability of  $\frac{3}{4}$  each. Thus the probability she escapes detection in all  $k$  events is

$$P(\text{escape after measuring } k \text{ checked photons}) = \left(\frac{3}{4}\right)^k.$$

Let's assume that  $N'$  photon events remain after step **(4)** of the protocol and that Eve made measurements on  $K$  of these photons. We now calculate the probability that Eve's measurements **will not** be detected in the security check on  $M$  randomly chosen photon events, for  $K \leq M$ .

The probability of escaping detection is the probability of choosing

0 checked photons **and** escape detection, and  
 1 checked photon **and** escape detection, and  
 2 checked photons **and** escape detection, and  
 $\vdots$   
 $K$  checked photons **and** escape detection.

Each of the lines represents an independent event and so the total probability is the sum of the probability of each event:

$$P_{\text{escape}}(K) = \sum_{k=0}^K P(\text{choosing } k \text{ checked photons and escape detection})$$

where

$$\begin{aligned} P(\text{choosing } k \text{ checked photons and escape detection}) &= P(\text{choosing } k \text{ checked photons}) \\ &\quad \times P(\text{escape detection on measuring } k \text{ checked photons}) \end{aligned}$$

The probability of choosing  $k$  checked photons is

$$P(\text{choosing } k \text{ checked photons}) = \frac{(\# \text{ ways to choose } k \text{ from } M)(\# \text{ ways to choose remainder})}{(\# \text{ ways to choose } K \text{ from } N')}$$

where

$$\begin{aligned} \# \text{ ways to choose } k \text{ from } M &= \binom{M}{k} \\ \# \text{ ways to choose remainder} &= \# \text{ ways to choose } (K - k) \text{ from } (N' - M) \\ &= \binom{N' - M}{K - k} \\ \# \text{ ways to choose } K \text{ from } N' &= \binom{N'}{K} \end{aligned}$$

Here  $\binom{N}{n} = \frac{N!}{n!(N-n)!} = {}^N C_n$  is the binomial coefficient. Thus

$$P(\text{choosing } k \text{ checked photons}) = \frac{\binom{M}{k} \binom{N' - M}{K - k}}{\binom{N'}{K}}.$$

Using

$$P(\text{escape after measuring } k \text{ checked photons}) = \left(\frac{3}{4}\right)^k$$

from above, we now have

$$\begin{aligned} P(\text{choosing } k \text{ checked photons and escape detection}) &= \frac{\binom{M}{k} \binom{N' - M}{K - k}}{\binom{N'}{K}} \times \left(\frac{3}{4}\right)^k. \end{aligned}$$

Hence we find the probability of Eve escaping detection when she

measures  $K$  photons is

$$P_{\text{escape}}(K) = \sum_{k=0}^K \frac{\binom{M}{k} \binom{N'-M}{K-k}}{\binom{N'}{K}} \left(\frac{3}{4}\right)^k.$$

Finally the probability of Eve's activity being detected is just the complement of the previous result,

$$\begin{aligned} P(K) &= 1 - P_{\text{escape}}(K) \\ &= 1 - \sum_{k=0}^K \frac{\binom{M}{k} \binom{N'-M}{K-k}}{\binom{N'}{K}} \left(\frac{3}{4}\right)^k. \end{aligned} \quad (3.10)$$

# Problem Sheet

- (1) Given that  $\{|\varphi_n\rangle : n = 1, 2, \dots\}$  is an orthonormal basis, check that the operators

$$\begin{aligned}\hat{P}_1 &= |\varphi_1\rangle\langle\varphi_1|, \\ \hat{P}_2 &= |\varphi_1\rangle\langle\varphi_1| + |\varphi_5\rangle\langle\varphi_5|\end{aligned}$$

satisfy the properties of a projection operator in Eqs. (3.1) and (3.2).

- (2) The state of a spin-1 particle is given by

$$|\psi\rangle = \frac{1}{2}|-1\rangle_z + \frac{1}{\sqrt{2}}|0\rangle_z + \frac{1}{2}|1\rangle_z.$$

Calculate the final states and their probability if a measurement of the operator

$$\hat{O} = |-1\rangle_{zz}\langle -1| + 2\left(|0\rangle_{zz}\langle 0| + |1\rangle_{zz}\langle 1|\right)$$

is made on the particle where  $\{|-1\rangle_z, |0\rangle_z, |1\rangle_z\}$  is an orthonormal set of eigenkets of the  $z$  component of spin  $\hat{S}_z$ .

**[HINTS:**

- Note that  $\hat{O}$  is **not** equal to  $\hat{S}_z$  which has the following expansion

$$\begin{aligned}\hat{S}_z &= -\hbar|-1\rangle_{zz}\langle -1| + 0|0\rangle_{zz}\langle 0| + \hbar|1\rangle_{zz}\langle 1| \\ &= -\hbar|-1\rangle_{zz}\langle -1| + \hbar|1\rangle_{zz}\langle 1|\end{aligned}$$

and eigenvalues  $-\hbar, 0, \hbar$ . In contrast, the eigenvalues of  $\hat{O}$  are just 1 and 2.

- First identify the projection operators associated with each eigenvalue. Then determine the outcomes of the measurement using these projection operators.]

- (3) Show that a unitary operator does not change the trace of a density operator, i.e. show that

$$\text{Tr}\left(\hat{U}\hat{\rho}\hat{U}^\dagger\right) = \text{Tr}(\hat{\rho}).$$

- (4) Check that the POVM  $\{\hat{E}_i = \hat{M}_i^\dagger \hat{M}_i : i = 1, 2, 3\}$  comprising the measurement operators defined on page 12 satisfies  $\sum_i \hat{E}_i = \hat{\mathbb{1}}$ .

- (5) Consider a system whose Hilbert space is spanned by the 3 orthonormal states  $\{|\varphi_1\rangle, |\varphi_2\rangle, |\varphi_3\rangle\}$ . A measurement is to be performed on the system, however the set of measurement

operators is incomplete. The following elements are known,

$$\begin{aligned}\hat{M}_1 &= \frac{1}{2}|\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| , \\ \hat{M}_2 &= \frac{1}{2}|\varphi_1\rangle\langle\varphi_1| + \frac{1}{\sqrt{2}}|\varphi_3\rangle\langle\varphi_3| , \\ \hat{M}_3 &= \frac{1}{\sqrt{8}}\left(|\varphi_1\rangle + |\varphi_3\rangle\right)\left(\langle\varphi_1| + \langle\varphi_3|\right) ,\end{aligned}\quad (3.11)$$

but the fourth element  $\hat{M}_4$  is yet to be determined.

- (a) Calculate the POVM associated with the measurement.
- (b) Hence find a possible solution for  $\hat{M}_4$ .
- (c) Write down the format for the general solution for  $\hat{M}_4$ .

# Advanced Quantum Theory

## 3 Measurement Theory

### Solutions to Problems

- (1)** Noting that  $\{|\varphi_n\rangle : n = 1, 2, \dots\}$  is an orthonormal basis we find

$$\begin{aligned}
 (\hat{P}_1)^2 &= \left( |\varphi_1\rangle\langle\varphi_1| \right) \left( |\varphi_1\rangle\langle\varphi_1| \right) = |\varphi_1\rangle\langle\varphi_1|\varphi_1\rangle\langle\varphi_1| = |\varphi_1\rangle\langle\varphi_1| \\
 &= \hat{P}_1 \quad \checkmark \\
 (\hat{P}_1)^\dagger &= \left( |\varphi_1\rangle\langle\varphi_1| \right)^\dagger = |\varphi_1\rangle\langle\varphi_1| \\
 &= \hat{P}_1 \quad \checkmark \\
 (\hat{P}_2)^2 &= \left( |\varphi_1\rangle\langle\varphi_1| + |\varphi_5\rangle\langle\varphi_5| \right) \left( |\varphi_1\rangle\langle\varphi_1| + |\varphi_5\rangle\langle\varphi_5| \right) \\
 &= |\varphi_1\rangle\langle\varphi_1|\varphi_1\rangle\langle\varphi_1| + |\varphi_1\rangle\langle\varphi_1|\varphi_5\rangle\langle\varphi_5| \\
 &\quad + |\varphi_5\rangle\langle\varphi_5|\varphi_1\rangle\langle\varphi_1| + |\varphi_5\rangle\langle\varphi_5|\varphi_5\rangle\langle\varphi_5| \\
 &= |\varphi_1\rangle\langle\varphi_1| + |\varphi_5\rangle\langle\varphi_5| \\
 &= \hat{P}_2 \quad \checkmark \\
 (\hat{P}_2)^\dagger &= \left( |\varphi_1\rangle\langle\varphi_1| + |\varphi_5\rangle\langle\varphi_5| \right)^\dagger = |\varphi_1\rangle\langle\varphi_1| + |\varphi_5\rangle\langle\varphi_5| \\
 &= \hat{P}_2 \quad \checkmark
 \end{aligned}$$

- (2)** The operator  $\hat{O}$  can be written in terms of projection operators as

$$\hat{O} = 1\hat{Q}_1 + 2\hat{Q}_2$$

where

$$\begin{aligned}
 \hat{Q}_1 &= |-1\rangle_{zz}\langle -1| \\
 \hat{Q}_2 &= |0\rangle_{zz}\langle 0| + |1\rangle_{zz}\langle 1|
 \end{aligned}$$

**CHECK:** These are projection operators because

$$\begin{aligned}
 (\hat{Q}_1)^2 &= \hat{Q}_1 \text{ by inspection and} \\
 (\hat{Q}_2)^2 &= |0\rangle_{zz}\langle 0| |0\rangle_{zz}\langle 0| + |0\rangle_{zz}\langle 0| |1\rangle_{zz}\langle 1| \\
 &\quad + |1\rangle_{zz}\langle 1| |0\rangle_{zz}\langle 0| + |1\rangle_{zz}\langle 1| |1\rangle_{zz}\langle 1| \\
 &= |0\rangle_{zz}\langle 0| + |1\rangle_{zz}\langle 1| = \hat{Q}_2
 \end{aligned}$$

#### Measurement description

Initial state of the system :  $|\psi\rangle = \frac{1}{2}|-1\rangle_z + \frac{1}{\sqrt{2}}|0\rangle_z + \frac{1}{2}|1\rangle_z$

Projection operators:  $\{\hat{Q}_1, \hat{Q}_2\}$

Outcome 1	probability:	$p_1 = \ \hat{Q}_1 \psi\rangle\ ^2 = (\frac{1}{2})^2 = \frac{1}{4}$
	final state:	$\hat{Q}_1 \psi\rangle/\sqrt{p_1} =  -1\rangle_z$

Outcome 2	probability:	$p_2 = \ \hat{Q}_2 \psi\rangle\ ^2 = \frac{3}{4}$
	final state:	$\hat{Q}_1 \psi\rangle/\sqrt{p_1} = \frac{2}{\sqrt{6}} 0\rangle_z + \frac{1}{\sqrt{3}} 1\rangle_z$
Averaged final state:		$\hat{Q}_1 \psi\rangle\langle\psi \hat{Q}_1 + \hat{Q}_2 \psi\rangle\langle\psi \hat{Q}_2$
		$= \frac{1}{4}  -1 \rangle_{zz}\langle -1   + \left( \frac{1}{\sqrt{2}} 0\rangle_z + \frac{1}{2} 1\rangle_z \right) \left( \frac{1}{\sqrt{2}}z\langle 0   + \frac{1}{2}z\langle 1   \right).$

(3) Using the cyclic property of the trace we find

$$\begin{aligned} \text{Tr}\left(\hat{U}\rho\hat{U}^\dagger\right) &= \text{Tr}\left(\hat{U}^\dagger\hat{U}\rho\right) \\ &= \text{Tr}\left(\hat{\mathbb{1}}\rho\right) \\ &= \text{Tr}(\hat{\rho}) \end{aligned}$$

and so the action of  $\hat{U}$  leaves the trace unchanged.

(4) First evaluate each POVM element

$$\begin{aligned} \hat{E}_1 &= \hat{M}_1^\dagger \hat{M}_1 = T^2 |\mathsf{H}\rangle\langle\mathsf{H}| , \\ \hat{E}_2 &= \hat{M}_2^\dagger \hat{M}_2 \\ &= \frac{1}{2(R^2+1)} \left[ \left( R|\mathsf{H}\rangle + |\mathsf{V}\rangle \right) \left( R\langle\mathsf{H}| + \langle\mathsf{V}| \right) \right]^\dagger \\ &\quad \times \left[ \left( R|\mathsf{H}\rangle + |\mathsf{V}\rangle \right) \left( R\langle\mathsf{H}| + \langle\mathsf{V}| \right) \right] \\ &= \frac{1}{2(R^2+1)} \left[ \left( R|\mathsf{H}\rangle + |\mathsf{V}\rangle \right) \left( R^2 + 1 \right) \left( R\langle\mathsf{H}| + \langle\mathsf{V}| \right) \right] \\ &= \frac{1}{2} \left( R|\mathsf{H}\rangle + |\mathsf{V}\rangle \right) \left( R\langle\mathsf{H}| + \langle\mathsf{V}| \right) , \\ \\ \hat{E}_3 &= \hat{M}_3^\dagger \hat{M}_3 \\ &= \frac{1}{2(R^2+1)} \left[ \left( R|\mathsf{H}\rangle - |\mathsf{V}\rangle \right) \left( R\langle\mathsf{H}| - \langle\mathsf{V}| \right) \right]^\dagger \\ &\quad \times \left[ \left( R|\mathsf{H}\rangle - |\mathsf{V}\rangle \right) \left( R\langle\mathsf{H}| - \langle\mathsf{V}| \right) \right] \\ &= \frac{1}{2(R^2+1)} \left[ \left( R|\mathsf{H}\rangle - |\mathsf{V}\rangle \right) \left( R^2 + 1 \right) \left( R\langle\mathsf{H}| - \langle\mathsf{V}| \right) \right] \\ &= \frac{1}{2} \left( R|\mathsf{H}\rangle - |\mathsf{V}\rangle \right) \left( R\langle\mathsf{H}| - \langle\mathsf{V}| \right) \end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^3 \hat{E}_i &= T^2 |\mathsf{H}\rangle\langle\mathsf{H}| + \frac{1}{2} \left( R|\mathsf{H}\rangle + |\mathsf{V}\rangle \right) \left( R\langle\mathsf{H}| + \langle\mathsf{V}| \right) \\
&\quad + \frac{1}{2} \left( R|\mathsf{H}\rangle - |\mathsf{V}\rangle \right) \left( R\langle\mathsf{H}| - \langle\mathsf{V}| \right) \\
&= \left( T^2 + R^2 \right) |\mathsf{H}\rangle\langle\mathsf{H}| + \left( \frac{1}{2} + \frac{1}{2} \right) |\mathsf{V}\rangle\langle\mathsf{V}| \\
&= |\mathsf{H}\rangle\langle\mathsf{H}| + |\mathsf{V}\rangle\langle\mathsf{V}| \\
&= \hat{\mathbb{1}}
\end{aligned}$$

where we have used  $T^2 + R^2 = 1$ .

(5) (a) First evaluate each **known** POVM element

$$\begin{aligned}
\hat{E}_1 &= \hat{M}_1^\dagger \hat{M}_1 \\
&= \left( \frac{1}{2} |\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| \right)^\dagger \left( \frac{1}{2} |\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| \right) \\
&= \frac{1}{4} |\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| \\
\hat{E}_2 &= \hat{M}_2^\dagger \hat{M}_2 \\
&= \left( \frac{1}{2} |\varphi_1\rangle\langle\varphi_1| + \frac{1}{\sqrt{2}} |\varphi_3\rangle\langle\varphi_3| \right)^\dagger \left( \frac{1}{2} |\varphi_1\rangle\langle\varphi_1| + \frac{1}{\sqrt{2}} |\varphi_3\rangle\langle\varphi_3| \right) \\
&= \frac{1}{4} |\varphi_1\rangle\langle\varphi_1| + \frac{1}{2} |\varphi_3\rangle\langle\varphi_3| \\
\hat{E}_3 &= \hat{M}_3^\dagger \hat{M}_3 \\
&= \left[ \frac{1}{\sqrt{8}} \left( |\varphi_1\rangle + |\varphi_3\rangle \right) \left( \langle\varphi_1| + \langle\varphi_3| \right) \right]^\dagger \\
&\quad \times \left[ \frac{1}{\sqrt{8}} \left( |\varphi_1\rangle + |\varphi_3\rangle \right) \left( \langle\varphi_1| + \langle\varphi_3| \right) \right] \\
&= \frac{1}{8} \left( |\varphi_1\rangle + |\varphi_3\rangle \right) \left( 2 \right) \left( \langle\varphi_1| + \langle\varphi_3| \right) \\
&= \frac{1}{4} \left( |\varphi_1\rangle + |\varphi_3\rangle \right) \left( \langle\varphi_1| + \langle\varphi_3| \right)
\end{aligned}$$

Next evaluate the sum of these three POVM elements

$$\begin{aligned}
\sum_{i=1}^3 \hat{E}_i &= \frac{1}{4} |\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| + \frac{1}{4} |\varphi_1\rangle\langle\varphi_1| + \frac{1}{2} |\varphi_3\rangle\langle\varphi_3| \\
&\quad + \frac{1}{4} \left( |\varphi_1\rangle + |\varphi_3\rangle \right) \left( \langle\varphi_1| + \langle\varphi_3| \right) \\
&= \frac{3}{4} |\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| + \frac{3}{4} |\varphi_3\rangle\langle\varphi_3| \\
&\quad + \frac{1}{4} |\varphi_1\rangle\langle\varphi_3| + \frac{1}{4} |\varphi_3\rangle\langle\varphi_1| .
\end{aligned}$$

This is clearly not equal to  $\hat{\mathbb{1}} = |\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| + |\varphi_3\rangle\langle\varphi_3|$  and so the POVM must contain at least one more element. Let this element be  $\hat{E}_4$  such that  $\sum_{i=1}^4 \hat{E}_i = \hat{\mathbb{1}}$ . In that case we have

$$\begin{aligned}\hat{\mathbb{1}} &= \left( \sum_{i=1}^3 \hat{E}_i \right) + \hat{E}_4, \quad \text{and so} \\ \hat{E}_4 &= \hat{\mathbb{1}} - \left( \sum_{i=1}^3 \hat{E}_i \right) \\ &= \hat{\mathbb{1}} - \left( \frac{3}{4}|\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| + \frac{3}{4}|\varphi_3\rangle\langle\varphi_3| \right. \\ &\quad \left. + \frac{1}{4}|\varphi_1\rangle\langle\varphi_3| + \frac{1}{4}|\varphi_3\rangle\langle\varphi_1| \right) \\ &= \frac{1}{4}|\varphi_1\rangle\langle\varphi_1| + \frac{1}{4}|\varphi_3\rangle\langle\varphi_3| - \frac{1}{4}|\varphi_1\rangle\langle\varphi_3| - \frac{1}{4}|\varphi_3\rangle\langle\varphi_1| \\ &= \frac{1}{4}(|\varphi_1\rangle - |\varphi_3\rangle)(\langle\varphi_1| - \langle\varphi_3|)\end{aligned}$$

The POVM associated with the measurement is  $\{\hat{E}_1, \hat{E}_2, \hat{E}_3, \hat{E}_4\}$  where the elements  $\hat{E}_i$  are given above.

**(b)** We need to solve

$$\hat{M}_4^\dagger \hat{M}_4 = \hat{E}_4 = \frac{1}{4}(|\varphi_1\rangle - |\varphi_3\rangle)(\langle\varphi_1| - \langle\varphi_3|) \quad (3.12)$$

for  $\hat{M}_4$ . From knowledge of previous calculations a reasonable guess would be

$$\hat{M}_4 \stackrel{?}{=} C(|\varphi_1\rangle - |\varphi_3\rangle)(\langle\varphi_1| - \langle\varphi_3|)$$

where  $C$  is a constant because then

$$\begin{aligned}\hat{M}_4^\dagger \hat{M}_4 &\stackrel{?}{=} |C|^2 \left[ (|\varphi_1\rangle - |\varphi_3\rangle)(\langle\varphi_1| - \langle\varphi_3|) \right]^\dagger \\ &\quad \left[ (|\varphi_1\rangle - |\varphi_3\rangle)(\langle\varphi_1| - \langle\varphi_3|) \right] \\ &\stackrel{?}{=} 2|C|^2 (|\varphi_1\rangle - |\varphi_3\rangle)(\langle\varphi_1| - \langle\varphi_3|).\end{aligned}$$

Comparing with Eq. (3.12) shows that  $2|C|^2 = \frac{1}{4}$  and so  $C = \frac{1}{\sqrt{8}}$ . Hence a possible solution is

$$\hat{M}_4 = \frac{1}{\sqrt{8}} (|\varphi_1\rangle - |\varphi_3\rangle)(\langle\varphi_1| - \langle\varphi_3|).$$

**(c)** Performing an additional unitary operator gives the most

general form, i.e.

$$\hat{M}_4 = \hat{U} \frac{1}{\sqrt{8}} \left( |\varphi_1\rangle - |\varphi_3\rangle \right) \left( \langle \varphi_1| - \langle \varphi_3| \right).$$

# 4. Quantum Entanglement

## a. Composite systems

We now consider the quantum description of composite systems. These are systems which are composed of a number of smaller individual subsystems. Examples are a system composed of a number of silver atoms (each modelled as a spin- $\frac{1}{2}$  particle), a system composed of photons travelling in different directions, or even a system composed of a photon and a silver atom etc. Typically we consider composites of just 2 subsystems, but the approach can be applied to composite systems composed of any number of subsystems.

### Terminology: systems and subsystems...

We use the terms system and subsystem interchangeably. The **larger** system will usually be called the total, whole or combined system and its **components** will be called the subsystems of just systems.

### Tensor products of states

Consider the possible outcomes for the toss of two coins:

coin 1	coin 2
heads	heads
heads	tails
tails	heads
tails	tails

Notice that we give the state of the outcome as

$$(\text{state of coin 1}) \times (\text{state of coin 2}).$$

Similarly for the toss of two dice, the state of the outcome is

$$(\text{state of die 1}) \times (\text{state of die 2})$$

and so possible states of two dice are

$$\begin{array}{ll} 1 & \text{and } 4, \text{ or} \\ 6 & \text{and } 2, \text{ etc.} \end{array}$$

In quantum physics we similarly represent the state of a **composite system** comprising two systems as

$$(\text{state of system 1}) \otimes (\text{state of system 2}).$$

The symbol  $\otimes$  indicates a **tensor product**. For example consider

two spin- $\frac{1}{2}$  particles. If the state of the first is  $|\frac{1}{2}\rangle_z$  and the second is  $|-\frac{1}{2}\rangle_y$  then the state of the combined systems is

insert  
1

$$|\psi\rangle = |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_y$$

The states of the combined systems lies in the tensor product of the Hilbert spaces of the individual systems, viz.  $H_{\text{com}} = H_1 \otimes H_2$  where  $H_{\text{com}}$ ,  $H_1$  and  $H_2$  are the Hilbert spaces of the combined systems, system 1 and system 2, respectively. There are many bases for  $H_{\text{com}}$ . The simplest is just a tensor product of the bases of the individual systems such as  $\{|\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z, |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z, |-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z, |-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z\}$ , or writing out in table form, the set of states:

$$\begin{aligned} & |\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z , \\ & |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z , \\ & |-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z , \\ & |-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z . \end{aligned}$$

### Symbol $\otimes$

You should think of the symbol  $\otimes$  as combining the states of two systems in the same way as the outcomes of two dice are combined.

## Tensor products of operators

Let the operators  $\hat{A}_1$  and  $\hat{A}_2$  represent operations on system 1 and 2, respectively. That is,  $\hat{A}_1$  operates on states in the Hilbert space  $H_1$  and  $\hat{A}_2$  on states in  $H_2$ . Thus if initially system 1 is in the state  $|\psi_1\rangle$  and system 2 is in the state  $|\psi_2\rangle$ , the action of these operators is to produce the final states  $\hat{A}_1|\psi_1\rangle$  and  $\hat{A}_2|\psi_2\rangle$ . Expressed as a tensor product, this means the initial combined state

$$|\psi_1\rangle \otimes |\psi_2\rangle$$

becomes the final combined state

$$\left(\hat{A}_1|\psi_1\rangle\right) \otimes \left(\hat{A}_2|\psi_2\rangle\right).$$

We write the combined operator as a tensor product of operators  $\hat{A}_1 \otimes \hat{A}_2$ . Thus

$$\left(\hat{A}_1 \otimes \hat{A}_2\right)\left(|\psi_1\rangle \otimes |\psi_2\rangle\right) = \left(\hat{A}_1|\psi_1\rangle\right) \otimes \left(\hat{A}_2|\psi_2\rangle\right).$$

If operation  $\hat{A}_1$  only performed on system 1 and system 2 is left unchanged, the combined operations is  $\hat{A}_1 \otimes \mathbb{1}_2$  where  $\mathbb{1}_2$  is the

identity operator for system 2 (and Hilbert space  $H_2$ ). This gives

$$\begin{aligned} (\hat{A}_1 \otimes \mathbf{1}_2) (|\psi_1\rangle \otimes |\psi_2\rangle) &= (\hat{A}_1|\psi_1\rangle) \otimes (\mathbf{1}_2|\psi_2\rangle) \\ &= (\hat{A}_1|\psi_1\rangle) \otimes |\psi_2\rangle . \end{aligned}$$

The matrix elements of a tensor product of operators is just the product of the matrix elements of the individual operators. For example, let the sets

insert  
2

$$\{|\varphi_{1,i}\rangle : i = 1, 2, \dots\} \text{ and } \{|\varphi_{2,i}\rangle : i = 1, 2, \dots\}$$

be orthonormal bases of system 1 and system 2, respectively. The basis of the combined system is

$$\{|\varphi_{1,i}\rangle \otimes |\varphi_{2,j}\rangle : i = 1, 2, \dots, j = 1, 2, \dots\} .$$

The matrix elements of the combined operator  $\hat{A} = \hat{A}_1 \otimes \hat{A}_2$  is given by

$$\begin{aligned} A_{i,j,m,n} &= (\langle \varphi_{1,i} | \otimes \langle \varphi_{2,j} |) \hat{A} (|\varphi_{1,m}\rangle \otimes |\varphi_{2,n}\rangle) \\ &= (\langle \varphi_{1,i} | \otimes \langle \varphi_{2,j} |) (\hat{A}_1 \otimes \hat{A}_2) (|\varphi_{1,m}\rangle \otimes |\varphi_{2,n}\rangle) \\ &= (\langle \varphi_{1,i} | \otimes \langle \varphi_{2,j} |) (\hat{A}_1|\varphi_{1,m}\rangle) \otimes (\hat{A}_2|\varphi_{2,n}\rangle) \\ &= (\langle \varphi_{1,i} | \hat{A}_1 | \varphi_{1,m} \rangle) \otimes (\langle \varphi_{1,j} | \hat{A}_2 | \varphi_{2,n} \rangle) \\ &= (\hat{A}_1)_{i,m} (\hat{A}_2)_{j,n} \end{aligned}$$

where

$$\begin{aligned} (\hat{A}_1)_{i,m} &= \langle \varphi_{1,i} | \hat{A}_1 | \varphi_{1,m} \rangle \\ (\hat{A}_2)_{j,n} &= \langle \varphi_{2,j} | \hat{A}_2 | \varphi_{2,n} \rangle . \end{aligned}$$

The tensor product operation is also used to describe the general mixed state of two systems. If the systems have no correlations then the density operator for the combined system is

$$\hat{\rho} = \hat{\rho}_1 \otimes \hat{\rho}_2$$

where  $\hat{\rho}_1$  and  $\hat{\rho}_2$  are the density operators of the respective systems.

### Implicit notation

People often omit writing the tensor product explicitly and, for instance, write the state  $|\psi_1\rangle \otimes |\phi_2\rangle$  as  $|\psi_1\rangle|\phi_2\rangle$  with the tensor product operation being implicit.

## The partial trace and the reduced density operator

If we are interested in expectation values for one system only then we can simplify the problem using the **partial trace**. For example, imagine the state of the combined system is given by  $\hat{\rho}$  and we want to calculate the expectation value of  $\hat{A}_1$  for the first system only. Using the trace form of calculating expectation values gives

$$\begin{aligned}
 \langle \hat{A}_1 \rangle &= \langle \hat{A}_1 \otimes \mathbb{1}_2 \rangle \\
 &= \text{Tr}[\hat{\rho}(\hat{A}_1 \otimes \mathbb{1}_2)] \\
 &= \sum_i \sum_j (\langle \varphi_{1,i} | \otimes \langle \varphi_{2,j} |) [\hat{\rho}(\hat{A}_1 \otimes \mathbb{1}_2)] (|\varphi_{1,i}\rangle \otimes |\varphi_{2,j}\rangle) \\
 &= \sum_i \sum_j (\langle \varphi_{1,i} | \otimes \langle \varphi_{2,j} |) \hat{\rho}(\hat{A}_1 |\varphi_{1,i}\rangle \otimes |\varphi_{2,j}\rangle) \\
 &= \sum_i \langle \varphi_{1,i} | \left( \sum_j \langle \varphi_{2,j} | \hat{\rho} | \varphi_{2,j} \rangle \right) \hat{A}_1 |\varphi_{1,i}\rangle \\
 &= \sum_i \langle \varphi_{1,i} | \hat{\rho}_1 \hat{A}_1 | \varphi_{1,i} \rangle
 \end{aligned}$$

where  $\hat{\rho}_1$  is defined by the following.

### Reduced density operator for system 1

$$\begin{aligned}
 \hat{\rho}_1 &= \sum_j \langle \varphi_{2,j} | \hat{\rho} | \varphi_{2,j} \rangle \\
 &= \text{Tr}_2(\hat{\rho}) \quad (\text{partial trace over system 2})
 \end{aligned}$$

Here  $\hat{\rho}_1$  is an operator acting on Hilbert space  $H_1$  and represents the state of system 1 for calculations involving system 1 only. It is called the **reduced density operator for system 1**. Notice that the trace  $\text{Tr}_2$  is over system 2 basis states only; this operation is called the **partial trace over system 2**.

### Partial trace of single terms

For kets  $|x\rangle_2$ ,  $|y\rangle_2$  belonging to an orthonormal basis such as  $\{|\phi_n\rangle_2 : n = 1, 2, \dots\}$  the following is true

$$\text{Tr}_2 \left[ \left( |A\rangle_1 \otimes |x\rangle_2 \right) \left( {}_1\langle B| \otimes {}_2\langle y| \right) \right] = \begin{cases} |A\rangle_1 \langle B|, & \text{if } x = y \\ 0, & \text{if } x \neq y \end{cases}. \quad (4.1)$$

**Problem**

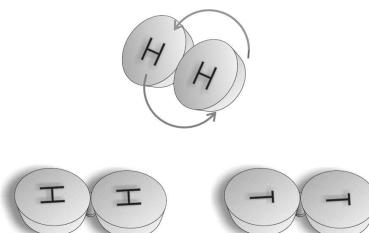
Prove Eq. (4.1).

## b. Correlations

So far we have considered two systems which are completely separate and uncorrelated. By uncorrelated we mean that if a measurement is made on one system it tells us nothing about the other system. We now want to consider correlated systems. There are two forms of correlation, **classical correlations** and **quantum correlations**. We need to distinguish between them because it is the quantum correlations that are the most interesting.

### Mixed states and “classical” uncertainty

Classical correlations arise when there is some uncertainty about the state of a composite system. For example after the toss of 2 coins we do not know the state of the coins (the outcome). But imagine that the coins are glued side by side as shown in the figure. If one lands heads up the other is heads up. Hence knowing the state of one system after the toss immediately tells us the state of the other coin.



**Fig 4.1:** Correlated coins. The two coins are glued together so that they always show the same face.

This is rather a contrived example. A better one is to consider a box holding a red and a green ball. Person 1 picks out a ball at random leaving the remaining ball for person 2. If person 1 reveals the colour of his/her ball, we immediately know the colour of person 2's ball. In this case the balls held by the two people are **anti-correlated** as the colour of person 1's ball (e.g. red) implies that person 2's ball has the **opposite** colour (green). There is a 50% chance that person 1 will have the red ball and person 2 the green ball, and 50% chance that person 1 will have the green ball and person 2 the red ball.

This situation can be found in quantum systems as well. Consider two spin- $\frac{1}{2}$  particles. The analogy of the anti-correlated red-green balls is a 50% of the spins being in the tensor product state  $|\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z$  and 50% chance of the spins being in  $|-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z$ .

The combined system density operator is therefore

$$\hat{\rho} = \frac{1}{2} \left( \left| \frac{1}{2} \right\rangle_z \otimes \left| -\frac{1}{2} \right\rangle_z \right) \left( z \langle \frac{1}{2} | \otimes z \langle -\frac{1}{2} | \right) + \frac{1}{2} \left( \left| -\frac{1}{2} \right\rangle_z \otimes \left| \frac{1}{2} \right\rangle_z \right) \left( z \langle -\frac{1}{2} | \otimes z \langle \frac{1}{2} | \right).$$

This kind of correlation is not so interesting. It is the familiar classical correlation.

### Pure states of composite systems - quantum correlations (entanglement)

In quantum physics, the combined state can be **pure** and yet there can be **correlations** between the subsystems. Pure states represent perfect knowledge of the state of the system (but not necessarily its observables). This means, the state of the combined system is known and certain, and yet the states of the subsystems can be uncertain and correlated.

#### Nonclassical feature

This situation is not possible in classical physics: any uncertainty in the subsystems implies uncertainty in the state of the combined system. For example, for the red-green balls example, if the colour of person 1's ball is uncertain (could be red or green) then the combined system of person 1's ball and person 2's ball is uncertain.

An example of a quantum correlation is given by the **pure** state version of the anti-correlated spin- $\frac{1}{2}$  particles:

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left( \left| \frac{1}{2} \right\rangle_z \otimes \left| -\frac{1}{2} \right\rangle_z \right) + \frac{1}{\sqrt{2}} \left( \left| -\frac{1}{2} \right\rangle_z \otimes \left| \frac{1}{2} \right\rangle_z \right).$$

The state of the combined system is **certain**. Appropriate measurements can confirm this. However, the state of the first particle is uncertain as is the state of the second particle. If a measurement on the first particle finds it in the state  $|\frac{1}{2}\rangle_z$ , after the measurement the state of the combined system would be

insert  
3

$$|\psi_{\text{after}}\rangle = \left( \left| \frac{1}{2} \right\rangle_z \otimes \left| -\frac{1}{2} \right\rangle_z \right)$$

and so a measurement on the second particle will find it in the (opposite) state  $|\frac{1}{2}\rangle_z$ , and vice versa. Clearly there is a strong **anti-correlation** between the particles.

Systems exhibiting quantum correlations are called **entangled** and the amount of quantum correlation is called the **quantum entanglement** or just the **entanglement**.

### **Classical and quantum correlations**

There can be both kinds of correlations present at the same time, both quantum and classical. In this case the state of the combined system would be in a mixed state. However the analysis is difficult in this general case. In this course we will only consider the quantum version, which corresponds to pure states of the combined system.

## **c. Entanglement**

**Entanglement** is a central, defining feature of quantum physics. In its simplest form it arises as a superposition of the states of separated systems. Yet the consequences of this are quite profound. It has shaken our understanding of the physical world, and led some to question the very reality of observables and even the completeness of quantum theory. Whenever entanglement is present in a quantum system, the system possesses correlations beyond the scope of classical physics. That is, it has correlations that no classical description can reproduce. More about these aspects will be discussed in the section **Bell's Theorem**. The ramifications of this simple superposition state of composite systems are still being explored. It is the subject of a large and ongoing research programme in both theoretical and experimental physics. It remains an enigma from the theoretical perspective, and an experimental challenge to produce and exploit in new quantum technologies such as quantum cryptographic schemes, quantum computers and quantum teleportation...

The term entanglement was first used by Schrödinger in around 1936. His way of thinking about entanglement was like this. Entanglement is a property possessed, jointly, by two quantum systems. When it is present in a pair of quantum systems, this means that each particle is so intimately intertwined, linked or **entangled** with the other one that **it does not exist in its own well-defined state**. Only the combined state of both particles together is well-defined. This idea is somewhat at odds with our everyday experience. Surely, you might say, if the state of the both particles as a whole is well-defined then the states of each of the two particles making up this whole must necessarily be well-defined. But nature suggests that this is not true. One natural way to understand entanglement is to think of it as a **holistic** property. That is, when two particles are entangled with each other, we should think of them a one single whole instead of thinking of them in **reductionistic** manner as two separate entities.

### Definition of entangled states

A quantum system comprising two subsystems is said to be **entangled** when the system's total (pure) state  $|\Psi\rangle$  cannot be written as a tensor product of the states of the subsystems, that is, when

$$|\Psi\rangle \neq |\psi\rangle \otimes |\phi\rangle .$$

The two subsystems are said to be **entangled with each other**.

### Definition of separable states

Two subsystems which are not entangled are said to be in a **separable state** which has the form of a tensor product, i.e.,

$$|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle .$$

- **Example (1).**

Consider two spin- $\frac{1}{2}$  particles in the combined state  $|\Psi\rangle = |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z$ . This state is of the tensor product form and hence it is **not** entangled. Clearly it is separable.

• • •

- **Example (2).**

Consider the combined state that we used before

insert  
4

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left( |-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \right). \quad (4.2)$$

We prove that  $|\Psi\rangle$  is an entangled state by *reductio ad absurdum*. We begin by assuming that we **can** write  $|\Psi\rangle$  as a separable state. In that case we can express  $|\Psi\rangle$  as

$$|\Psi\rangle = \left( a|\frac{1}{2}\rangle_z + b|-\frac{1}{2}\rangle_z \right) \otimes \left( c|\frac{1}{2}\rangle_z + d|-\frac{1}{2}\rangle_z \right), \quad (4.3)$$

where  $a, b, c$  and  $d$  are complex numbers satisfying  $|a|^2 + |b|^2 = |c|^2 + |d|^2 = 1$ . Expanding the right-hand side of Eq. (4.3) gives

$$\begin{aligned} |\Psi\rangle &= ac|\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + ad|\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \\ &\quad + bc|-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + bd|-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z . \end{aligned}$$

Equating coefficients in the expressions for  $|\Psi\rangle$  in Eq. (4.2) and Eq. (4.4) yields the conditions  $ac = 0$ ,  $bd = 0$ ,  $bc = \frac{1}{\sqrt{2}}$  and  $ad = \frac{1}{\sqrt{2}}$ .

As  $ac = 0$ , then either (i)  $a = 0$  or (ii)  $c = 0$  or (iii) both are zero. However,  $a$  cannot be zero because  $ad = \frac{1}{\sqrt{2}}$  and  $c$  cannot be zero because  $bc = \frac{1}{\sqrt{2}}$ . The conditions on the coefficients are inconsistent. Thus Eq. (4.3) cannot be a true representation of

$|\Psi\rangle$  and so our assumption is false. We have thus proved that  $|\Psi\rangle$  is entangled.

• • •

• **Example (3).**

Any state of the forms

$$\begin{aligned} |\Psi_1\rangle &= \alpha|\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z + \beta|-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z \\ |\Psi_2\rangle &= \alpha|\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + \beta|-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \end{aligned}$$

is an entangled state for  $|\alpha| \neq 0$  and  $|\beta| \neq 0$ .

• • •

• **Example (4).**

Is the following state entangled?

$$\begin{aligned} |\Psi\rangle &= \frac{1}{2}|\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + \frac{1}{2}|\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \\ &\quad + \frac{1}{2}|-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + \frac{1}{2}|-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z . \end{aligned}$$

At first it might seem that the answer is “yes” but in fact the state can be factorised as

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z + |-\frac{1}{2}\rangle_z \right) \otimes \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z + |-\frac{1}{2}\rangle_z \right)$$

insert  
5

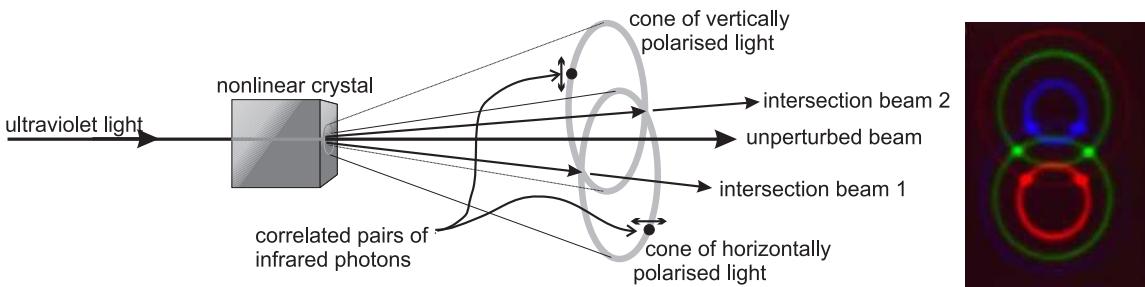
$$= |\frac{1}{2}\rangle_x \otimes |\frac{1}{2}\rangle_x$$

and so  $|\Psi\rangle$  is separable. (Indeed if we used the method of the proof in example 2 we would find the coefficients  $a = b = c = d = \frac{1}{\sqrt{2}}$  which is a consistent solution.)

• • •

### A physical example of an entangled state

Consider a strong laser beam of ultraviolet light shone on a crystal of calcite, a birefringent material. Most of the photons will go straight through the crystal and come out the other side with the same frequency.



**Fig 4.2:** Down conversion light cones: the two beams at the intersections are entangled in polarisation. The image on right from Paul Kwiat's web site <http://research.physics.uiuc.edu/QI/Photonics/index.html>

However, about 1 in  $10^{10}$  of them are absorbed by the crystal and **down converted** into two lower-energy photons. Given a suitable arrangement of the crystal and the incident beam, the down converted photons in the two beams at the **intersections of the cones** shown in Fig. 4.2 are in the entangled state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|H\rangle \otimes |V\rangle + |V\rangle \otimes |H\rangle), \quad (4.4)$$

where  $|H\rangle$  and  $|V\rangle$  denote, respectively, horizontally and vertically polarised photons, and  $|\text{beam } 1\rangle \otimes |\text{beam } 2\rangle$  gives the corresponding state of the two beams. One intuitive way of thinking about this state is to say that it is one in which the state of both photons, together, is well-defined in the sense that **both have mutually orthogonal polarisations**, however, the state of each individual photon is not well-defined as it is neither definitely horizontal or vertical.

### The linear-entropy measure of entanglement

The reduced density operator (for a **pure** combined-system state) is different for entangled and separable states. For example, consider the separable combined-system state

insert  
6

$$|\Psi_s\rangle = |\phi_1\rangle \otimes |\phi_2\rangle .$$

The reduced density operator of the first system is given by the partial trace over the second system, i.e.

$$\begin{aligned} \hat{\rho}_1 &= \text{Tr}_2(|\Psi_s\rangle\langle\Psi_s|) \\ &= \text{Tr}_2\left[\left(|\phi_1\rangle \otimes |\phi_2\rangle\right)\left(\langle\phi_1| \otimes \langle\phi_2|\right)\right] \\ &= \left(|\phi_1\rangle\langle\phi_1|\right)\text{Tr}_2\left(|\phi_2\rangle\langle\phi_2|\right) \\ &= |\phi_1\rangle\langle\phi_1| \end{aligned}$$

which is a **pure state**. In contrast, for the entangled state

$$|\Psi_e\rangle = \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + |-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \right). \quad (4.5)$$

the reduced density operator for the first system is given by

$$\begin{aligned} & \hat{\rho}'_1 \\ &= \text{Tr}_2(|\Psi_e\rangle\langle\Psi_e|) \\ &= \frac{1}{2} \text{Tr}_2 \left[ \left( |\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + |-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \right) \left( z\langle\frac{1}{2}| \otimes z\langle\frac{1}{2}| + z\langle-\frac{1}{2}| \otimes z\langle-\frac{1}{2}| \right) \right] \\ &= \frac{1}{2} \text{Tr}_2 \left[ \left( |\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z \right) \left( z\langle\frac{1}{2}| \otimes z\langle\frac{1}{2}| \right) + \left( |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \right) \left( z\langle-\frac{1}{2}| \otimes z\langle-\frac{1}{2}| \right) \right. \\ &\quad \left. + \left( |-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \right) \left( z\langle\frac{1}{2}| \otimes z\langle\frac{1}{2}| \right) + \left( |-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \right) \left( z\langle-\frac{1}{2}| \otimes z\langle-\frac{1}{2}| \right) \right] \\ &= \frac{1}{2} \left[ \left( |\frac{1}{2}\rangle_{zz} \langle\frac{1}{2}| \right) \text{Tr}_2 \left( |\frac{1}{2}\rangle_{zz} \langle\frac{1}{2}| \right) + \left( |\frac{1}{2}\rangle_{zz} \langle-\frac{1}{2}| \right) \text{Tr}_2 \left( |\frac{1}{2}\rangle_{zz} \langle-\frac{1}{2}| \right) \right. \\ &\quad \left. + \left( |-\frac{1}{2}\rangle_{zz} \langle\frac{1}{2}| \right) \text{Tr}_2 \left( |-\frac{1}{2}\rangle_{zz} \langle\frac{1}{2}| \right) + \left( |-\frac{1}{2}\rangle_{zz} \langle-\frac{1}{2}| \right) \text{Tr}_2 \left( |-\frac{1}{2}\rangle_{zz} \langle-\frac{1}{2}| \right) \right] \\ &= \frac{1}{2} \left( |\frac{1}{2}\rangle_{zz} \langle\frac{1}{2}| + |-\frac{1}{2}\rangle_{zz} \langle-\frac{1}{2}| \right) \end{aligned}$$

which is a **mixed state**. This is a general result.

### Partial trace of entangled states

Pure entangled states have **mixed** reduced density operators.

We can use this fact in order to quantify the amount of entanglement. One convenient way is to base the measure of the entanglement on the purity,  $\text{Tr}(\hat{\rho}_{\text{reduced}}^2)$ , of the reduced density operator  $\hat{\rho}_{\text{reduced}}$ . We want the measure to be zero for separable states and 1 for the maximally entangled state in Eq. (4.5). The negative of the logarithm (base 2) of the purity of the reduced density operator has these properties. This measure is called the **linear entropy of entanglement**.

### Linear-entropy of entanglement

$$E_{\text{lin}}(|\Psi\rangle) = -\log_2[\text{Tr}(\hat{\rho}_{\text{reduced}}^2)]$$

where  $\hat{\rho}_{\text{reduced}}$  is the reduced density operator for one system, for example,

$$\hat{\rho}_{\text{reduced}} = \text{Tr}_2(|\Psi\rangle\langle\Psi|).$$

Consider the state  $|\Psi_1\rangle$  of Example (3) above. The linear entropy measure is insensitive to which system we trace over. Let's trace over the states of the second particle. This gives

$$\hat{\rho}_1 = \text{Tr}_2(|\Psi_1\rangle\langle\Psi_1|) = |\alpha|^2|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| + |\beta|^2|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| ,$$

and so the purity is

$$\text{Tr}_1(\hat{\rho}_1^2) = \text{Tr}\left(|\alpha|^4|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| + |\beta|^4|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}|\right) = |\alpha|^4 + |\beta|^4 ,$$

and thus the linear entropy of entanglement is

$$\begin{aligned} E_{\text{lin}}(|\Psi_1\rangle) &= -\log_2[\text{Tr}_1(\hat{\rho}_1^2)] \\ &= -\log_2(|\alpha|^4 + |\beta|^4) . \end{aligned}$$

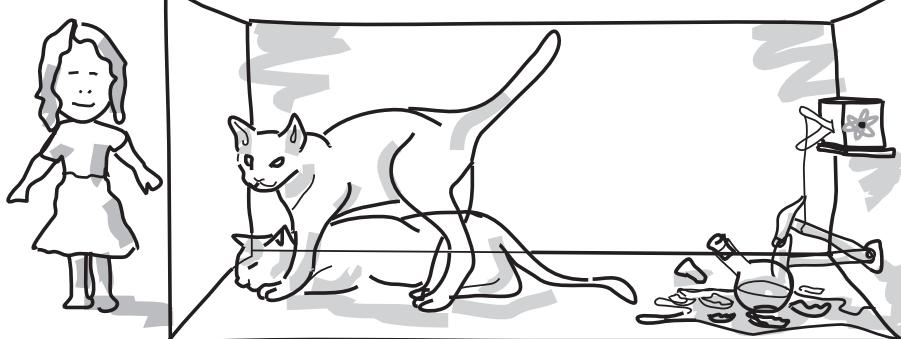
This has a minimum of zero when  $\alpha = 0$  or  $\beta = 0$  and a maximum of 1 for  $|\alpha| = |\beta| = \frac{1}{\sqrt{2}}$ .

## d. Schrödinger's cat and Many Worlds

Schrödinger was troubled by the implications of quantum theory. At one point he remarked

*I dont like it and I'm sorry I ever had anything to do with it.*

He introduced the following thought experiment to illustrate the absurd repercussions if quantum theory was applied to macroscopic systems. [See E. Schrödinger, *Naturwissenschaften* **23** 807-12, 812-28, 844-49 (1935). An English translation appears in *Proceedings of the American Philosophical Society* **124**, 323 (1980).]



**Fig 4.3:** Schrödinger's Cat

A **cat** is confined in a sealed box which also contains a vial of poison. The poison is released by an **atom** undergoing radioactive-decay event that has a 50-50 chance of occurring in an hour. What is the state of the cat at the end of the hour?

For this exercise we consider only the state of the cat and the atom and we assume they both begin in pure states (this is really stretching the reality!). The possible states of the cat are  $|\text{alive}\rangle_C$  and  $|\text{dead}\rangle_C$  and the states of the atom are  $|\text{undecayed}\rangle_A$  and  $|\text{decayed}\rangle_A$ , where the subscripts C and A stand for cat and atom,

respectively. The initial state of the combined cat+atom system is presumed to be

$$|\Psi(0)\rangle_{CA} = |\text{alive}\rangle_C \otimes |\text{undecayed}\rangle_A .$$

If the box is not disturbed, the state of the cat and the atom is after 1 hour is presumed to be

$$\begin{aligned} & |\Psi(1\text{hr})\rangle_{CA} \\ &= \frac{1}{\sqrt{2}} \left( |\text{alive}\rangle_C \otimes |\text{undecayed}\rangle_A + |\text{dead}\rangle_C \otimes |\text{decayed}\rangle_A \right) \end{aligned} \quad (4.6)$$

and the cat is in some quasi **alive-dead state**.

When someone opens the box and observes (i.e. measures) the state of the cat the wave function in Eq. (4.6) supposedly **collapses** to one member of the superposition. This is the conventional interpretation of quantum measurements which has become known the **Copenhagen Interpretation**. It was advocated by Niels Bohr and Werner Heisenberg who developed it through collaboration in Copenhagen sometime around 1927. It is by no means the only interpretation, or even the most widely accepted interpretation. Importantly, the collapse is not described by Schrödinger's equation.<sup>§</sup>

Nevertheless, if someone does make such an observation but doesn't tell you the outcome, the density operator you would assign to the state of the cat would be the equally-weighted mixture

$$\frac{1}{2} \left( |\text{alive}\rangle_{CC} \langle \text{alive}| + |\text{dead}\rangle_{CC} \langle \text{dead}| \right) \quad (4.7)$$

so it **appears** to you to be **either alive or dead** but you just don't know which. Moreover, if we are interested only in the state of the cat and not the atom, we should consider the **reduced density operator** for the cat only. This is given by taking the **partial trace** with respect to the states of the atom.

---

<sup>§</sup>The outcome of the collapse is random, whereas the Schrödinger equation describes deterministic evolution.

### Problem

The density operator for the combined cat-atom system is given by

$$\hat{\rho}_{CA}(1\text{hr}) = \left( |\Psi(1\text{hr})\rangle_{CA} \right) \left( {}_{CA}\langle \Psi(1\text{hr})| \right).$$

The **partial trace** with respect to the atom's states is given by

$$\begin{aligned} \hat{\rho}_C^{(\text{reduced})}(1\text{hr}) &= {}_A\langle \text{undecayed} | \hat{\rho}_{CA}(1\text{hr}) | \text{undecayed} \rangle_A \\ &\quad + {}_A\langle \text{decayed} | \hat{\rho}_{CA}(1\text{hr}) | \text{decayed} \rangle_A \end{aligned}$$

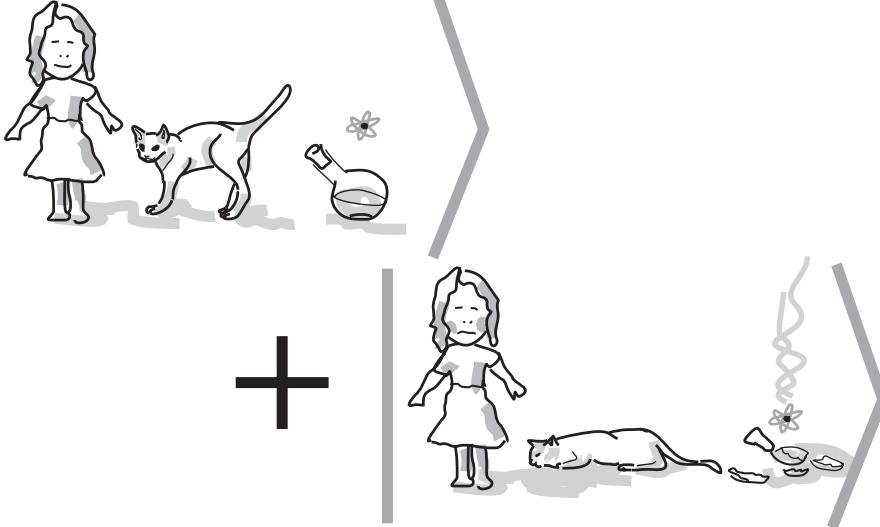
where  $\hat{\rho}_C^{(\text{reduced})}(1\text{hr})$  is the **reduced density operator** which describes the state of the cat only. Show that this expression equates to

$$\hat{\rho}_C^{(\text{reduced})}(1\text{hr}) = \frac{1}{2} \left( |\text{alive}\rangle_{CC} \langle \text{alive}| + |\text{dead}\rangle_{CC} \langle \text{dead}| \right).$$

Hence, comparing Eq. (4.7) with Eq. (4.10) shows that if we **consider the cat alone, it appears to be either dead or alive** and not in a quasi dead-alive state. The superposition of dead and alive states only occurs if we consider the state of the **combined atom-cat system**.

On the other hand, the evolution described by Schrödinger's equation is markedly different. The action of a **person** observing the state of the cat, like any physical process in the universe, should be described by this equation. However the practical calculation of the evolution of macroscopic and complex objects is not even remotely feasible. We need to make simplifying approximations that retain the kernel of the issue. Again we stretch the limits of rationality and assume that we can describe the person just by the **state of a memory element** in the person's brain which records the observation of the cat. Ignoring the rest of the person is like what we did in the Stern-Gerlach experiment where we considered only the spin state of the unpaired electron in a silver atom and neglected the rest of the atom. Further assume that the possible states of the memory element are  $|\text{no observation}\rangle_M$ ,  $|\text{see alive}\rangle_M$  and  $|\text{see dead}\rangle_M$  and that the memory element is in the "no observation" state before the observation, where here the subscript M labels the state of the memory element. The process of observation at the 1 hour point yields the correlation:

$$\begin{aligned} \frac{1}{\sqrt{2}} &\left( |\text{see alive}\rangle_M \otimes |\text{alive}\rangle_C \otimes |\text{undecayed}\rangle_A \right. \\ &\quad \left. + |\text{see dead}\rangle_M \otimes |\text{dead}\rangle_C \otimes |\text{decayed}\rangle_A \right). \quad (4.8) \end{aligned}$$



**Fig 4.4:** A cartoon of the state in Eq. (4.8).

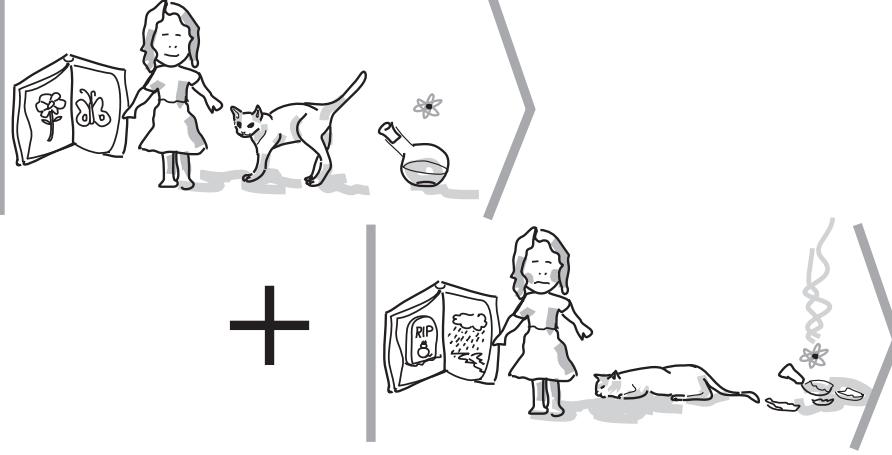
If we take the partial trace over the atom again we get the mixture

$$\frac{1}{2} \left[ \left( | \text{see alive} \rangle_M \otimes | \text{alive} \rangle_C \right) \left( M \langle \text{see alive} | \otimes C \langle \text{alive} | \right) + \left( | \text{see dead} \rangle_M \otimes | \text{dead} \rangle_C \right) \left( M \langle \text{see dead} | \otimes C \langle \text{dead} | \right) \right]$$

which represents a statistical mixture (not a superposition). Hence the memory element is statistically correlated with the state of the cat. This corresponds to what we would expect to see in “normal experience”.

However, the “super” composite system of memory-cat-atom is actually in the superposition state Eq. (4.8) and, in particular, we can’t assign individual states to any of the component systems, the memory, cat or atom. This **entanglement** grows as more objects interact with the person, cat or atom, for example if the person writes his/her experience in a diary (subscript D) the situation would be the perfect correlation

$$\frac{1}{\sqrt{2}} \left( | \text{I'm happy} \rangle_D \otimes | \text{see alive} \rangle_M \otimes | \text{alive} \rangle_C \otimes | \text{undecayed} \rangle_A + | \text{I'm sad} \rangle_D \otimes | \text{see dead} \rangle_M \otimes | \text{dead} \rangle_C \otimes | \text{decayed} \rangle_A \right).$$



**Fig 4.5:** A superposition of two “worlds” that evolve independently.

Moreover the linearity of Schrödinger’s equation ensures that **this superposition exists forever** in the following sense. The linearity means if  $|\psi_1(t)\rangle$  and  $|\psi_2(t)\rangle$  are two solutions to Schrödinger’s equation then so is the superposition  $\frac{1}{\sqrt{2}}[|\psi_1(t)\rangle + |\psi_2(t)\rangle]$ . Thus if at  $t = 0$

$$|\psi_1(0)\rangle = |I'm\ happy\rangle_D \otimes |see\ alive\rangle_M \otimes |alive\rangle_C \otimes |undecayed\rangle_A$$

$$|\psi_2(0)\rangle = |I'm\ sad\rangle_D \otimes |see\ dead\rangle_M \otimes |dead\rangle_C \otimes |decayed\rangle_A$$

and  $|\psi_1(t)\rangle$  and  $|\psi_2(t)\rangle$  represent the distant future evolution of each of these states under an arbitrary Hamiltonian, then the superposition state

$$\begin{aligned} & \frac{1}{\sqrt{2}} [|\psi_1(0)\rangle + |\psi_2(0)\rangle] \\ &= \frac{1}{\sqrt{2}} \left( |I'm\ happy\rangle_D \otimes |see\ alive\rangle_M \otimes |alive\rangle_C \otimes |undecayed\rangle_A \right. \\ & \quad \left. + |I'm\ sad\rangle_D \otimes |see\ dead\rangle_M \otimes |dead\rangle_C \otimes |decayed\rangle_A \right) . \end{aligned}$$

evolves to

$$\frac{1}{\sqrt{2}} \left[ \underbrace{|\psi_1(t)\rangle}_{\text{world I}} + \underbrace{|\psi_2(t)\rangle}_{\text{world II}} \right]$$

in the distant future under the same Hamiltonian. It’s as if  $|\psi_1(t)\rangle$  and  $|\psi_2(t)\rangle$  are **two worlds evolving independently**, one in which the cat died after one hour, and the other in which it didn’t. Indeed these two states are initially orthogonal and under the evolution of the Schrödinger equation, they remain orthogonal,

$$\langle \psi_1(t) | \psi_2(t) \rangle = 0 .$$

This is the basis of the **Many Worlds** interpretation of quantum mechanics introduced by Everett.<sup>¶</sup> In this interpretation measurements do not reduce superpositions to a single measurement outcome but rather generate entanglement. The system becomes ever more entangled such that its state splits into ever more “branches” like  $|\psi_1(t)\rangle$  and  $|\psi_2(t)\rangle$ . The different worlds arise from considering Eq. (4.8) to represent one world in which there is the memory of a dead cat along with the dead cat and decayed atom and another world in which there is the memory of a live cat along with the live cat and undecayed atom. There are presumably many such worlds to describe the universe. Our personal experience of just one world, corresponding to a single outcome of the observation, is an **illusion**. It is so strong that we **cannot escape** from it.

It is useful at this point to recount the words of Everett from a footnote in his paper:<sup>¶</sup>

“

In reply to a preprint of this article some correspondents have raised the question of the “transition from possible to actual,” arguing that in “reality” there is—as our experience testifies—no such splitting of observer states, so that only one branch can ever actually exist. Since this point may occur to other readers the following is offered in explanation.

The whole issue of the transition from “possible” to “actual” is taken care of in the theory in a very simple way—there is no such transition, nor is such a transition necessary for the theory to be in accord with our experience. From the viewpoint of the theory *all* elements of a superposition (all “branches”) are “actual,” none any more “real” than the rest. It is unnecessary to suppose that all but one are somehow destroyed, since all the separate elements of a superposition individually obey the wave equation with complete indifference to the presence of absence (“actuality” or not) of any other elements. This total lack of effect of one branch on another also implies that no observer will ever be aware of any “splitting” process.

Arguments that the world picture presented by this theory is contradicted by experience, because we are unaware of any branching process, are like the criticism of the Copernican theory that the mobility of the earth as a real physical fact is incompatible with common sense interpretation of nature because we feel no such motion. In both cases the argument fails when it is shown that the theory itself predicts that our experience will be what in fact it is. (In the Copernican case the addition of Newtonian physics was required to be able to show that the earth’s inhabitants would be unaware of any motion of the earth.)

, ,

---

<sup>¶</sup>H. Everett, “*Relative State*” Formulation of Quantum Mechanics, Rev. Mod. Phys. **29**, 454 (1957). See also J.A. Wheeler, Assessment of Everett’s “*Relative State*” Formulation of Quantum Theory, Rev. Mod. Phys. **29**, 463 (1957).

# Problem Sheet

(1) Prove Eq. (4.1).

(2) Prove that the state

$$|\Psi\rangle = \frac{1}{\sqrt{5}}|1\rangle_z \otimes |1\rangle_z + \frac{1}{\sqrt{5}}|0\rangle_z \otimes |-1\rangle_z + \frac{\sqrt{3}}{\sqrt{5}}|-1\rangle_z \otimes |0\rangle_z$$

of a pair of spin-1 particles is entangled.

[**HINT:** You will need to consider a tensor product of the form  $(a|1\rangle_z + b|0\rangle_z + c|-1\rangle_z) \otimes (d|1\rangle_z + e|0\rangle_z + f|-1\rangle_z)$ .]

(3) Calculate the linear entropy of entanglement for the following states:

$$\begin{aligned} |\Psi_i\rangle &= \frac{1}{\sqrt{5}}|1\rangle_z \otimes |1\rangle_z + \frac{1}{\sqrt{5}}|0\rangle_z \otimes |-1\rangle_z + \frac{\sqrt{3}}{\sqrt{5}}|-1\rangle_z \otimes |0\rangle_z , \\ |\Psi_{ii}\rangle &= |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z , \\ |\Psi_{iii}\rangle &= \frac{1}{\sqrt{2}} \left( |-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \right) . \end{aligned}$$

(4) The density operator for the combined cat-atom system is given by

$$\hat{\rho}_{CA}(1\text{hr}) = \left( |\Psi(1\text{hr})\rangle_{CA} \right) \left( {}_{CA}\langle \Psi(1\text{hr})| \right)$$

where

$$\begin{aligned} &|\Psi(1\text{hr})\rangle_{CA} \\ &= \frac{1}{\sqrt{2}} \left( |\text{alive}\rangle_C \otimes |\text{undecayed}\rangle_A + |\text{dead}\rangle_C \otimes |\text{decayed}\rangle_A \right) . \end{aligned}$$

The **partial trace** with respect to the atom's states is given by

$$\begin{aligned} \hat{\rho}_C^{(\text{reduced})}(1\text{hr}) &= {}_A\langle \text{undecayed} | \hat{\rho}_{CA}(1\text{hr}) | \text{undecayed}\rangle_A \\ &\quad + {}_A\langle \text{decayed} | \hat{\rho}_{CA}(1\text{hr}) | \text{decayed}\rangle_A \quad (4.9) \end{aligned}$$

where  $\hat{\rho}_C^{(\text{reduced})}(1\text{hr})$  is the **reduced density operator** which describes the state of the cat only. Show that this expression equates to

$$\hat{\rho}_C^{(\text{reduced})}(1\text{hr}) = \frac{1}{2} \left( |\text{alive}\rangle_C \langle \text{alive}| + |\text{dead}\rangle_C \langle \text{dead}| \right) . \quad (4.10)$$

# Advanced Quantum Theory

## 4 Quantum Entanglement

### Solutions to Problems

(1) We can rewrite the partial trace as

$$\begin{aligned}
 & \text{Tr}_2 \left[ \left( |A\rangle_1 \otimes |x\rangle_2 \right) \left( {}_1\langle B| \otimes {}_2\langle y| \right) \right] \\
 &= \text{Tr}_2 \left[ \left( |A\rangle_{11}\langle B| \right) \otimes \left( |x\rangle_{22}\langle y| \right) \right] \\
 &= \left( |A\rangle_{11}\langle B| \right) \text{Tr}_2 \left( |x\rangle_{22}\langle y| \right) \\
 &= \left( |A\rangle_{11}\langle B| \right) \sum_n {}_2\langle \phi_n | \left( |x\rangle_{22}\langle y| \right) |\phi_n\rangle_2 \\
 &= \left( |A\rangle_{11}\langle B| \right) \sum_n {}_2\langle \phi_n | x \rangle_{22} \langle y | \phi_n \rangle_2
 \end{aligned}$$

Reordering the inner products (which are simple complex numbers) yields

$$\begin{aligned}
 & \text{Tr}_2 \left[ \left( |A\rangle_1 \otimes |x\rangle_2 \right) \left( {}_1\langle B| \otimes {}_2\langle y| \right) \right] \\
 &= \left( |A\rangle_{11}\langle B| \right) \sum_n {}_2\langle y | \phi_n \rangle_{22} \langle \phi_n | x \rangle_2 \\
 &= \left( |A\rangle_{11}\langle B| \right) {}_2\langle y | \left( \sum_n |\phi_n\rangle_{22}\langle \phi_n| \right) |x\rangle_2 ,
 \end{aligned}$$

then recognising the resolution of the identity in the  $\{|\phi_n\rangle_2\}$  basis we find

$$\begin{aligned}
 \text{Tr}_2 \left[ \left( |A\rangle_1 \otimes |x\rangle_2 \right) \left( {}_1\langle B| \otimes {}_2\langle y| \right) \right] &= \left( |A\rangle_{11}\langle B| \right) {}_2\langle y | \hat{\mathbf{1}} | x \rangle_2 \\
 &= \left( |A\rangle_{11}\langle B| \right) {}_2\langle y | x \rangle_2 ,
 \end{aligned}$$

and so

$$\text{Tr}_2 \left[ \left( |A\rangle_1 \otimes |x\rangle_2 \right) \left( {}_1\langle B| \otimes {}_2\langle y| \right) \right] = \begin{cases} |A\rangle_{11}\langle B| , & \text{if } x = y \\ 0 , & \text{if } x \neq y \end{cases} .$$

(2) As in **Example (2)** on page 8 of the lecture notes we use *reductio ad absurdum*. We first **assume** that  $|\psi\rangle$  can be written as a separable state of the form suggested by the

Hint:

$$\begin{aligned}
 |\psi\rangle &= \left( a|1\rangle_z + b|0\rangle_z + c|-1\rangle_z \right) \otimes \left( d|1\rangle_z + e|0\rangle_z + f|-1\rangle_z \right) \\
 &= ad|1\rangle_z \otimes |1\rangle_z + ae|1\rangle_z \otimes |0\rangle_z + af|1\rangle_z \otimes |-1\rangle_z \\
 &\quad + bd|0\rangle_z \otimes |1\rangle_z + be|0\rangle_z \otimes |0\rangle_z + bf|0\rangle_z \otimes |-1\rangle_z \\
 &\quad + cd|-1\rangle_z \otimes |1\rangle_z + ce|-1\rangle_z \otimes |0\rangle_z + cf|-1\rangle_z \otimes |-1\rangle_z
 \end{aligned}$$

Comparing with

$$\begin{aligned}
 |\Psi\rangle &= \frac{1}{\sqrt{5}}|1\rangle_z \otimes |1\rangle_z + \frac{1}{\sqrt{5}}|0\rangle_z \otimes |-1\rangle_z \\
 &\quad + \frac{\sqrt{3}}{\sqrt{5}}|-1\rangle_z \otimes |0\rangle_z
 \end{aligned}$$

shows that

$$\begin{aligned}
 ad &= \frac{1}{\sqrt{5}} & ae &= 0 & af &= 0 \\
 bd &= 0 & be &= 0 & bf &= \frac{1}{\sqrt{5}} \\
 cd &= 0 & ce &= \frac{\sqrt{3}}{\sqrt{5}} & cf &= 0
 \end{aligned} .$$

Note that  $ad = \frac{1}{\sqrt{5}}$  implies that both  $a$  and  $d$  are non-zero,  $bf = \frac{1}{\sqrt{5}}$  implies that both  $b$  and  $f$  are non-zero and  $ce = \frac{\sqrt{3}}{\sqrt{5}}$  implies that both  $c$  and  $f$  are non-zero. Hence all 6 coefficients,  $a, b, c, d, e, f$  are nonzero. But if this were true then it is impossible that any products, such as  $ae$ , is zero. Hence the assumption leads to a **contradiction** and so is false. Thus the state is inseparable, i.e. **it is entangled**.

(3) For  $|\Psi_i\rangle$  the partial trace over system 2 is given by

$$\begin{aligned}
 \hat{\rho}_1 &= \text{Tr}_2(|\Psi_i\rangle\langle\Psi_i|) \\
 &= \text{Tr}_2\left[\left(\frac{1}{\sqrt{5}}|1\rangle_z \otimes |1\rangle_z + \frac{1}{\sqrt{5}}|0\rangle_z \otimes |-1\rangle_z + \frac{\sqrt{3}}{\sqrt{5}}|-1\rangle_z \otimes |0\rangle_z\right)\right. \\
 &\quad \left.\left(\frac{1}{\sqrt{5}}z\langle 1| \otimes z\langle 1| + \frac{1}{\sqrt{5}}z\langle 0| \otimes z\langle -1| + \frac{\sqrt{3}}{\sqrt{5}}z\langle -1| \otimes z\langle 0|\right)\right]
 \end{aligned}$$

and using Eq. (4.1) we find

$$\hat{\rho}_1 = \frac{1}{5}|1\rangle_{zz}\langle 1| + \frac{1}{5}|0\rangle_{zz}\langle 0| + \frac{3}{5}|-1\rangle_{zz}\langle -1| .$$

The purity of this reduced density operators is

$$\begin{aligned}
 \text{Tr}_1\left(\hat{\rho}_1^{(i)}\right)^2 &= \text{Tr}_1\left[\left(\frac{1}{5}|1\rangle_{zz}\langle 1| + \frac{1}{5}|0\rangle_{zz}\langle 0| + \frac{3}{5}|-1\rangle_{zz}\langle -1|\right)\right. \\
 &\quad \times \left.\left(\frac{1}{5}|1\rangle_{zz}\langle 1| + \frac{1}{5}|0\rangle_{zz}\langle 0| + \frac{3}{5}|-1\rangle_{zz}\langle -1|\right)\right] \\
 &= \text{Tr}_1\left[\frac{1}{25}|1\rangle_{zz}\langle 1| + \frac{1}{25}|0\rangle_{zz}\langle 0| + \frac{9}{25}|-1\rangle_{zz}\langle -1|\right] \\
 &= \frac{1}{25} + \frac{1}{25} + \frac{9}{25} = \frac{11}{25} .
 \end{aligned}$$

Hence the linear entropy of entanglement is

$$E_{\text{lin}}(|\Psi_i\rangle) = -\log_2(\frac{11}{25}) \approx 1.2 .$$

For  $|\Psi_{ii}\rangle$  the partial trace over system 2 is given by

$$\begin{aligned} \hat{\rho}_1 &= \text{Tr}_2\left(|\Psi_{ii}\rangle\langle\Psi_{ii}|\right) \\ &= \text{Tr}_2\left[\left(|\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z\right)\left(z\langle\frac{1}{2}| \otimes z\langle-\frac{1}{2}|\right)\right] \\ &= |\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| , \end{aligned}$$

for which the purity is

$$\begin{aligned} \text{Tr}_1\left(\hat{\rho}_1\right)^2 &= \text{Tr}_1\left[\left(|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right)\left(|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right)\right] \\ &= \text{Tr}_1\left[|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right] \\ &= 1 , \end{aligned}$$

and so the linear entropy of entanglement is

$$E_{\text{lin}}(|\Psi_{ii}\rangle) = -\log_2(1) = 0 .$$

For  $|\Psi_{iii}\rangle$  the partial trace over system 2 is given by

$$\begin{aligned} \hat{\rho}_1 &= \text{Tr}_2\left(|\Psi_{iii}\rangle\langle\Psi_{iii}|\right) \\ &= \text{Tr}_2\left[\frac{1}{\sqrt{2}}\left(|-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z + |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z\right)\right. \\ &\quad \times \frac{1}{\sqrt{2}}\left(z\langle-\frac{1}{2}| \otimes z\langle\frac{1}{2}| + z\langle\frac{1}{2}| \otimes z\langle-\frac{1}{2}|\right)\left.\right] \\ &= \frac{1}{2}\left(|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + |\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right) , \end{aligned}$$

for which the purity is

$$\begin{aligned} \text{Tr}_1\left(\hat{\rho}_1\right)^2 &= \text{Tr}_1\left[\frac{1}{2}\left(|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + |\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right)\right. \\ &\quad \times \frac{1}{2}\left(|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + |\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right)\left.\right] \\ &= \text{Tr}_1\left[\frac{1}{4}|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}| + \frac{1}{4}|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}|\right] \\ &= \frac{1}{2} , \end{aligned}$$

and so the linear entropy of entanglement is

$$E_{\text{lin}}(|\Psi_{ii}\rangle) = -\log_2(\frac{1}{2}) = 1 .$$

**(4)** We evaluate each term in Eq. (4.9). Using Eq. (4.6) we

find

$$\begin{aligned} {}_A\langle \text{undecayed} | \hat{\rho}_{CA}(1\text{hr}) | \text{undecayed} \rangle_A \\ = {}_A\langle \text{undecayed} | \left( |\Psi(1\text{hr})\rangle_{CA} \right) \left( {}_{CA}\langle \Psi(1\text{hr})| \right) | \text{undecayed} \rangle_A . \end{aligned}$$

Evaluating each inner product separately, i.e.

$$\begin{aligned} {}_A\langle \text{undecayed} | \left( |\Psi(1\text{hr})\rangle_{CA} \right) \\ = {}_A\langle \text{undecayed} | \left[ \frac{1}{\sqrt{2}} \left( |\text{alive}\rangle_C \otimes |\text{undecayed}\rangle_A \right. \right. \\ \left. \left. + |\text{dead}\rangle_C \otimes |\text{decayed}\rangle_A \right) \right] \\ = \frac{1}{\sqrt{2}} |\text{alive}\rangle_C \end{aligned}$$

and similarly

$$\left( {}_{CA}\langle \Psi(1\text{hr}) | \right) | \text{undecayed} \rangle_A = \frac{1}{\sqrt{2}} \left( {}_C\langle \text{alive} | \right) ,$$

shows that

$${}_A\langle \text{undecayed} | \hat{\rho}_{CA}(1\text{hr}) | \text{undecayed} \rangle_A = \frac{1}{2} \left( |\text{alive}\rangle_C \right) \left( {}_C\langle \text{alive} | \right) .$$

The same procedure leads to

$${}_A\langle \text{decayed} | \hat{\rho}_{CA}(1\text{hr}) | \text{decayed} \rangle_A = \frac{1}{2} \left( |\text{dead}\rangle_C \right) \left( {}_C\langle \text{dead} | \right)$$

and so from Eq. (4.9)

$$\hat{\rho}_C^{(\text{reduced})}(1\text{hr}) = \frac{1}{2} \left( |\text{alive}\rangle_{CC} \langle \text{alive}| + |\text{dead}\rangle_{CC} \langle \text{dead}| \right)$$

as required.

# 5. Nonlocality and Bell's Theorem

## a. Reality and Locality in Quantum Interpretations

In Chapter 4 you were introduced to two interpretations of quantum mechanics:

1. **Copenhagen:** Macroscopic observers exist independent of the quantum systems, and collapse the wavefunction of a quantum system whenever they observe it.
2. **Many-Worlds:** Macroscopic observers are quantum systems too, and the wavefunction describes a multiplicity of worlds. Our personal experience of a single world is an **illusion**.

There are obvious objections that can (and have) been raised with both of these interpretations.

### Copenhagen:

1. How can macroscopic observers exist independent of the quantum systems when they are made up of atoms etc. (which are quantum systems)?
2. What defines an observer? (can it be a cat? a baby? an undergraduate?)
3. What defines the moment of observation? (is it seeing? believing? writing it down?)

### Many-Worlds:

1. The evidence for any physical theory comes from observing outcomes of experiments. How can we then adopt an interpretation which says that the outcomes we observed are no more real than all the other possible outcomes which take place in other “worlds” ?
2. What does it mean to say that one outcome is more likely to happen than another outcome, if every outcome happens in some “world” ?

## Einstein on Reality, and Locality

Not surprisingly, many physicists have sought alternative interpretations to quantum mechanics from the above two. The most famous of these physicists is **Einstein**. Einstein was definitely a **realist**. That is, he believed that things had “independent real situations”. Thus it seems almost certain that Einstein would have rejected the many-worlds interpretation (although we’ll never know for sure because he died in 1955, while Everett did not publish until 1957).

Einstein certainly rejected the Copenhagen interpretation. Like Schrödinger, he objected to its vagueness (as set out above, and as illustrated by the Schrödinger cat paradox). However, his chief objection was different to those listed above: it was that the Copenhagen interpretation involved **action-at-a-distance (aaad)**. This was pointed out by Einstein in 1927, only a month after Bohr proposed what became known as the Copenhagen interpretation.\*

Einstein’s argument was very simple, involving just a single quantum particle. He pointed out that the wavefunction  $\psi$  could spread to two (or more) arbitrarily separated locations over time. But the particle can only be detected by one observer. In the Copenhagen interpretation this is explained by *wavefunction collapse* or *wavepacket reduction*. The detection of the particle by an observer at one location *instantaneously* changes the wavefunction over all space, making it zero everywhere except at the location where the particle was observed. Not only is this action-at-a-distance counter-intuitive, it also violates the principle of relativistic causality, that no effect can propagate faster than light. Einstein thought that violating this principle was as “unacceptable” as abandoning reality.

(Einstein also did not like the randomness or indeterminacy in the Copenhagen interpretation, famously saying “God does not play dice.” However most Einstein scholars think that this was a less important objection to him than those mentioned above.)

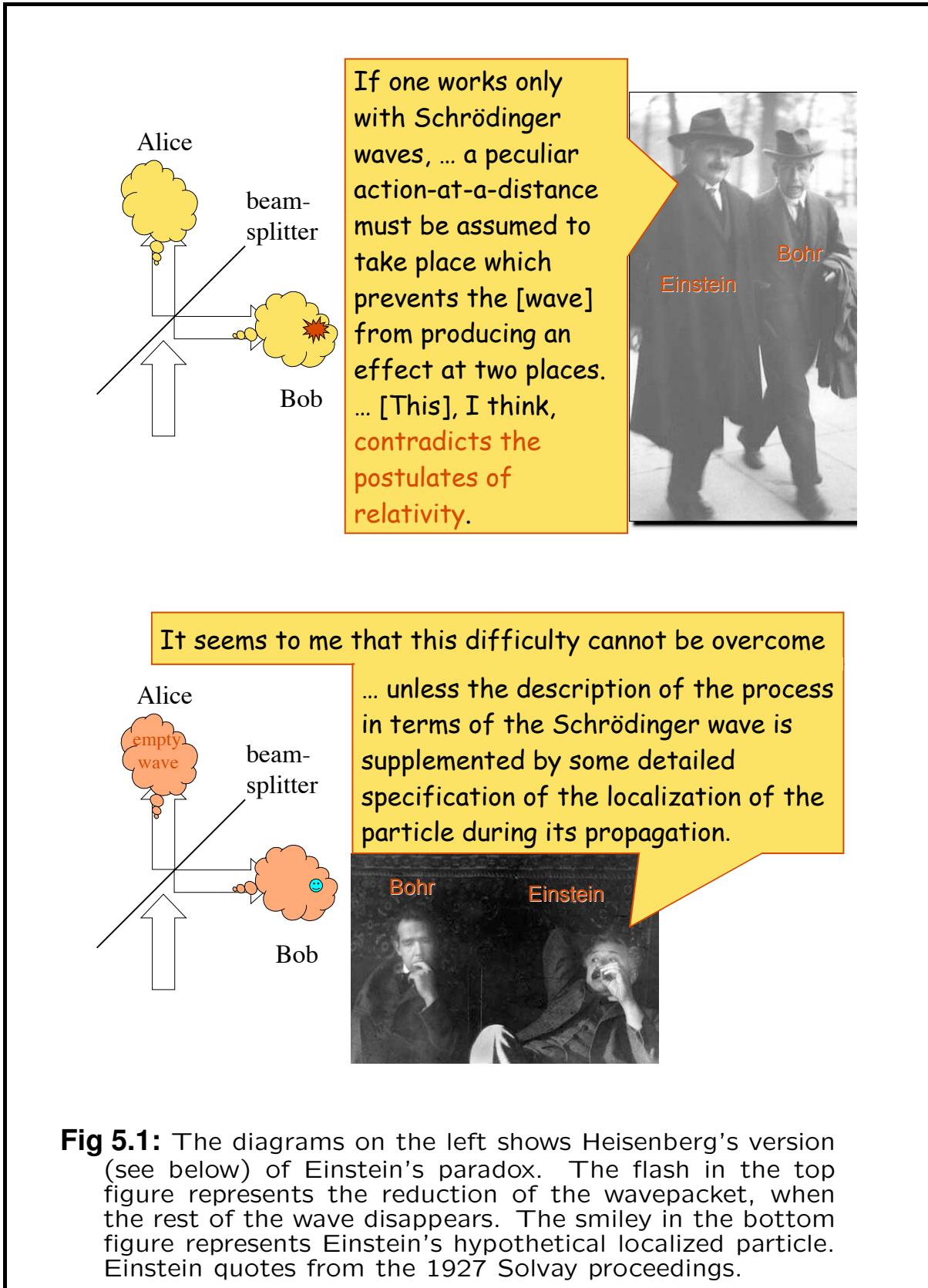
## Einstein and Incompleteness

As Einstein pointed out, a much more natural interpretation for the fact that only one observer sees the particle is that the particle has a real position: the observer (say Bob) who detects

---

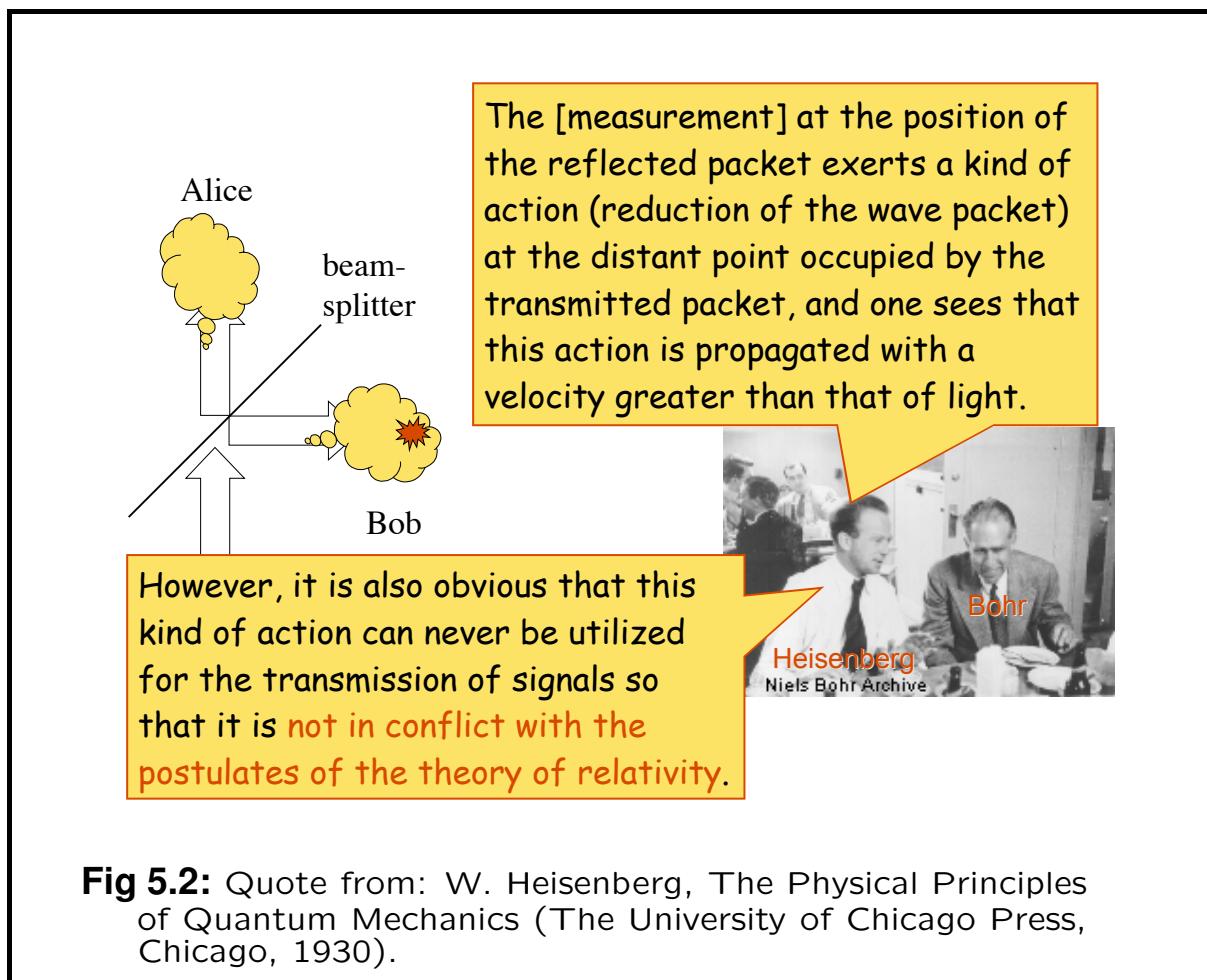
\*See G. Bacciagaluppi & A. Valentini, *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference* (Cambridge University Press, 2009). Several years later, Einstein and his colleagues published a more elaborate and experimentally relevant argument which used a two-particle state. This is the famous **EPR** paper: A. Einstein, B. Podolsky and N. Rosen, “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?”, Phys. Rev. **47**, 777 (1935).

the particle does so because the particle entered his detector, and another observer (say Alice) cannot also detect the particle because in this case it (obviously) did not enter her detector. That is to say, the mysterious action-at-a-distance vanishes if one abandons the central idea of the Copenhagen interpretation that  $\psi$  is a *complete* description of the particle's state. See Figure 1.



## The Response from Copenhagen

The proponents of the Copenhagen interpretation were rather blasé about Einstein's criticisms. They were convinced that the state  $|\psi\rangle$  was a *complete* description of the quantum world, and should not be "supplemented" with any other variables. They did not deny that their interpretation involved action-at-a-distance, but instead argued that it was of a harmless kind. Heisenberg, considering an example like Einstein's, but with the wavefunction split into parts (transmitted and reflected) by a beam splitter. See Figure 2.



**Fig 5.2:** Quote from: W. Heisenberg, The Physical Principles of Quantum Mechanics (The University of Chicago Press, Chicago, 1930).

### Types of Nonlocality

These two quotations above show that physicists can disagree radically over the meaning of words and phrases such as "the postulates of relativity". To avoid misunderstanding it is necessary to be clear. Heisenberg's idea was that the postulates of relativity forbid signalling faster than light. This can be called **signal-locality**. By contrast, Einstein thought these postulates (which were of course first formulated by him!) forbid **aaad**. The latter is closely related to another concept, called **local causality**, which we'll get to eventually. For these reasons I will avoid using the terms "nonlocality" or "locality" in any technical sense, since they mean different things to different people.

## de Broglie–Bohm Mechanics

When Einstein talked about “supplementing” Schrödinger’s wave with a particle with a specific location, he was not talking completely hypothetically. Another founder of quantum mechanics, **de Broglie**, had just (in 1927) proposed an interpretation with exactly this feature, and Einstein said “I think Monsieur de Broglie is right in searching in this direction”. In face of opposition from the Copenhagen school, de Broglie gave up his idea, but it was resurrected and completed by Bohm in 1952.\* For this reason, the theory is known as de Broglie–Bohm (B–B) mechanics.

Consider a single particle with wavefunction  $\psi(\mathbf{r}; t) = \langle \mathbf{r} | \psi(t) \rangle$ . In the Copenhagen interpretation  $P(\mathbf{r}; t) = |\psi(\mathbf{r}; t)|^2$  is the probability to *find* the particle at position  $\mathbf{r} = (x, y, z)$ . In B–B mechanics, the particle has a **pre-existing real position**  $\mathbf{r}$ , and  $P(\mathbf{r}; t)$  is the probability distribution for that position (before you look).

The position  $\mathbf{r}$  evolves according to

$$\dot{\mathbf{r}} = \mathbf{v}(\mathbf{r}; t) \equiv \mathbf{j}(\mathbf{r}; t)/P(\mathbf{r}; t). \quad (5.1)$$

Here

$$\mathbf{j}(\mathbf{r}; t) = (1/m) \operatorname{Re} \psi^*(\mathbf{r}; t) (-i\hbar) \nabla \psi(\mathbf{r}; t) \quad (5.2)$$

is known as the *probability current*. This equation ensures that  $P(\mathbf{r}; t)$  stays equal to  $|\psi(\mathbf{r}; t)|^2$ , because of the continuity equation

$$\dot{P}(\mathbf{r}; t) + \nabla \cdot \mathbf{j}(\mathbf{r}; t) = 0. \quad (5.3)$$

### Problem

Prove that this continuity equation is obeyed for  $P$  and  $j$  as defined above.

Thus in B–B mechanics, the wavefunction *guides* or *pilots* the real position of the particle. For this reason it is also called **pilot-wave theory**. Some example trajectories in B–B mechanics are shown in Figure 3.

For  $N$  particles, the B–B theory describes reality by  $\mathbf{r}$ , now a huge vector containing the positions of all the particles:  $\mathbf{r} = (x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N)$ .

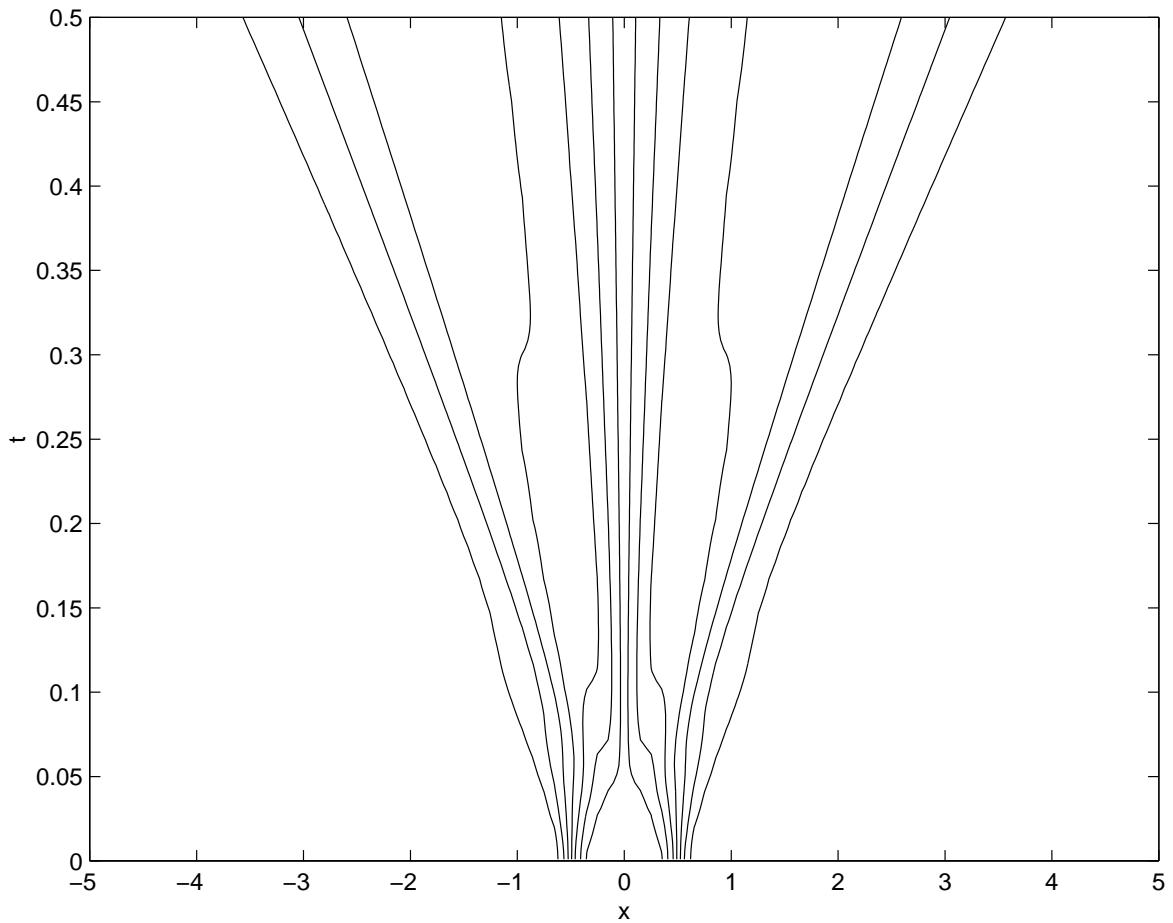
\begin{nonexaminable} The equation for the  $n$ th component  $r_n$  is

$$\dot{r}_n = v_n(\mathbf{r}; t) = \operatorname{Re} \frac{\langle \Psi(t) | \mathbf{r} \rangle \langle \mathbf{r} | i[\hat{H}, \hat{r}_n] | \Psi(t) \rangle}{\hbar \langle \Psi(t) | \mathbf{r} \rangle \langle \mathbf{r} | \Psi(t) \rangle}. \quad (5.4)$$

\end{nonexaminable}

---

\*D. Bohm, “A suggested interpretation of the quantum theory in terms of ‘hidden’ variables” (parts I and II) Phys. Rev. **85**, 166–193 (1952).



**Fig 5.3:** A selection of trajectories of a particle in a double-slit experiment, according to the B–B theory. Note how the particles do not travel in straight lines even after they have left the slits, but instead are guided to form the interference fringes (3 maxima in this case) in the far-field.

Here  $|\Psi(t)\rangle$  is the wavefunction of the universe, including observers. In this respect, the theory is like the **many-worlds** interpretation. The difference is that B–B interpretation has only **one world** which is described by the positions of the particles  $\mathbf{r}$ . All interpretations, (many-worlds, Copenhagen, and B–B mechanics) are **experimentally indistinguishable**.

### Nonlocality in de Broglie–Bohm Mechanics

Note that in Eq. (5.4), the rate of change of the position of the first particle,  $(x_1, y_1, z_1)$ , depends not only on its position, but on the position of *all other particles*, via  $\mathbf{r}$ . In particular this happens if the particles are in an *entangled* state, which could be shared across separate laboratories. The motion of the particles in one lab can (not surprisingly) be affected by the actions experimenters take there, and in Bohmian mechanics this can end up affecting the motion of particles in a distant lab. That is, Bohmian mechanics implies **aaad** — exactly the problem with the Copenhagen interpretation that Einstein was trying to avoid! This type of nonlocality is one reason that many physicists do

not like the B–B interpretation (even though it does not violate *signal-locality*). But remember: the Copenhagen interpretation also has **aaad**, whereas the B–B interpretation avoids the other problems for the Copenhagen interpretation (listed at the start of this chapter), and also those of the many-worlds interpretation. We can summarize the interpretation as:

- 3. de Broglie–Bohm:** Quantum mechanics applies at all scales, but particles have **real** positions. Our experience corresponds to this reality, and is not an illusion.

## Hidden Variable Interpretations

B–B mechanics is an example of a Hidden Variable Interpretation (HVI). In it  $\mathbf{r}$  is the “hidden variable”. It is hidden in the sense that it is not possible to follow the evolution of  $\mathbf{r}$ . That is, it is not possible to see the individual trajectories in Figure 3. (If we could then we could violate signal-locality; that is, we could signal faster than light.) All we can do is to repeat the preparation, and make a measurement at some particular time, in which case we will see the usual quantum statistical distribution  $P(\mathbf{r}; t)$ .

There are in fact infinitely many HVIs that are compatible with our experiencing a real world.<sup>§</sup> This is another reason that some physicists do not like the B–B interpretation, because it is not unique.<sup>¶</sup> A third argument against HVIs in general is that they have more “entities” than either Copenhagen or many-worlds, and so should be cut by Occam’s razor.

## John Bell

The fact that B–B mechanics fails to solve the **aaad** problem in the Copenhagen interpretation led the Belfast-born physicist, John Bell, to ask in 1964: can *any* interpretation solve this problem? By some fairly elementary maths, he managed to prove that the answer is: **no!**\* The presence of **aaad** (or violation of local causality) in the Copenhagen interpretation and in the B–B interpretation is an inherent part of quantum mechanics. This has been called the most profound result of 20th century science. We will talk more about its implications later, but first we will prove his theorem (this will take a while).

---

<sup>§</sup>See: J. Gambetta and H. M. Wiseman, “Modal Dynamics for Positive Operator Measures” Found. of Physics **34**, 419–448 (2004).

<sup>¶</sup>Although see: H. M. Wiseman, “Grounding Bohmian Mechanics in Weak Values and Bayesianism” New Journal of Physics **9**, 165 (2007)

\*J. S. Bell, “On the Einstein Podolsky Rosen Paradox,” Physics **1**, 195–200 (1964).

## b. Bell's Inequality

### Causality

The assumption(s) behind Bell's theorem can be stated in many ways, but the idea Bell eventually settled on was **local causality**<sup>¶</sup>. Before explaining this idea, we start with the simpler idea of causality. This is just the idea that events have causes, an idea that would seem indispensable for doing science. Consider an experimenter Alice, and the following *macroscopic* events:

- Preparation procedure  $c$
- Alice's measurement procedure  $x$  (*freely chosen* by her)
- Alice's measurement outcome  $a$ .

If  $a$  is determined by  $c$  and  $x$ , via some function  $A$ :

$$a = A(x, c), \quad (5.5)$$

then clearly we have causality. Note that we are using the same letters to denote events (points in space time) and their values, but this should not cause confusion.

In many phenomena (including quantum phenomena), the outcome is not predictable given the preparation procedure and the measurement procedure. Rather, all we can say is that there is some probability distribution  $P(a|x, c)$ . The idea of causality is, more or less<sup>†</sup>, that even in this unpredictable case there is a hidden cause  $\lambda$  for the outcome, so that

$$a = A(x, \lambda, c), \quad (5.6)$$

for some function  $A$  (different, obviously from that in Eq. (5.5)). It is always possible to introduce such causes. For example, if  $a$  is a binary variable (e.g. equal to 'heads' or 'tails') we can simply say that  $\lambda$  is a random number  $\in [0, 1]$ , and  $A$  is the function

$$A(x, \lambda, c) = \begin{cases} \text{'heads'} & \text{if } \lambda < P(\text{'heads'}|x, c) \\ \text{'tails'} & \text{otherwise} \end{cases}. \quad (5.7)$$

---

<sup>¶</sup>J. S. Bell, Epistemological Lett. **9**, 11-24 (1976).

<sup>†</sup>The concept defined here is really determinism at a hidden level (not to be confused with predictability from macroscopic events). Causality can be defined as a weaker concept (i.e. one that follows from determinism, but which does not necessarily require it), but the definition is quite complex. See for example H. M. Wiseman and E. G. Cavalcanti, "Causarum Investigatio and the Two Bell's Theorems of John Bell", arXiv:1503.06413 (2015). For our purposes, the definition here will do.

In general, there will be some physical theory which yields more interesting examples of  $A$  and  $\lambda$ . For example,  $c$  could be tossing a coin,  $x$  could be catching it at some particular time  $t_x$ , and  $a$  the outcome (heads or tails). In this case we believe that  $a$  is determined by all sorts of variables like the initial velocity and angular velocity of the coin, gusts of wind etc. that have definite values, but which are not determined precisely by the experimenter and so are effectively random. All of these can be wrapped up in the symbol  $\lambda$ . Given a preparation procedure, all we can say is that there is some probability distribution  $P(\lambda|c)$ , which gives rise to the distribution  $P(a|x, c)$  via the function  $a = A(x, \lambda, c)$ . Specifically,

$$P(a|x, c) = \sum_{\lambda} \delta_{a, A(x, \lambda, c)} P(\lambda|c). \quad (5.8)$$

In a quantum context, we have the following:

- $c$  corresponds to preparing a quantum state  $\rho_c$
- $x$  corresponds to a measuring some observable<sup>§</sup>  $\hat{a}_x$
- $a$  corresponds to an outcome, an eigenvalue of  $\hat{a}_x$
- $\lambda$  (if it is considered) is known as a hidden variable.

B–B theory is an example of a causal theory in which the particle position  $r$ , plus the positions of all the particles in the detector, comprise the hidden variable  $\lambda$  which determines the outcomes for all measurements on the particle. This is true regardless of whether it is a measurement of position (in which case the outcome is just  $r$ ), or whether it is a measurement of momentum, or spin, or any other observable.

## Local Causality

To define the concept of local causality we have to consider two observers. Calling the second observer Bob, we now also have:

- Bob's measurement procedure  $y$  (*freely chosen* by him)
- Bob's measurement outcome  $b$ .

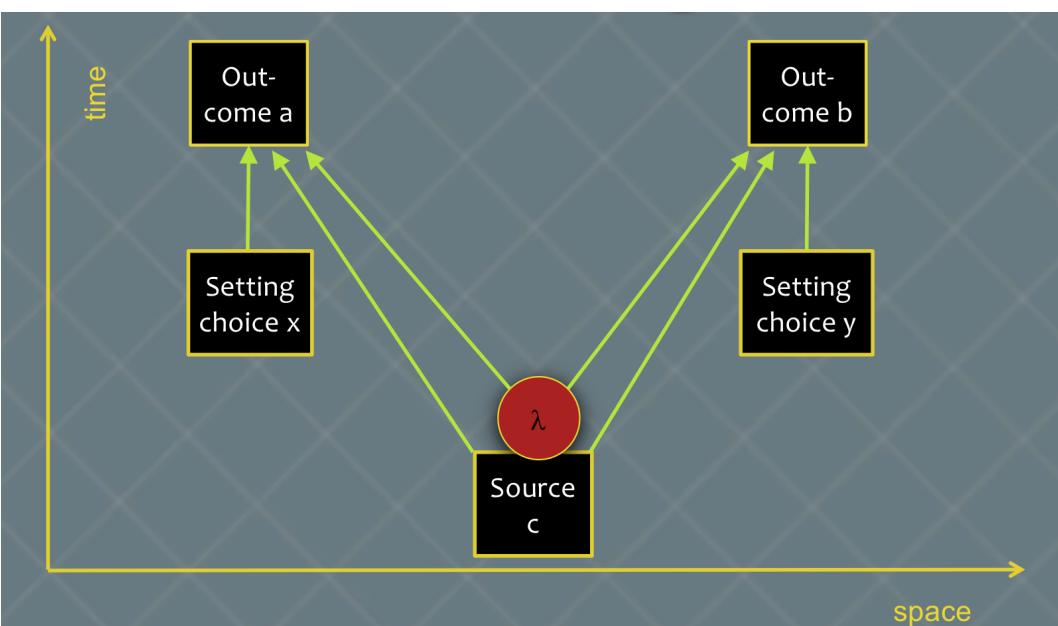
If Alice and Bob were in the same laboratory, then what Bob chooses to do might influence what result Alice gets, and vice versa. Thus in general, instead of having  $a = A(x, \lambda, c)$  we would have  $a = A(x, y, \lambda, c)$  for some new function  $A$ . Similarly we would have  $b = B(y, x, \lambda, c)$ . Note that the preparation  $c$  and the hidden variables  $\lambda$  are common to Alice and Bob.

---

<sup>§</sup>Of course Alice could make a measurement described by a POVM, but for simplicity we restrict to projective measurements in this illustration.

To define the concept of local causality, we have to assume that the macroscopic events  $y$  and  $b$  are *space-like separated* from  $x$  and  $a$ . That is, for example, Alice's could be on the moon while Bob is on the earth, which is 1.3 light-second away. Then if Alice chooses at 12:00:00 which measurement  $a$  she is going to make, and gets the result  $a$  at 12:00:01, and Bob does the same ( $y$  at 12:00:00,  $b$  at 12:00:01), then there is no way a signal can get from Alice to Bob which would allow  $x$  to influence  $b$ , or from Bob to Alice to allow  $y$  to influence  $a$ . Thus, in a locally causal theory, there exist functions  $A$  and  $B$  such that

$$a = A(x, \lambda, c); \quad b = B(y, \lambda, c). \quad (5.9)$$



**Fig 5.4:** Minkowski diagram showing the relative locations of the macroscopic events in space-time.

## Correlations

In the above formulation, Alice's and Bob's results have *common random causes*:  $\lambda$ . Thus it is not surprising that there can be *correlations* between their results. An example from everyday life is the following: Before they open their morning papers, Alice and Bob are in complete ignorance of how much the stock market has gone up or down the day before. As far as they are concerned, this is a random variable  $\lambda$ . But when they open their morning papers at the finance section, if Alice reads a headline "market up" (call this result  $a = +1$ ) then Bob reads the same thing (call this  $b = +1$ ). And if Alice reads "market down" ( $a = -1$ ) then so does Bob ( $b = -1$ ). And this happens day after day after day! The correlation is perfect:

$$\langle AB \rangle = +1 \quad (5.10)$$

even though the individual results are random:  $\langle A \rangle = \langle B \rangle = 0$  (or close to it).

The point is that correlations *per se* are not mysterious. The above correlation has an obvious cause. Other correlations may be more puzzling but still be *locally explicable*. For example:



The spooky correlations in Bell's theorem are often described as being like two dice that, no matter how distant, always roll the same.



**Fig 5.5:** The “New Scientist” explanation of the spooky non-locality in quantum mechanics

This is certainly an odd (and valuable) pair of dice, but actually there is nothing about this that violates local causality. There could be a motor in each die cleverly designed to make the die fall showing a particular number, determined from a list  $\lambda$  of random numbers between 1 and 6, one number for each time  $t$  (in seconds) the die might be rolled. If each die has the same list  $\lambda$  this can explain the correlations.

### Impossible Correlations

In the above examples, the correlations are perfect. Bell’s genius was to recognize that, by using more complicated scenarios in which Alice and Bob have to exercise their choice as to what measurement to perform, local causality puts bounds on how much correlation you can have — bounds which quantum mechanics can violate. Consider for example the following situation:

- Alice has two measurement choices, say  $x = 1$  and  $x = 2$ .
- Bob has two measurement choices, say  $y = 1$  and  $y = 2$ .
- Alice’s outcome  $a$  is always either  $+1$  or  $-1$ .
- Bob’s outcome  $b$  is always either  $+1$  or  $-1$ .

Now Alice randomly chooses whether to do her first measurement ( $x = 1$ ) or her second ( $x = 2$ ), and Bob likewise randomly chooses. Can the results be correlated so that they agree every time *except* when Alice and Bob both choose their second measurement, in which case they always disagree? That is, can we come up with a preparation procedure  $c$  to ensure that every time

$$ab = -1 \quad \text{if } x = y = 2 \tag{5.11}$$

$$ab = +1 \quad \text{otherwise} \tag{5.12}$$

Under the assumption of local causality, Alice's and Bob's measurement results are caused by the shared randomness  $\lambda$ , according to functions  $A$  and  $B$ . For brevity of notation we will suppress the dependence on  $c$  (which we will take to be fixed) and consider four different functions of  $\lambda$ ,

$$A_x(\lambda) \equiv A(x, \lambda, c) \quad (5.13)$$

$$B_y(\lambda) \equiv B(y, \lambda, c). \quad (5.14)$$

If the above conditions (5.11) and (5.12) are to hold every time the preparation and measurement are performed, then they must be true for every value of  $\lambda$ . That is, we require for all  $\lambda$ ,

$$A_2(\lambda)B_2(\lambda) = -1 \quad (5.15)$$

$$A_1(\lambda)B_1(\lambda) = 1; \quad A_1(\lambda)B_2(\lambda) = 1; \quad A_2(\lambda)B_1(\lambda) = 1 \quad (5.16)$$

A little bit of thought shows that it is impossible to satisfy conditions (5.15) and (5.16).

### Problem

Prove that it is impossible to satisfy conditions (5.15) and (5.16). Remember, in these equations,  $A_1(\lambda)$ , for example, is just a number — the value of the function  $A$  for some  $\lambda$  and for  $x = 1$  and  $c$  fixed. For ease of notation, you could drop the  $(\lambda)$  and start by considering a  $\lambda$  for which  $A_1 = 1$ .

If it were possible to satisfy all four conditions simultaneously, then the following product would equal 4:

$$A_1(\lambda)B_1(\lambda) + A_1(\lambda)B_2(\lambda) + A_2(\lambda)B_1(\lambda) - A_2(\lambda)B_2(\lambda), \quad (5.17)$$

because each term equals +1. However, a little further thought shows that it is only ever possible to satisfy three of the four conditions at the same time. Whichever one that is *not* satisfied will have the *wrong sign*, and so will contribute  $-1$  to the sum. Thus local causality means that *at most* Eq. (5.17) equals 2. In an experiment we don't know  $\lambda$ , so we can't check what Eq. (5.17) equals. But if it is at most 2 for every  $\lambda$  then we know that the correlation we can observe (averaging over  $\lambda$ ) satisfies

$$\langle A_1B_1 + A_1B_2 + A_2B_1 - A_2B_2 \rangle \leq 2 \quad (5.18)$$

To reiterate, for any  $x$  and  $y$ ,

$$\langle A_x B_y \rangle = \sum_{\lambda} P(\lambda|c) A_x(\lambda) A_y(\lambda), \quad (5.19)$$

and this is how the bound (2) of the inequality is derived.

Now, under the assumption of local causality,

$$\langle ab \rangle_{xy} = \langle A_x B_y \rangle, \quad (5.20)$$

where  $a$  and  $b$  are the outcomes (random variables) resulting from

measurements with setting  $x$  and  $y$  by Alice and Bob. Thus we obtain, under the assumption of local causality,

$$\langle ab \rangle_{11} + \langle ab \rangle_{12} + \langle ab \rangle_{21} - \langle ab \rangle_{22} \leq 2 \quad (5.21)$$

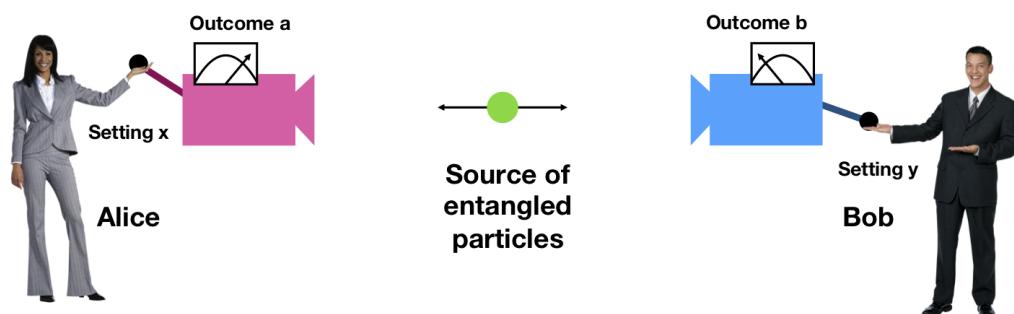
Eq. (5.21) is an example of a **Bell inequality**. (Specifically, this is the Clauser-Horne-Shimony-Holt or CHSH inequality from 1969.) Note that in an experiment we *don't* assume local causality. We don't assume that  $a$  is unaffected by  $y$ , or  $b$  unaffected by  $x$ . We simply repeat the preparation many times, with Alice and Bob randomly choosing the measurements, to create a large ensemble of results. Then later they can pool their results to work out the four separate terms, such as  $\langle ab \rangle_{11}$ , from the four subensembles. While Eq. (5.21) is therefore the correct way to write a Bell inequality for experiment, we will sometimes use the form in Eq. (5.18) as if it applies to experiment (even though that is an abuse of notation) because it is more convenient for calculations.

### c. Bell's Theorem

Bell's theorem is that quantum mechanics allows the above inequality to be violated. That is, it violates local causality.

#### Entangled states

This is only possible using *entangled states*. The explanation in the Copenhagen interpretation is that when one party (Alice) measures her system, that causes the state of Bob's system to instantaneously collapse — a faster-than-light effect.



**Fig 5.6:** Cartoon of a Bell experiment.

Consider the following state for two particles with entangled spin states:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle|\uparrow\rangle + |\downarrow\rangle|\downarrow\rangle). \quad (5.22)$$

Note that we have omitted the  $\otimes$  symbol, as is common. This could represent a pair of electrons, with  $|\uparrow\rangle$  being a spin-up and

$|\downarrow\rangle$  being a spin-down state. The up and down states are orthogonal:  $\langle \uparrow|\downarrow\rangle = 0$ . The electrons also have a wavefunction for their position, and we assume that the first electron (in the order that the kets above are written) is well-localized in Alice's lab, while the second is well-localized in Bob's lab. To produce a state like this, the electrons could have been made to interact at some time in the past, and then been carefully transported to the two different labs <sup>&</sup>. All of this is similar to the entangled state for photons (also with two internal states, the two orthogonal polarizations) considered in a previous chapter, but the notation for spin makes it easier to see how Bell's inequality is violated.

## Perfect correlations with no measurement choice

Clearly if Alice and Bob both measure their electron spin in the  $z$ -direction, that is the  $\hat{z}$  direction, they will get perfectly correlated results. If Alice gets the result  $\uparrow$ , that will collapses Bob's state into the  $|\uparrow\rangle$  state. Formally, we can take Alice's two measurement operators (which are projectors) to be

$$\hat{M}_\uparrow = |\uparrow\rangle\langle\uparrow| \otimes I, \quad (5.23)$$

$$\hat{M}_\downarrow = |\downarrow\rangle\langle\downarrow| \otimes I. \quad (5.24)$$

### Problem

Prove that these measurement operators obey the completeness relation (i.e. that the projectors sum to the identity).

Then, for example, if Alice gets the result  $\uparrow$ , the system collapses into the new state

$$\hat{M}_\uparrow|\Psi\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle|\uparrow\rangle, \quad (5.25)$$

so that Bob will also get the result  $\uparrow$ . The norm of this unnormalized state gives the probability of this result:

$$P_{\uparrow\uparrow} = \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2} \quad (5.26)$$

The other possible result is of course  $\downarrow\downarrow$ , also with probability  $\frac{1}{2}$ . Note the **aaad** in the collapse: Bob's reduced state collapses instantaneously from  $\rho_B = I/2$  to  $\rho_B = |\uparrow\rangle\langle\uparrow|$  or  $|\downarrow\rangle\langle\downarrow|$ .

We can also calculate this perfect correlation by defining observables represented by operators. If we assign the number  $+1$  to the result  $\uparrow$  when Alice measures in the  $\hat{z}$  direction, and  $-1$  to the result  $\downarrow$ , then this measurement is represented by the operator

$$\hat{a}_\uparrow = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|. \quad (5.27)$$

Note that the subscript indicates which direction is assigned  $+1$ ;

---

<sup>&</sup>Though this is not how it was actually done in B. Hensen *et al.*, "Loophole-free Bell inequality violation using electron spins separated by 1.3 Kilometres", Nature **526**, 682 (2015) (see later.)

the opposite direction is automatically assigned  $-1$ . Similarly for Bob we define (where now the states represent Bob's states of course)

$$\hat{b}_\uparrow = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow| \quad (5.28)$$

Now we find (with that abuse of notation referred to earlier)

$$\langle A_\uparrow B_\uparrow \rangle = \langle \Psi | \hat{a}_\uparrow \otimes \hat{b}_\uparrow | \Psi \rangle \quad (5.29)$$

$$= \frac{1}{2} (\langle \uparrow|\langle\uparrow| + \langle\downarrow|\langle\downarrow|) \hat{a}_\uparrow \otimes \hat{b}_\uparrow (|\uparrow\rangle|\uparrow\rangle + |\downarrow\rangle|\downarrow\rangle) \quad (5.30)$$

$$= \frac{1}{2} (\langle \uparrow|\langle\uparrow| + \langle\downarrow|\langle\downarrow|)(|\uparrow\rangle|\uparrow\rangle + (-1)^2 |\downarrow\rangle|\downarrow\rangle) \quad (5.31)$$

$$= \frac{1}{2} (\langle \uparrow|\langle\uparrow| + \langle\downarrow|\langle\downarrow|)(|\uparrow\rangle|\uparrow\rangle + |\downarrow\rangle|\downarrow\rangle) \quad (5.32)$$

$$= +1. \quad (5.33)$$

If this was all that Alice and Bob could do, then the situation would be exactly analogous to the scenario Einstein and Heisenberg discussed with a single particle. The correlations, Einstein would argue, are not mysterious; they are just explained in a mysterious way in the Copenhagen interpretation. The natural explanation of the above phenomenon is that the two electrons are either both spinning up, or both spinning down, and the measurements just *reveal* this fact. (This is also just like the correlated stock market reports in the newspapers that Alice and Bob read, as in Eq. (5.10).) That is, there is a local explanation of the correlations. To find an example where the correlations *cannot* be explained in this way, we have to consider the situation where Alice and Bob can choose between different measurements.

### Perfect (or no) correlations with measurement choice

As a warm up, consider if Alice and Bob can choose instead to measure spin in the  $\leftrightarrow$  direction. This is a basis change, with

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle) \quad (5.34)$$

$$|\leftarrow\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle) \quad (5.35)$$

Assigning result  $+1$  to the  $\rightarrow$  direction and  $-1$  to the  $\leftarrow$  direction, this is represented by the operator

$$\hat{a}_\rightarrow = |\rightarrow\rangle\langle\rightarrow| - |\leftarrow\rangle\langle\leftarrow| \quad (5.36)$$

$$= |\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow| \quad (5.37)$$

#### Problem

Show the second line here from the first line.

We can define analogously for Bob

$$\hat{b}_\rightarrow = |\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|. \quad (5.38)$$

Now as well as Eq. (5.33) we can work out correlations like

$$\langle A_{\uparrow}B_{\rightarrow} \rangle = \langle \Psi | \hat{a}_{\uparrow} \otimes \hat{b}_{\rightarrow} | \Psi \rangle \quad (5.39)$$

$$= \frac{1}{2}(\langle \uparrow | \langle \uparrow | + \langle \downarrow | \langle \downarrow |) \hat{a}_{\uparrow} \otimes \hat{b}_{\rightarrow} (\langle \uparrow | \uparrow \rangle + \langle \downarrow | \downarrow \rangle) \quad (5.40)$$

$$= \frac{1}{2}(\langle \uparrow | \langle \uparrow | + \langle \downarrow | \langle \downarrow |)(\langle \uparrow | \downarrow \rangle - \langle \downarrow | \uparrow \rangle) \quad (5.41)$$

$$= 0 \quad (5.42)$$

and, similarly

$$\langle A_{\rightarrow}B_{\uparrow} \rangle = 0, \quad (5.43)$$

while

$$\langle A_{\rightarrow}B_{\rightarrow} \rangle = +1 \quad (5.44)$$

### Problem

Show Eq. (5.44).

## Imperfect correlations violating a Bell Inequality

These correlations when Alice and Bob can each make two different measurements are not quite so easy to explain as the case where they can only make one measurement. But in fact they are also explicable locally, as Einstein pointed out using a similar example in the EPR paper of 1935<sup>¶</sup>. However, something interesting happens, which Einstein didn't expect, if we change Bob's two measurements from  $B_{\uparrow}$  and  $B_{\rightarrow}$  to  $B_{\nearrow}$  and  $B_{\nwarrow}$ . These measurements are represented by the operators

$$\hat{b}_{\nearrow} = (\hat{b}_{\uparrow} + \hat{b}_{\rightarrow})/\sqrt{2}, \quad (5.45)$$

$$\hat{b}_{\nwarrow} = (\hat{b}_{\uparrow} - \hat{b}_{\rightarrow})/\sqrt{2}. \quad (5.46)$$

### Problem

Show that these operators have eigenvalues of  $\pm 1$ .

**Hint:** One way is to write each operator as a  $2 \times 2$  matrix in the  $|\uparrow\rangle, |\downarrow\rangle$  basis, for example, and find the eigenvalues in the usual way.

**Hint:** Another way is, from the operator representation (or the matrix representation) show that each operator squares to the identity. Remember, for an Hermitian operator  $\hat{X}$ , the eigenvalues of  $\hat{X}^2$  are the squares of the eigenvalues of  $\hat{X}$ . (This is obvious if you think in the diagonal representation.)

Why would we consider these particular measurements? Well consider the CHSH inequality, where we take  $A_1 = A_{\uparrow}$  and  $A_2 = A_{\rightarrow}$  (the observables we have been considering all have results  $\pm 1$

---

<sup>¶</sup>A. Einstein, B. Podolsky and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?", Phys. Rev. **47**, 777 (1935).

as assumed in deriving the inequality):

$$\langle A_{\uparrow}B_1 + A_{\uparrow}B_2 + A_{\rightarrow}B_1 - A_{\rightarrow}B_2 \rangle \leq 2 \quad (5.47)$$

If we want to maximize the LHS, in order to try to violate the inequality, we want  $B_1$  to be correlated with  $A_{\uparrow}$  and  $A_{\rightarrow}$ , and we want  $B_2$  to be correlated with  $A_{\uparrow}$ , but we want it to be *anticorrelated* with  $A_{\rightarrow}$ . Since we know that  $B_{\rightarrow}$  is correlated with  $A_{\rightarrow}$  and likewise  $B_{\uparrow}$  with  $A_{\uparrow}$ , it makes sense to choose  $B_1$  to be pointing midway between  $B_{\uparrow}$  and  $B_{\rightarrow}$ , but  $B_2$  to be midway between  $B_{\uparrow}$  and  $-B_{\rightarrow}$ . That is, we want to choose  $\hat{b}_1 = \hat{b}_{\nearrow}$  and  $\hat{b}_2 = \hat{b}_{\nwarrow}$ .

Thus we want to evaluate

$$C = \langle A_{\uparrow}B_{\nearrow} + A_{\uparrow}B_{\nwarrow} + A_{\rightarrow}B_{\nearrow} - A_{\rightarrow}B_{\nwarrow} \rangle \quad (5.48)$$

Using the result (5.33), (5.42), (5.43), and (5.44), plus the definitions (5.45) and (5.46), we have

$$\langle A_{\uparrow}B_{\nearrow} \rangle = \langle \hat{a}_{\uparrow} \otimes \hat{b}_{\nearrow} \rangle \quad (5.49)$$

$$= \langle \hat{a}_{\uparrow} \otimes \frac{1}{\sqrt{2}}(\hat{b}_{\uparrow} + \hat{b}_{\rightarrow}) \rangle \quad (5.50)$$

$$= \frac{1}{\sqrt{2}}\langle \hat{a}_{\uparrow} \otimes \hat{b}_{\uparrow} \rangle + \frac{1}{\sqrt{2}}\langle \hat{a}_{\uparrow} \otimes \hat{b}_{\rightarrow} \rangle \quad (5.51)$$

$$= \frac{1}{\sqrt{2}} + 0 = \frac{1}{\sqrt{2}}. \quad (5.52)$$

That is, the correlations are *imperfect* (less than one). Similarly,

$$\langle A_{\uparrow}B_{\nwarrow} \rangle = \frac{1}{\sqrt{2}} \quad (5.53)$$

$$\langle A_{\rightarrow}B_{\nearrow} \rangle = \frac{1}{\sqrt{2}} \quad (5.54)$$

$$\langle A_{\rightarrow}B_{\nwarrow} \rangle = -\frac{1}{\sqrt{2}}. \quad (5.55)$$

### Problem

Show these.

Thus finally we have

$$C = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - \left(-\frac{1}{\sqrt{2}}\right) = 2\sqrt{2} > 2 \quad (5.56)$$

That is, the Bell inequality is violated!

### Experimental confirmation

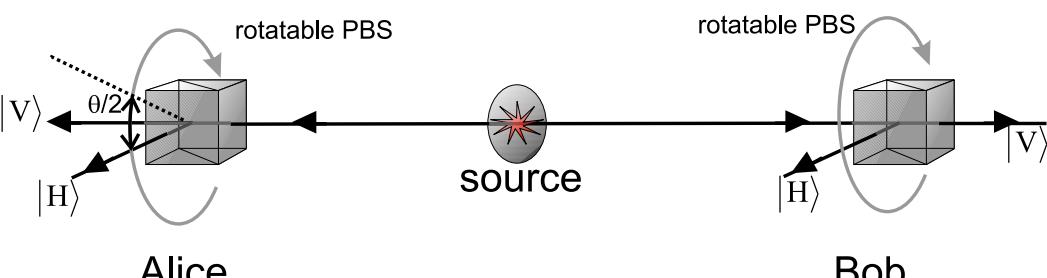
Since the 1970s, physicists have done experiments to test this prediction of quantum mechanics. The landmark experiments are generally regarded as those performed in the early 1980s by Aspect and colleagues<sup>§</sup>. But it took more than a generation

---

<sup>§</sup>A. Aspect, P. Grangier and G. Roger, "Experimental Tests of Realistic Local Theories via Bell's Theorem", Phys. Rev. Lett. **47**, 460 (1981)

until an experiment was done that closed (essentially) all of the ‘loopholes’ in those early experiments<sup>&</sup>. All of these experiments are consistent with quantum theory and show violations of Bell inequalities.

Most experiments have been done with photons, as illustrated in Fig. 4. Here H and V correspond to two orthogonal results, such as  $\uparrow$  and  $\downarrow$ . The different measurement bases required are easily achieved by rotating a polarizing beam-splitter (PBS), similar to the BB84 protocol discussed in an earlier chapter. Indeed, Bell’s theorem enables quantum key distribution (QKD) which is even more secure than the BB84 protocol. In the BB84 protocol, the parties, Alice and Bob, have to trust that their devices are, respectively, preparing and measuring in the basis that the theory requires. Using suitable entangled states and Bell-inequality violation, it is possible to do **device-independent** QKD. However, this is even harder to achieve than merely violating a Bell inequality, and has not yet been demonstrated in a practical setting.



**Fig 5.7:** Experimental set up used for testing whether Bell’s inequality is violated, using entangled photon pairs.

## Beyond two particles

More than two particles can be entangled, and this is also interesting for Bell’s theorem. An example is the 3-particle state “GHZ-state”\*.

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left( |\uparrow\rangle_z \otimes |\uparrow\rangle_z \otimes |\uparrow\rangle_z + |\downarrow\rangle_z \otimes |\downarrow\rangle_z \otimes |\downarrow\rangle_z \right) \quad (5.57)$$

### Problem

Prove it is entangled by showing that it cannot be written in the separable form

$$|\psi\rangle = (\alpha|\uparrow\rangle_z + \beta|\downarrow\rangle_z) \otimes (\gamma|\uparrow\rangle_z + \delta|\downarrow\rangle_z) \otimes (\epsilon|\uparrow\rangle_z + \zeta|\downarrow\rangle_z) . \quad (5.58)$$

<sup>&B. Hensen *etal.*, “Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres”, Nature **526**, 682 (2015)</sup>

<sup>\*D.M. Greenberger, M.A. Horne, and A. Zeilinger, “Going Beyond Bell’s Theorem,” in Bell’s Theorem, Quantum Theory, and Conceptions of the Universe, (Kluwer, Dordrecht, The Netherlands, 1989), pp. 73–76.</sup>

This state allows for an elegant proof of Bell's theorem using three parties without using probabilities (all the correlations involved are perfect)<sup>¶</sup>. Moreover, it can be shown that the Bell-nonlocality arising from this state allows those three parties to do useful things that would not be possible in a locally causal universe. As a light-hearted example, it could allow three alleged criminals to evade jail ☺<sup>§</sup>.

## d. What does this violation mean?

The experimental violation of Bell's inequality means that the assumption that went into its derivation must be false. So far we have just stated that assumption as *local causality*. But since we know (now) that this is false, it is useful to break it up into a combination of more elementary assumptions, since this will give us more options about what to give up. That is, (at least!) one of the assumptions must be false, but different interpretations of quantum mechanics will have a different 'opinion' about which one it is. Below we examine all of the assumptions that might reasonably be questioned. Whichever one it is, it means that nature is profoundly counter-intuitive.

### No action-at-a-distance

The most obvious solution is to allow action-at-a-distance. This is what is given up in two of the most popular interpretations, the **Copenhagen** interpretation and the **de Broglie–Bohm** interpretation. There are also **retrocausal** interpretations in which influences travel backwards in time, which are also nonlocal (remember that in relativity travel faster than light and travel backwards in time are indistinguishable).

If we allowed faster than light influences then there would be no reason to assume that  $a = A_x(\lambda) = A(x, \lambda, c)$ . Instead we would have  $a = A_{x,y}(\lambda) = A(x, y, \lambda, c)$ . That is, Alice's result could be influenced by what Bob chooses to measure. Then it would be easy to violate Bell's inequality. Indeed, by choosing

$$A_{11} = B_{11} = A_{12} = B_{12} = A_{21} = B_{21} = A_{22} = 1, \quad (5.59)$$

and

$$B_{22} = -1, \quad (5.60)$$

we could obtain

$$C = \langle (AB)_{11} + (AB)_{12} + (AB)_{21} - (AB)_{22} \rangle = 4. \quad (5.61)$$

That is, we could violate Bell's inequality even more than quantum mechanics does.

<sup>¶</sup>N.D. Mermin, "Quantum Mysteries Revisited," Am. J. Phys. **58**, 731–734 (1990)

<sup>§</sup>K. A. Jacobs and H. M. Wiseman, "An entangled web of crime: Bell's theorem as a short story", Am. J. Phys **73**, 932-937 (2005).

The problem with allowing **aaad** is that it goes against the spirit of the theory of relativity. (Whether it goes against the postulates is a matter of opinion, as we've seen.) This theory definitely requires (and experiment seems to confirm) that it is impossible to *signal* faster than light. So the question is: why is **aaad** built into the world but in such a sneaky way as to prevent us from using it to signal (or using it to discover the preferred reference frame etc.)

## Macro-reality

Another possibility is to give up macro-reality. That is, if we refuse to admit that definite values  $a$  and  $b$  are obtained for the measurements performed by Alice and Bob in their distant laboratories, then the correlation in Bell's inequality is between non-existent entities, and so its violation has no significance. This is the path taken in the popular **Many worlds** interpretation, which says that getting particular measurement results is an illusion. **Solipsism** also takes this path, but in this interpretation not all results are illusions. *My* results are real. But yours are not real, or at least not real until *I* have observed them, when you have communicated them to me.

These approaches might have worrying moral implications for how we treat other people and ourselves.

## Causal explanations

A third possibility is to give up causal explanations. That is, to deny that events are due to causes, to deny that correlations require explanation. If we give this up we are saying that we are not interested in explanations, only prediction. In the quantum context, this is sometimes called a **neo-Copenhagen** or **Quantum Bayesian** interpretation. It means that we should not think of the quantum state  $|\Psi\rangle$  as anything real, but rather just as a way of calculating probabilities in laboratory experiments.

The biggest problem with this philosophy is that it seems incompatible with the scientific method — the way we discover new phenomena and theories is by trying to *explain* observations. Indeed, this is the way we reason in everyday life. These interpretations also have all of the same problems as detailed for the Copenhagen interpretation at the beginning of this chapter.

## Freedom of choice

The last assumption is one that has been implicit so far: that Alice and Bob are free to choose which measurement to perform. If that is not true, then Bell's inequality cannot be derived. For instance, if Alice's and Bob's decisions are determined by  $\lambda$  then  $A_x = A(x, \lambda, c)$  can also depend on  $y$  because  $y$  is determined by  $\lambda$ .

That is, it is really  $A_{x,y}$ . In this case we get back to the analysis above under the *No action-at-a-distance* heading. A lack of freedom of choice means that Alice and Bob are constrained in some way such that their choices can not be regarded as independent and random.

One might hope that this explanation might require giving up only some freedom of choice — that is, that Alice's and Bob's choices might be influenced by  $\lambda$  but would still be free to some extent. However, it has recently been shown\* that this is not the case. Some Bell inequality violations can only be explained (in this way) if their choices are completely determined by  $\lambda$ .

Such constraints have serious implications for science. If we can no longer assume the freedom of choice we need for the Bell theorem, then we cannot assume that the choices made in randomized trials are in fact random. Again, the implications are not limited to science — if all of our actions are predestined this raises questions about how we regard ourselves and the universe.

Whatever way you look at it, Bell's theorem is disturbing.

---

\*G. Pütz, D. Rosset, T. J. Barnea, Y. C. Liang, and N. Gisin, Phys. Rev. Lett. **113**, 190402 (2014)

# Problem Sheet

- (1) Prove that the B–B continuity equation is obeyed for  $P$  and  $j$  as defined above.
- (2) Prove that it is impossible to satisfy conditions (5.15) and (5.16).
- (3) Prove that the measurement operators  $\hat{M}_\uparrow$  and  $\hat{M}_\downarrow$  in Eq. (5.23) and Eq. (5.24) obey the completeness relation (i.e. that the projectors sum to the identity).
- (4) Show that the two expressions on the right sides are equal:

$$\begin{aligned}\hat{a}_\rightarrow &= |\rightarrow\rangle\langle\rightarrow| - |\leftarrow\rangle\langle\leftarrow| \\ &= |\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|\end{aligned}$$

- (5) Show Eq. (5.44).
- (6) Show that the operators  $\hat{b}_\nearrow$  and  $\hat{b}_\nwarrow$  in Eq. (5.45) and Eq. (5.46) have eigenvalues of  $\pm 1$ .
- (7) Show that

$$\begin{aligned}\langle A_\uparrow B_\nwarrow \rangle &= \frac{1}{\sqrt{2}} \\ \langle A_\rightarrow B_\nearrow \rangle &= \frac{1}{\sqrt{2}} \\ \langle A_\rightarrow B_\nwarrow \rangle &= -\frac{1}{\sqrt{2}}.\end{aligned}$$

- (8) Prove that the 3-particle state state

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left( |\uparrow\rangle_z \otimes |\uparrow\rangle_z \otimes |\uparrow\rangle_z + |\downarrow\rangle_z \otimes |\downarrow\rangle_z \otimes |\downarrow\rangle_z \right)$$

is entangled by showing that it cannot be written in the separable form

$$|\psi\rangle = (\alpha|\uparrow\rangle_z + \beta|\downarrow\rangle_z) \otimes (\gamma|\uparrow\rangle_z + \delta|\downarrow\rangle_z) \otimes (\epsilon|\uparrow\rangle_z + \zeta|\downarrow\rangle_z).$$

# 6. Angular Momentum Operators

The treatment of **angular momentum** in quantum physics is important for two reasons. The first is that many system we study experimentally involve the angular momentum, e.g. the states typically used in atomic physics describe the angular momentum of electrons in orbit and the intrinsic spin of the electrons and the nucleus. The conservation of angular momentum allows us to transfer angular momentum in “lumps” of  $\hbar$  from, say, light to atoms or ions and back again. States of different angular momentum are also reasonably easy to distinguish experimentally.

The second reason is that it is really a beautiful theory. Some calculations might seem a little tedious at times, but the theory as a whole has strong patterns and symmetries. The theory of angular momentum is one particular application of **group theory**, which itself is a rich and beautiful theory.

We can categorise angular momentum into 3 classes: intrinsic angular momentum (spin), orbital angular momentum and total angular momentum. We have already discussed spin and the spin operators  $\hat{S}_x$ ,  $\hat{S}_y$  and  $\hat{S}_z$  earlier in this course. In topic **9 Addition of Angular Momentum** we will discuss total angular momentum.

## a. Orbital Angular Momentum

### Review of QMII

Orbital angular momentum was introduced in the Quantum Mechanics II part of 2303BPS Classical and Quantum Physics II. We will briefly review the main results in this section. Reread your Quantum Mechanics II notes for full details.

In classical physics, orbital angular momentum is defined by the equation  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ . Breaking  $\mathbf{L}$  into its  $x$ ,  $y$  and  $z$  components we have:

$$\begin{aligned} L_x &= yp_z - zp_y \\ L_y &= xp_z - zp_x \\ L_z &= xp_y - yp_x \end{aligned}$$

The corresponding angular momentum operators in quantum mechanics are given by replacing the quantities on the right-hand side of Eq. (6.1) by their quantum-mechanical analogues, i.e.

$$\begin{aligned} \hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{L}_y &= \hat{x}\hat{p}_z - \hat{z}\hat{p}_x \\ \hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x. \end{aligned}$$

The operators  $\hat{L}_x$ ,  $\hat{L}_y$  and  $\hat{L}_z$  obey the commutation relation

insert  
1

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z,$$

and, in addition, all **cyclical permutations** of it, that is,

### Commutation relations

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= i\hbar\hat{L}_z \\ [\hat{L}_z, \hat{L}_x] &= i\hbar\hat{L}_y \\ [\hat{L}_y, \hat{L}_z] &= i\hbar\hat{L}_x . \end{aligned} \quad \left. \right\} (6.1)$$

For other combinations we note that as, for any two operators  $\hat{A}$  and  $\hat{B}$ ,  $[\hat{B}, \hat{A}] = -[\hat{A}, \hat{B}]$ , it follows that  $[\hat{L}_y, \hat{L}_x] = -i\hbar\hat{L}_z$  etc.

### Same algebra for different angular momenta

The commutation relations for the three classes of angular momentum, orbital  $\hat{L}$ , spin  $\hat{S}$  and total  $\hat{J}$ , are of the same form, e.g.

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= i\hbar\hat{L}_z , \\ [\hat{S}_x, \hat{S}_y] &= i\hbar\hat{S}_z , \\ [\hat{J}_x, \hat{J}_y] &= i\hbar\hat{J}_z . \end{aligned}$$

The reason for the close similarities is that  $\hat{L}$ ,  $\hat{S}$  and  $\hat{J}$  possess the same underlying algebraic structure due to being **representations of the generators of rotation** in different dimensional Hilbert spaces.

The square of the total orbital angular momentum is given by

$$\hat{\mathbf{L}}^2 \equiv |\hat{\mathbf{L}}|^2 = \hat{\mathbf{L}} \cdot \hat{\mathbf{L}} = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 . \quad (6.2)$$

### Notation

“ $\hat{\mathbf{L}}^2$ ” is the **square of the modulus** of the vector operator  $\hat{\mathbf{L}}$ , and so it is a **scalar operator** as illustrated by the right-hand side of Eq. (6.2). Strictly speaking we should write it as  $|\hat{\mathbf{L}}|^2$ , but this makes the symbolism a little tedious after a while. Many texts adopt the shorthand symbol  $\hat{\mathbf{L}}^2$  which we shall use here. But remember,  $\hat{\mathbf{L}}^2$  is the scalar operator  $\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$  and **not** a vector operator.

$\hat{\mathbf{L}}^2$  commutes with all other angular momentum operators, i.e.

insert  
2

$$[\hat{\mathbf{L}}^2, \hat{L}_\mu] = 0$$

where  $\mu = x, y, z$ . That is,  $\hat{\mathbf{L}}^2$  commutes with each of  $\hat{L}_x$ ,  $\hat{L}_y$  and  $\hat{L}_z$ . This implies that there exist simultaneous eigenstates of  $\hat{\mathbf{L}}^2$  and  $\hat{L}_\mu$ . Typically we choose to find the simultaneous eigenstates

and eigenvalues of  $\hat{L}_z$  and  $\hat{\mathbf{L}}^2$ , i.e.  $|l, m\rangle$  where

$$\begin{aligned}\hat{L}_z|l, m\rangle &= m\hbar|l, m\rangle , \\ \hat{\mathbf{L}}^2|l, m\rangle &= l(l+1)\hbar^2|l, m\rangle .\end{aligned}$$

The symbols  $l$  and  $m$  represent quantum numbers characterising the eigenstates of  $\hat{\mathbf{L}}^2$  and  $L_z$ . The symbol  $l$  is known as the **orbital angular momentum quantum number** and  $m$  is called the **magnetic quantum number** as it can be important to the behaviour of charged particles in magnetic fields. The respective allowed values are

$$\begin{aligned}l &= 0, 1, 2, \dots \\ m &= 0, \pm 1, \pm 2, \dots, \pm l , \quad \text{i.e. } -l \leq m \leq l .\end{aligned}$$

The state  $|l, m\rangle$  represents a state possessing definite values for both orbital angular momentum  $|\hat{\mathbf{L}}|$  and the  $z$  component of orbital angular momentum  $L_z$ , i.e.  $\hbar\sqrt{l(l+1)}$  and  $m\hbar$ , respectively.

For instance, the state  $|2, 1\rangle$  represents an electron with  $\sqrt{6}\hbar$  units of angular momentum whose  $z$  component of angular momentum is  $\hbar$ .

### Review of Spherical Harmonics from ANP

The state  $|l, m\rangle$  has a “concrete” representation as a function of the spatial coordinates. This representation was derived in Atomic and Nuclear Physics. The representation is given by

$$|l, m\rangle = \int_{\mathbf{x}} d^3\mathbf{x} Y_{lm}(\theta, \phi) |\mathbf{x}\rangle ,$$

where  $\mathbf{x}$  is a three-dimensional position vector, and  $\theta$  and  $\phi$  denote, respectively, the radial and azimuthal (declination) angles for spherical polar co-ordinates. The function  $Y_{lm}$  is known as a **spherical harmonic** and, for  $m \geq 0$ , it is given by

$$Y_{lm}(\theta, \phi) = (-1)^m \left[ \frac{(2l+1)(l-m)!}{4\pi(l+m)!} \right]^{1/2} P_l^m(\cos \theta) e^{im\phi} ,$$

where  $P_l^m$  is known as an **associated Legendre function**. This function is defined as the solution to the Legendre equation which is, for any variable  $w$ ,

$$\left\{ (1-w^2) \frac{d^2}{dw^2} - 2w \frac{d}{dw} \left[ l(l-1) - \frac{m^2}{1-w^2} \right] \right\} P_m^l(w) = 0 .$$

When  $m < 0$ , spherical harmonics are defined by the relation

$$Y_{lm} = (-1)^m Y_{l-m}(\theta, \phi) .$$

Note that for many systems the total state  $|\psi\rangle$  is the tensor product of the angular state  $|l, m\rangle$  and some radial state  $|\phi(r)\rangle$ . That is  $|\psi\rangle = |\phi(r)\rangle \otimes |l, m\rangle$ .

## Raising and lowering operators (new material)

Two further operators are needed in angular momentum; these are the **raising** and **lowering** operators:

### Definitions

$$\left. \begin{aligned} \hat{L}_+ &= \hat{L}_x + i\hat{L}_y && \text{(raising operator)} \\ \hat{L}_- &= \hat{L}_x - i\hat{L}_y && \text{(lowering operator)} \\ \hat{L}_+ &= (\hat{L}_-)^{\dagger} \end{aligned} \right\} \quad (6.3)$$

They get their names from the fact that they perform analogous operations to the raising and lowering operators for the electromagnetic field  $\hat{a}^\dagger$  and  $\hat{a}$  that you encountered in QM III.

The commutation-relation properties of  $\hat{L}_+$  and  $\hat{L}_-$  are:

### Commutators

$$\left. \begin{aligned} [\hat{L}^2, \hat{L}_{\pm}] &= 0, \\ [\hat{L}_-, \hat{L}_+] &= -2\hbar\hat{L}_z, \\ \text{and } [\hat{L}_z, \hat{L}_+] &= \hbar\hat{L}_+. \end{aligned} \right\} \quad (6.4)$$

The first is easily proved: as  $[\hat{L}^2, \hat{L}_x] = [\hat{L}^2, \hat{L}_y] = 0$  then  $[\hat{L}^2, \hat{L}_{\pm}] = [\hat{L}^2, \hat{L}_x + i\hat{L}_y] = [\hat{L}^2, \hat{L}_x] + i[\hat{L}^2, \hat{L}_y] = 0$ .

### Problem

Use the communication relations Eq. (6.1) and the definitions Eq. (6.3) to show  $[\hat{L}_-, \hat{L}_+] = -2\hbar\hat{L}_z$  and  $[\hat{L}_z, \hat{L}_+] = \hbar\hat{L}_+$ .

Using these equations, we can determine the action of  $\hat{L}_+$  and  $\hat{L}_-$  on the state  $|l, m\rangle$ . Due to the fact that  $[\hat{L}_z, \hat{L}_+] = \hbar\hat{L}_+$ , it follows that  $\hat{L}_z\hat{L}_+ = \hat{L}_+\hat{L}_z + \hbar\hat{L}_+$ . Using this we find

$$\begin{aligned} [\hat{L}_z, \hat{L}_+]|l, m\rangle &= \hbar\hat{L}_+|l, m\rangle \\ \therefore \hat{L}_z\hat{L}_+|l, m\rangle &= \hat{L}_+\hat{L}_z|l, m\rangle + \hbar\hat{L}_+|l, m\rangle \\ &= \hat{L}_+m\hbar|l, m\rangle + \hbar\hat{L}_+|l, m\rangle = (m+1)\hbar\hat{L}_+|l, m\rangle \end{aligned}$$

$$\text{thus } \hat{L}_z(\hat{L}_+|l, m\rangle) = (m+1)\hbar(\hat{L}_+|l, m\rangle)$$

and so  $(\hat{L}_+|l, m\rangle)$  is an eigenvector of  $\hat{L}_z$  with eigenvalue  $(m+1)\hbar$ .

Thus  $\hat{L}_+$  **raises** the  $m$ th eigenstate of  $\hat{L}_z$  to a state proportional to the  $(m+1)$ th eigenstate.

Performing a similar calculation with  $\hat{L}_-$  in place of  $\hat{L}_+$ , we find

that  $\hat{L}_z(\hat{L}_-|l,m\rangle) = (m-1)\hbar(\hat{L}_-|l,m\rangle)$ . That is,  $\hat{L}_-$  **lowers** the  $m$ th eigenstate of  $\hat{L}_z$  to a state proportional to the  $(m-1)$ th eigenstate.

We now determine the proportionality constant  $\alpha$  where

$$\hat{L}_+|l,m\rangle = \alpha|l,m+1\rangle. \quad (6.5)$$

Let the raised state be  $|\phi\rangle = \hat{L}_+|l,m\rangle = \alpha|l,m+1\rangle$ . Thus

$$\begin{aligned} |\alpha|^2 &= \langle\phi|\phi\rangle \\ &= (\hat{L}_+|l,m\rangle)^\dagger(\hat{L}_+|l,m\rangle) \\ &= (\langle l,m|\hat{L}_+^\dagger)(\hat{L}_+|l,m\rangle) \\ &= \langle l,m|\hat{L}_-L_+|l,m\rangle. \end{aligned} \quad (6.6)$$

Using Eq. (6.4) and the definitions Eq. (6.3) of  $\hat{L}_+$  and  $\hat{L}_-$ , we find that

$$\begin{aligned} [\hat{L}_-, \hat{L}_+] &= \hat{L}_-\hat{L}_+ - \hat{L}_+\hat{L}_- \quad \text{and so} \\ \hat{L}_-\hat{L}_+ &= [\hat{L}_-, \hat{L}_+] + \hat{L}_+\hat{L}_- \\ &= -2\hbar\hat{L}_z + (\hat{L}_x + i\hat{L}_y)(\hat{L}_x - i\hat{L}_y). \end{aligned} \quad (6.7)$$

Expanding the brackets on the right-hand side of the last line yields

$$\begin{aligned} (\hat{L}_x + i\hat{L}_y)(\hat{L}_x - i\hat{L}_y) &= \hat{L}_x^2 + \hat{L}_y^2 - i[\hat{L}_x, \hat{L}_y] \\ &= \hat{L}_x^2 + \hat{L}_y^2 + \hbar L_z \\ &= \hat{\mathbf{L}}^2 - \hat{L}_z^2 + \hbar L_z. \end{aligned}$$

Substituting this result into Eq. (6.7) we obtain

insert  
3

$$\hat{L}_-\hat{L}_+ = \hat{\mathbf{L}}^2 - \hat{L}_z^2 - \hbar\hat{L}_z.$$

Hence Eq. (6.6) becomes

$$\begin{aligned} |\alpha|^2 &= \langle m, l | (\hat{\mathbf{L}}^2 - \hat{L}_z^2 - \hbar\hat{L}_z) | l, m \rangle \\ &= [l(l+1)\hbar^2 - m^2\hbar^2 - m\hbar^2] \langle m, l | m, l \rangle \\ &= [l(l+1) - m(m+1)] \hbar^2 \\ &= [(l+m+1)(l-m)] \hbar^2, \end{aligned} \quad (6.8)$$

and so

$$\alpha = \hbar\sqrt{(l+m+1)(l-m)}$$

where we have assumed that  $\alpha$  is real and positive. Using this in

Eq. (6.5) gives the final result that

$$\hat{L}_+|l, m\rangle = \hbar\sqrt{(l+m+1)(l-m)}|l, m+1\rangle .$$

Performing an almost identical calculation replacing  $\hat{L}_+$  with  $\hat{L}_-$ , we obtain

$$\hat{L}_-|l, m\rangle = \hbar\sqrt{(l-m+1)(l+m)}|l, m-1\rangle .$$

### Action of ladder operators

$$\hat{L}_\pm|l, m\rangle = \hbar\sqrt{(l\pm m+1)(l\mp m)}|l, m\pm 1\rangle \quad (6.9)$$

### Range of $m_l$ values

Notice that Eq. (6.9) with  $\hat{L}_+$  and  $m = l$  is

$$\hat{L}_+|l, l\rangle = \hbar\sqrt{(l+l+1)(l-l)}|\dots\rangle = 0$$

and so  $\hat{L}_+$  cannot raise  $|l, m\rangle$  above  $|l, l\rangle$ . This implies that the largest value of  $m$  is  $l$ . Likewise Eq. (6.9) with  $\hat{L}_-$  and  $m = -l$  is

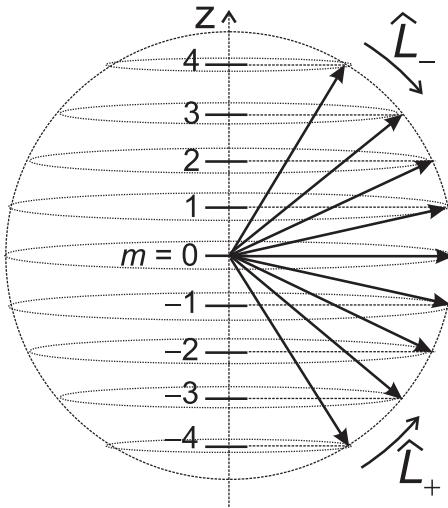
$$\hat{L}_-|l, -l\rangle = \hbar\sqrt{(l+l+1)(l-l)}|\dots\rangle = 0 ,$$

$\hat{L}_-$  cannot lower  $|l, m\rangle$  below  $|l, -l\rangle$  and so  $-l$  is the lowest value of  $m$ . Hence the values of  $m$  are bounded by

$$-l \leq m \leq l$$

which is a result from QMII and ANP.

Physically,  $\hat{L}_\pm$  raise and lower the values of the  $z$  component of orbital angular momentum while leaving the total orbital angular momentum unchanged. For instance, applying  $\hat{L}_+$  repeatedly to the state  $|l, -l\rangle$  yields the “ladder” of states proportional to  $|l, -l+1\rangle, |l, -l+2\rangle, |l, -l+3\rangle$  etc. until we finally arrive at a state proportional to  $|l, l\rangle$  which has the largest possible  $m$  value. This process can be pictorially represented as follows:



**Fig 6.1:** Picture representing the repeated application of  $\hat{L}_+$  increases the value of  $m$  by one while keeping  $l = 4$  constant. The length of each vector is  $\sqrt{4(4 + 1)} = \sqrt{20} \approx 4.5$  and remains unchanged. A similar situation occurs for the lowering operator  $\hat{L}_-$ . Units are chosen such that  $\hbar = 1$ .

## b. Spin and total angular momentum

We have focused on orbital angular momentum  $\hat{\mathbf{L}}$ , however, because spin  $\hat{\mathbf{S}}$  and total angular momentum  $\hat{\mathbf{J}}$  share the same commutations relations as  $\hat{\mathbf{L}}$ , all the results we have derived also apply to these angular momentum quantities as well. By simply replacing the symbols  $L$  and  $l$  with either  $S$  and  $s$  or  $J$  and  $j$  we obtain the analogous results for spin and total angular momentum. The only differences are that the quantum numbers  $l$  for the orbital case are integers whereas  $s$  and  $j$  for spin and total angular momentum can be half integer as well.

# Problem Sheet

- (1) Use the commutation relations Eq. (6.1) and the definitions Eq. (6.3) to show  $[\hat{L}_-, \hat{L}_+] = -2\hbar\hat{L}_z$  and  $[\hat{L}_z, \hat{L}_+] = \hbar\hat{L}_+$ .
- (2) For the state  $|l, m\rangle$ , what are the allowed values of  $m$  when  $l = 2$ . That is, what values can  $m$  be? If we measure  $\hat{\mathbf{L}}^2$  what value will we obtain?
- (3) Show that the commutation relation  $[\hat{L}_x, \hat{\mathbf{L}}^2] = 0$  holds.

[**HINT:** You may find the operator identity  $[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}]$  useful.]

# Advanced Quantum Theory

## 6 Angular Momentum Operators

### Solutions to Problems

- (1)** Substituting the definitions in Eq. (6.3) to write everything in terms of  $\hat{L}_x$ ,  $\hat{L}_y$  and  $\hat{L}_z$  and expanding using the bi-linearity of the commutator

$$\begin{aligned}
 [\hat{L}_-, \hat{L}_+] &= [\hat{L}_x - i\hat{L}_y, \hat{L}_x + i\hat{L}_y] \\
 &= [\hat{L}_x, \hat{L}_x] - i[\hat{L}_y, \hat{L}_x] + i[\hat{L}_x, \hat{L}_y] + [\hat{L}_y, \hat{L}_y] \\
 &= i[\hat{L}_x, \hat{L}_y] + i[\hat{L}_x, \hat{L}_y] \\
 &= 2i[\hat{L}_x, \hat{L}_y] \\
 &= -2\hbar\hat{L}_z ,
 \end{aligned}$$

where the final line comes from the definitions in equation Eq. (6.1). Similarly for the second commutator,

$$\begin{aligned}
 [\hat{L}_z, \hat{L}_+] &= [\hat{L}_z, \hat{L}_x + i\hat{L}_y] \\
 &= [\hat{L}_z, \hat{L}_x] + i[\hat{L}_z, \hat{L}_y] \\
 &= [\hat{L}_z, \hat{L}_x] - i[\hat{L}_y, \hat{L}_z] \\
 &= i\hbar\hat{L}_y + \hbar\hat{L}_x \\
 &= \hbar(\hat{L}_x + i\hat{L}_y) \\
 &= \hbar\hat{L}_+ .
 \end{aligned}$$

The final line here comes from applying the definitions in equation Eq. (6.3).

- (2)** Noting that  $-l \leq m \leq l$ , all of the  $m$  values must have an integer value as their difference and that the lowest value of  $m$  is  $-l$ , the values of  $m$  must be:

$$m = -2, -1, 0, 1, 2 .$$

The value of a measurement of  $\hat{\mathbf{L}}^2$  for an eigenstate of this operator is given by the corresponding eigenvalue. As

$$\hat{\mathbf{L}}^2|l, m\rangle = l(l+1)\hbar|l, m\rangle$$

and the eigenvalue is independent of  $m$  the value of a measurement for  $l = 2$  would be  $6\hbar$ .

- (3)** First substitute in the definition of  $\hat{\mathbf{L}}^2$  in terms of  $\hat{L}_x$ ,  $\hat{L}_y$  and  $\hat{L}_z$ .

$$\begin{aligned}
 [\hat{L}_x, \hat{\mathbf{L}}^2] &= [\hat{L}_x, \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2] \\
 &= [\hat{L}_x, \hat{L}_x\hat{L}_x + \hat{L}_y\hat{L}_y + \hat{L}_z\hat{L}_z] \\
 &= [\hat{L}_x, \hat{L}_x\hat{L}_x] + [\hat{L}_x, \hat{L}_y\hat{L}_y] + [\hat{L}_x, \hat{L}_z\hat{L}_z] .
 \end{aligned}$$

Considering each term on the right hand side and applying

the rule given in the hint,

$$\begin{aligned}
 [\hat{L}_x, \hat{L}_x \hat{L}_x] &= [\hat{L}_x, \hat{L}_x] \hat{L}_x + \hat{L}_x [\hat{L}_x, \hat{L}_x] \\
 &= 0 \\
 [\hat{L}_x, \hat{L}_y \hat{L}_y] &= [\hat{L}_x, \hat{L}_y] \hat{L}_y + \hat{L}_y [\hat{L}_x, \hat{L}_y] \\
 &= 2i\hbar(\hat{L}_z \hat{L}_y + \hat{L}_y \hat{L}_z) \\
 &= 2i\hbar\{\hat{L}_z, \hat{L}_y\} \\
 [\hat{L}_x, \hat{L}_z \hat{L}_z] &= [\hat{L}_x, \hat{L}_z] \hat{L}_z + \hat{L}_z [\hat{L}_x, \hat{L}_z] \\
 &= -2i\hbar(\hat{L}_y \hat{L}_z + \hat{L}_z \hat{L}_y) \\
 &= -2i\hbar\{\hat{L}_z, \hat{L}_y\}
 \end{aligned}$$

where the curly braces are the anti-commutator defined as  $\{A, B\} = AB + BA$ . Substituting these into the expansion of the commutator we are evaluating gives

$$\begin{aligned}
 [\hat{L}_x, \hat{\mathbf{L}}^2] &= [\hat{L}_x, \hat{L}_x \hat{L}_x] + [\hat{L}_x, \hat{L}_y \hat{L}_y] + [\hat{L}_x, \hat{L}_z \hat{L}_z] \\
 &= 0 + 2i\hbar\{\hat{L}_z, \hat{L}_y\} - 2i\hbar\{\hat{L}_z, \hat{L}_y\} \\
 &= 0 .
 \end{aligned}$$

# 7. Addition of angular momentum

## a. Composite systems involving angular momentum

We often consider collections of systems each of which possess their own angular momentum. For instance, we may be interested in two particles, 1 and 2, each with angular momenta  $\hat{\mathbf{J}}_1$  and  $\hat{\mathbf{J}}_2$  respectively. Note that here the angular momentum could be spin, orbital or some combination. Alternately, we may be interested in **single** particle with two forms of angular momentum, e.g. the orbital ( $\hat{\mathbf{L}}$ ) and spin ( $\hat{\mathbf{S}}$ ) angular momenta. In these cases we need to determine the total angular momentum of the whole system. For example, in the first case the total angular momentum  $\hat{\mathbf{J}}$  is the sum of angular momenta  $\hat{\mathbf{J}}_1$  and  $\hat{\mathbf{J}}_2$  and in the second  $\hat{\mathbf{J}}$  is the sum of orbital  $\hat{\mathbf{L}}$  and spin  $\hat{\mathbf{S}}$  momenta. This leads us to study **the addition of angular momentum**.

We begin by defining a number of relevant operators. Let the total angular momentum of the composite system be represented by  $\hat{\mathbf{J}}$  and the two component angular momenta by  $\hat{\mathbf{J}}_1$  and  $\hat{\mathbf{J}}_2$ , irrespective of whether the component angular momenta is orbital, spin or total. The total angular momentum is given by vector addition

insert  
1

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2 . \quad (7.1)$$

Let  $\hat{J}_{1z}$  and  $\hat{J}_{2z}$  denote the  $z$  components of angular momentum for system 1 and 2, respectively.

Let  $j_k$  and  $m_k$  be the quantum numbers associated with  $\hat{J}_k$  and  $\hat{J}_{kz}$ , respectively, where  $k = 1, 2$ . We write the eigenvalue equations as

$$\begin{aligned}\hat{J}_k^2 |j_1, j_2; m_1, m_2\rangle &= j_k(j_k + 1)\hbar^2 |j_1, j_2; m_1, m_2\rangle \\ \hat{J}_{kz} |j_1, j_2; m_1, m_2\rangle &= m_k \hbar |j_1, j_2; m_1, m_2\rangle ,\end{aligned}$$

where  $k = 1, 2$ . Here  $|j_1, j_2; m_1, m_2\rangle = |j_1, m_1\rangle \otimes |j_2, m_2\rangle$ . These equations show that  $|j_1, j_2; m_1, m_2\rangle$  is a simultaneous eigenstate of  $\hat{J}_1^2$ ,  $\hat{J}_2^2$ ,  $\hat{J}_{1z}$  and  $\hat{J}_{2z}$  and thus that the observables corresponding to these four operators are **compatible**.

### Eigenbasis of $\hat{J}_1^2$ , $\hat{J}_2^2$ , $\hat{J}_{1z}$ and $\hat{J}_{2z}$

Consider the subspace  $H_{j_1, j_2}$  spanned by the orthonormal states  $\{|j_1, j_2; m_1, m_2\rangle : m_1 = -j_1, \dots, j_1; m_2 = -j_2, \dots, j_2\}$  for fixed values of  $j_1$  and  $j_2$ . Each  $m_k$  ranges from  $-j_k$  to  $j_k$  in integer steps, and so there are  $(2j_k + 1)$  values of  $m_k$  for  $k = 1$  or  $2$ . Thus there are  $(2j_1 + 1) \times (2j_2 + 1)$  states in this set. This number is also the dimension of  $H_{j_1, j_2}$ .

## Eigenbasis of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$

There also exists another orthonormal basis for the subspace  $H_{j_1, j_2}$ . The alternate basis consists of the eigenstates of the total angular momentum  $\hat{\mathbf{J}}^2$  and  $z$  component  $\hat{J}_z$ . There many ways to write these eigenstates. We adopt the following notation:  $|J, M\rangle$  is an eigenstate of  $\hat{J}_z$  and  $\hat{\mathbf{J}}^2$  on this subspace (of fixed  $j_1$  and  $j_2$  values) with eigenvalues  $M\hbar$  and  $J(J+1)\hbar^2$ , respectively. That is,

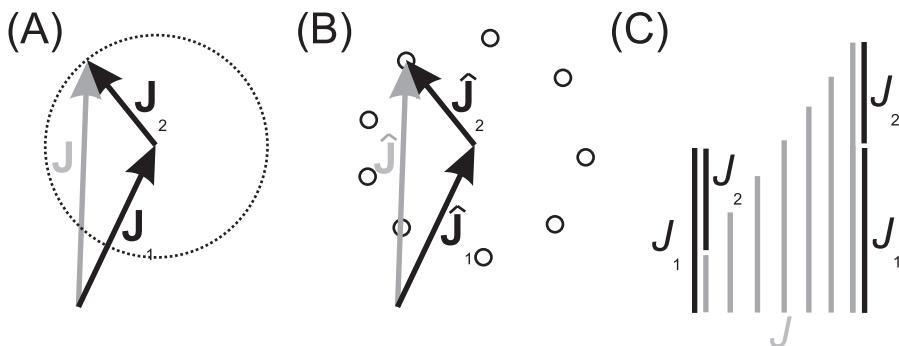
$$\begin{aligned}\hat{J}_z|J, M\rangle &= M\hbar|J, M\rangle \\ \hat{\mathbf{J}}^2|J, M\rangle &= J(J+1)\hbar^2|J, M\rangle.\end{aligned}$$

NOTE:

Some texts use the more-correct notation

$$|j_1, j_2; J, M\rangle$$

in place of our  $|J, M\rangle$  to indicate explicitly that the state belongs to the subspace  $H_{j_1, j_2}$ . However, this notation can become cumbersome. Here we will just remember that the state  $|J, M\rangle$  belongs to  $H_{j_1, j_2}$ .



**Fig 7.1:** Addition of angular momentum (A) classical (B) quantum (C) range of  $J$  values in the quantum case.

The angular momenta are vectors and add vectorially. In classical mechanics the combined angular momentum  $\mathbf{J}$  has a magnitude ranging from  $|\left(|\mathbf{J}_1| - |\mathbf{J}_2|\right)|$  to  $|\mathbf{J}_1| + |\mathbf{J}_2|$ , the actual value depending on the relative orientation of the two component angular momenta. A similar situation exists in the quantum regime. Here the quantum number  $J$  has the range

$$J = |j_1 - j_2|, (|j_1 - j_2| + 1), \dots, (j_1 + j_2). \quad (7.2)$$

Let  $j_1 \geq j_2$ . In this case there are  $2j_2 + 1$  terms in Eq. (7.2) (i.e. count from  $-j_2$  to  $j_2$ ). Thus there are  $2j_2 + 1$  values of  $J$ . For each value of  $J$  there are  $2J + 1$  values of  $M$  (i.e.  $M = -J, -J + 1, \dots, J$ ). Adding the number of  $M$  values for each

value of  $J$  gives

$$\begin{aligned} & \left[ 2(j_1 - j_2) + 1 \right] + \left[ 2(j_1 - j_2 + 1) + 1 \right] + \dots \\ & + \left[ 2(j_1 + j_2 - 1) + 1 \right] \left[ 2(j_1 + j_2) + 1 \right] \quad \left. \right\} (2j_2 + 1) \text{ terms} \\ & = \left[ 2(j_1) + 1 \right] + \left[ 2(j_1) + 1 \right] + \dots + \left[ 2(j_1) + 1 \right] + \left[ 2(j_1) + 1 \right] \\ & \quad (\text{by collecting the } j_2 \text{ terms which sum to zero}) \\ & = (2j_2 + 1) \times (2j_1 + 1). \end{aligned}$$

Hence there are the same number of states in the set  $\{|j_1, j_2 : J, M\rangle : J = |j_1 - j_2|, \dots, j_1 + j_2; M = -J, \dots, J\}$  as there are in  $\{|j_1, m_1, j_2, m_2\rangle : m_1 = -j_1, \dots, j_1; m_2 = -j_2, \dots, j_2\}$ . Both are orthonormal bases for the subspace  $H_{j_1, j_2}$ .

### Relationship between bases: Clebsch-Gordan coefficients

We now determine the relationship between the two bases

$$\{|J, M\rangle : J = |j_1 - j_2|, \dots, j_1 + j_2; M = -J, \dots, J\}$$

and

$$\{|j_1, j_2; m_1, m_2\rangle : m_1 = -j_1, \dots, j_1; m_2 = -j_2, \dots, j_2\}.$$

Using the resolution of the identity in either basis we find

$$|J, M\rangle = \sum_{m_1, m_2} \langle j_1, j_2; m_1, m_2 | J, M \rangle |j_1, j_2; m_1, m_2\rangle \quad (7.3)$$

$$|j_1, j_2; m_1, m_2\rangle = \sum_{J, M} \langle J, M | j_1, j_2; m_1, m_2 \rangle |J, M\rangle.$$

The inner products  $\langle j_1, j_2; m_1, m_2 | J, M \rangle$  (and the complex conjugate  $\langle J, M | j_1, j_2; m_1, m_2 \rangle$ ) are called **Clebsch-Gordan** coefficients. They allow us to transform from one basis to the other. They are calculated as follows.

Following directly from Eq. (7.1) and the definition that  $\hat{\mathbf{J}} = \hat{J}_x \mathbf{i} + \hat{J}_y \mathbf{j} + \hat{J}_z \mathbf{k}$ , and similar for  $\hat{\mathbf{J}}_1$  and  $\hat{\mathbf{J}}_2$ , the cartesian components of angular momenta are related by

$$\begin{aligned} \hat{J}_x &= \hat{J}_{1x} + \hat{J}_{2x} \\ \hat{J}_y &= \hat{J}_{1y} + \hat{J}_{2y} \\ \hat{J}_z &= \hat{J}_{1z} + \hat{J}_{2z}. \end{aligned} \quad \left. \right\} \quad (7.4)$$

In particular

$$\begin{aligned} \langle j_1, j_2; m_1, m_2 | \hat{J}_z | J, M \rangle &= \langle j_1, j_2; m_1, m_2 | \hat{J}_{1z} + \hat{J}_{2z} | J, M \rangle, \\ \therefore M \langle j_1, j_2; m_1, m_2 | J, M \rangle &= (m_1 + m_2) \langle j_1, j_2; m_1, m_2 | J, M \rangle. \end{aligned}$$

Thus

$$\left[ M - (m_1 + m_2) \right] \langle j_1, j_2; m_1, m_2 | J, M \rangle = 0$$

and so

**Restriction on  $m$  values** (conservation of “ $J_z$ ”)

The Clebsch-Gordan coefficient  $\langle j_1, j_2; m_1, m_2 | J, M \rangle$  is zero unless

$$M = m_1 + m_2 . \quad (7.5)$$

In Section **6 Angular Momentum Operators** we discussed the action of the raising and lowering operators  $\hat{L}_+$  and  $\hat{L}_-$ . The same expressions apply here for the angular momenta  $\hat{\mathbf{J}}$ ,  $\hat{\mathbf{J}}_1$  and  $\hat{\mathbf{J}}_2$ . That is, we define

$$\begin{aligned}\hat{J}_\pm &= \hat{J}_x \pm i\hat{J}_y \\ \hat{J}_{1\pm} &= \hat{J}_{1x} \pm i\hat{J}_{1y} \\ \hat{J}_{2\pm} &= \hat{J}_{2x} \pm i\hat{J}_{2y}\end{aligned}$$

and so from Eq. (7.4) we find

insert  
2

$$\hat{J}_\pm = \hat{J}_{1\pm} + \hat{J}_{2\pm} . \quad (7.6)$$

Also

$$\left. \begin{aligned}\hat{J}_\pm |J, M\rangle &= \hbar \sqrt{(J \pm M + 1)(J \mp M)} |J, M \pm 1\rangle \\ \hat{J}_{1\pm} |j_1, j_2; m_1, m_2\rangle &= \hbar \sqrt{(j_1 \pm m_1 + 1)(j_1 \mp m_1)} |j_1, j_2; m_1 \pm 1, m_2\rangle \\ \hat{J}_{2\pm} |j_1, j_2; m_1, m_2\rangle &= \hbar \sqrt{(j_2 \pm m_2 + 1)(j_2 \mp m_2)} |j_1, j_2; m_1, m_2 \pm 1\rangle .\end{aligned}\right\} (7.7)$$

Consider the inner product  $\langle j_1, j_2; m_1, m_2 | \hat{J}_\pm | J, M \rangle$ . From Eq. (7.6) we get

$$\begin{aligned}\langle j_1, j_2; m_1, m_2 | \hat{J}_\pm | J, M \rangle &= \langle j_1, j_2; m_1, m_2 | \hat{J}_{1\pm} + \hat{J}_{2\pm} | J, M \rangle \\ &= \langle j_1, j_2; m_1, m_2 | \hat{J}_{1\pm} | J, M \rangle + \langle j_1, j_2; m_1, m_2 | \hat{J}_{2\pm} | J, M \rangle .\end{aligned}$$

Making use of Eq. (7.7) we get

**Recurrence relation for Clebsch-Gordan coefficients**

$$\begin{aligned}&\sqrt{(J \pm M + 1)(J \mp M)} \langle j_1, j_2; m_1, m_2 | J, M \pm 1 \rangle \\ &= \sqrt{(j_1 \mp m_1 + 1)(j_1 \pm m_1)} \langle j_1, j_2; m_1 \mp 1, m_2 | J, M \rangle \\ &\quad + \sqrt{(j_2 \mp m_2 + 1)(j_2 \pm m_2)} \langle j_1, j_2; m_1, m_2 \mp 1 | J, M \rangle .\end{aligned}\quad (7.8)$$

This relation allows us to determine the Clebsch-Gordan coeffi-

cients.

### NOTE

- The Clebsch-Gordan coefficient is zero if any  $m$  value is outside the allowed range, e.g. if  $M + 1$  is greater than  $J$ , or if  $m_1 - 1$  is less than  $-j_1$ , etc.

- In deriving Eq. (7.8) we have used the fact that  $(\hat{J}_{1\pm})^\dagger = \hat{J}_{1\mp}$  and so

$$\begin{aligned}
 & \langle j_1, j_2; m_1, m_2 | \hat{J}_{1\pm} \\
 &= \left[ (\hat{J}_{1\pm})^\dagger | j_1, j_2; m_1, m_2 \rangle \right]^\dagger \\
 &= \left( \hat{J}_{1\mp} | j_1, j_2; m_1, m_2 \rangle \right)^\dagger \\
 &= \left( \hbar \sqrt{(j_1 \mp m_1 + 1)(j_1 \pm m_1)} | j_1, j_2; m_1 \mp 1, m_2 \rangle \right)^\dagger \\
 &= \hbar \sqrt{(j_1 \mp m_1 + 1)(j_1 \pm m_1)} \langle j_1, j_2; m_1 \mp 1, m_2 | .
 \end{aligned}$$

and similarly for  $\hat{J}_{2\pm}$ .

### • Example (1).

Consider a system of two spin- $\frac{1}{2}$  particles. The values of  $j_1$  and  $j_2$  are both  $\frac{1}{2}$ . The range of the  $J$  values are 0 and 1 (i.e. from  $|j_1 - j_2|$  to  $j_1 + j_2$  in integer steps). From Eq. (7.3) we have

$$|J, M\rangle = \sum_{m_1, m_2} \langle \frac{1}{2}, \frac{1}{2}; m_1, m_2 | J, M \rangle | \frac{1}{2}, \frac{1}{2}; m_1, m_2 \rangle . \quad (7.9)$$

Thus we want the coefficients

$$\langle \frac{1}{2}, \frac{1}{2}; m_1, m_2 | 1, M \rangle$$

for  $m_1 = \pm \frac{1}{2}$ ,  $m_2 = \pm \frac{1}{2}$ ,  $M = -1, 0, 1$ , and

$$\langle \frac{1}{2}, \frac{1}{2}; m_1, m_2 | 0, 0 \rangle$$

for  $m_1 = \pm \frac{1}{2}$ ,  $m_2 = \pm \frac{1}{2}$ . Determine the Clebsch-Gordan coefficients for the expansion with  $J = 0$  and  $M = 0$ , i.e. the coefficients in

$$|0, 0\rangle = \sum_{m_1, m_2} \langle j_1, j_2; m_1, m_2 | 0, 0 \rangle | j_1, j_2; m_1, m_2 \rangle .$$

### Solution

From Eq. (7.5) the coefficient  $\langle \frac{1}{2}, \frac{1}{2}; m_1, m_2 | 0, 0 \rangle$  is zero unless  $m_1 + m_2 = 0$  for which there are only two possibilities: either  $m_1 = \frac{1}{2}$  and  $m_2 = -\frac{1}{2}$ , or  $m_1 = -\frac{1}{2}$  and  $m_2 = \frac{1}{2}$ . These are the

only non-zero coefficient in the expansion of  $|0, 0\rangle$  in Eq. (7.9), and so

$$\begin{aligned} |0, 0\rangle &= \langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | 0, 0 \rangle | \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \rangle \\ &\quad + \langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} | 0, 0 \rangle | \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \rangle . \end{aligned} \quad (7.10)$$

Taking the upper sign of the recurrence relation with  $j_1 = j_2 = \frac{1}{2}$ ,  $m_1 = m_2 = \frac{1}{2}$  and  $J = M = 0$  gives

$$0 = \langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} | 0, 0 \rangle + \langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | 0, 0 \rangle$$

and so

$$\langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} | 0, 0 \rangle = -\langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | 0, 0 \rangle . \quad (7.11)$$

### Sign convention

We adopt the sign convention of Racah where the coefficient  $\langle j_1, j_2; j_1, J - j_1 | J, J \rangle$  is real and positive.

Thus, taking into account Eq. (7.10), Eq. (7.11), the sign convention and normalisation, we find

$$|0, 0\rangle = \frac{1}{\sqrt{2}} \left( | \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \rangle - | \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \rangle \right) .$$

• • •

### Problem

Continue working through Example (1) and calculate the coefficients for the expansion of  $|1, 1\rangle$ ,  $|1, 0\rangle$  and  $|1, -1\rangle$ . You should find

$$\begin{aligned} \langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2} | 1, 1 \rangle &= 1 \\ \langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} | 1, -1 \rangle &= 1 \\ \sqrt{2} \langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | 1, 0 \rangle &= \langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2} | 1, 1 \rangle \\ \sqrt{2} \langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} | 1, 0 \rangle &= \langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} | 1, -1 \rangle \end{aligned}$$

with all other coefficients being zero.

[Hint: Consider the restriction on the  $m$  values given by Eq. (7.5).]

Thus

$$\begin{aligned}
 |1, -1\rangle &= |\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2}\rangle \\
 |1, 0\rangle &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\rangle + |\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2}\rangle \right) \\
 |1, 1\rangle &= |\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2}\rangle
 \end{aligned}$$

These results can be summarised in tabular form as

**Clebsch-Gordan coefficients for  $j_1 = \frac{1}{2}, j_2 = \frac{1}{2}$ .**

		J	1	1	0	1
		M	1	0	0	-1
$m_1$	$m_2$					
$\frac{1}{2}$	$\frac{1}{2}$		1	0	0	0
$\frac{1}{2}$	$-\frac{1}{2}$		0	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	0
$-\frac{1}{2}$	$\frac{1}{2}$		0	$\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	0
$-\frac{1}{2}$	$-\frac{1}{2}$		0	0	0	1

### Notes

- Rewriting these results using the notation of topics **1 Introduction** and **5 Entanglement** yields the following:

$$\begin{aligned}
 |1, 1\rangle_z &= |\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z \\
 |1, 0\rangle_z &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z + |-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z \right) \\
 |1, -1\rangle_z &= |-\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z \\
 |0, 0\rangle_z &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}\rangle_z \otimes |-\frac{1}{2}\rangle_z - |-\frac{1}{2}\rangle_z \otimes |\frac{1}{2}\rangle_z \right).
 \end{aligned}$$

- Notice that there are **two** different ways to get  $M = 0$  for the combined system, one for  $J = 1$  ( $|1, 0\rangle_z$ ) and the other for  $J = 0$  ( $|0, 0\rangle_z$ ).
- Notice also that the above 4 states are mutually orthogonal (including  $|1, 0\rangle_z$  and  $|0, 0\rangle_z$ ).

The Clebsch-Gordan coefficient for the coupling of orbital angular momentum with  $l = 1$  and the spin of  $\frac{1}{2}$  (such as in an atom) is given by

### Clebsch-Gordan coefficients for $j_1 = 1, j_2 = \frac{1}{2}$

$J$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{3}{2}$
$M$	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$
$m_1$	$m_2$					
1	$\frac{1}{2}$	1	0	0	0	0
1	$-\frac{1}{2}$	0	$\frac{1}{\sqrt{3}}$	$\sqrt{\frac{2}{3}}$	0	0
0	$\frac{1}{2}$	0	$\sqrt{\frac{2}{3}}$	$-\frac{1}{\sqrt{3}}$	0	0
0	$-\frac{1}{2}$	0	0	0	$\sqrt{\frac{2}{3}}$	$\frac{1}{\sqrt{3}}$
-1	$\frac{1}{2}$	0	0	0	$\frac{1}{\sqrt{3}}$	$-\sqrt{\frac{2}{3}}$
-1	$-\frac{1}{2}$	0	0	0	0	1

For example

$$|\frac{3}{2}, \frac{1}{2}\rangle = \frac{1}{\sqrt{3}}|1, \frac{1}{2}; 1, -\frac{1}{2}\rangle + \sqrt{\frac{2}{3}}|1, \frac{1}{2}; 0, \frac{1}{2}\rangle .$$

More details and general formula can be found in, for example, A. R. Edmonds, *Angular momentum in quantum mechanics* (Princeton University Press, Princeton, N.J, 1974) or M. E. Rose *Elementary theory of angular momentum* (Wiley, New York, 1957).

# Problem Sheet

- (1)** Continue working through Example (1) on page 5 and calculate the coefficients for the expansion of  $|1, 1\rangle$ ,  $|1, 0\rangle$  and  $|1, -1\rangle$ . You should find

$$\begin{aligned}\langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2} | 1, 1 \rangle &= 1 \\ \langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} | 1, -1 \rangle &= 1 \\ \sqrt{2} \langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | 1, 0 \rangle &= \langle \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2} | 1, 1 \rangle \\ \sqrt{2} \langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} | 1, 0 \rangle &= \langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} | 1, -1 \rangle\end{aligned}$$

with all other coefficients being zero.

[**HINT:** Consider the restriction on the  $m$  values given by Eq. (7.5).]

- (2)** Show that

$$\hat{\mathbf{J}}^2 = \hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + 2\hat{J}_{1x}\hat{J}_{2x} + 2\hat{J}_{1y}\hat{J}_{2y} + 2\hat{J}_{1z}\hat{J}_{2z}. \quad (7.12)$$

- (3)** Evaluate the commutator  $[\hat{\mathbf{J}}^2, \hat{J}_{1z}]$ .

- (4)** Use the results of the Section **1 Review** to show

$$\left. \begin{aligned}\hat{J}_{1x}\hat{J}_{2x}|\frac{1}{2}, \frac{1}{2}; m_1, m_2\rangle &= \frac{1}{4}\hbar^2 |\frac{1}{2}, \frac{1}{2}; -m_1, -m_2\rangle \\ \hat{J}_{1y}\hat{J}_{2y}|\frac{1}{2}, \frac{1}{2}; m_1, m_2\rangle &= -m_1 m_2 \hbar^2 |\frac{1}{2}, \frac{1}{2}; -m_1, -m_2\rangle \\ \hat{J}_{1z}\hat{J}_{2z}|\frac{1}{2}, \frac{1}{2}; m_1, m_2\rangle &= m_1 m_2 \hbar^2 |\frac{1}{2}, \frac{1}{2}; m_1, m_2\rangle.\end{aligned}\right\} (7.13)$$

[Hint: Here the  $\hat{J}_{1x}$ ,  $\hat{J}_{1y}$ , and  $\hat{J}_{1z}$  operators are the equivalent of the spin- $\frac{1}{2}$  operators  $\hat{S}_x$ ,  $\hat{S}_y$ ,  $\hat{S}_z$  from Section **1 Review**. Thus their action on the first spin particle is given by

$$\begin{aligned}\hat{J}_{1x} &= \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{xx}\langle\frac{1}{2}| + \frac{1}{2}\hbar|\frac{1}{2}\rangle_{xx}\langle-\frac{1}{2}|, \\ \hat{J}_{1y} &= \frac{i}{2}\hbar|-\frac{1}{2}\rangle_{yy}\langle\frac{1}{2}| - \frac{i}{2}\hbar|\frac{1}{2}\rangle_{yy}\langle-\frac{1}{2}|, \\ \hat{J}_{1z} &= \frac{1}{2}\hbar|\frac{1}{2}\rangle_{zz}\langle\frac{1}{2}| - \frac{1}{2}\hbar|-\frac{1}{2}\rangle_{zz}\langle-\frac{1}{2}|.\end{aligned}$$

A similar relation holds for the second spin particle.]

- (5)** Using the expansion Eq. (7.12) and the results Eq. (7.13), show that the eigenvalue of  $\hat{\mathbf{J}}^2$  for the states on the right-hand sides is the  $J$  value on the left-hand sides of the fol-

lowing equations:

$$\begin{aligned} |1, -1\rangle &= |\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2}\rangle \\ |1, 0\rangle &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\rangle + |\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2}\rangle \right) \\ |1, 1\rangle &= |\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2}\rangle \\ |0, 0\rangle &= \frac{1}{\sqrt{2}} \left( |\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\rangle - |\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2}\rangle \right) . \end{aligned}$$

# Advanced Quantum Theory

## 7 Addition of angular momentum

### Solutions to Problems

- (1)** The only way to get  $M = m_1 + m_2$  for  $M = 1$  is  $m_1 = m_2 = \frac{1}{2}$  and so the expansion of  $|1, 1\rangle$  only contains one coefficient  $\langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} | 1, -1 \rangle$  which, by normalisation, is unity. A similar situation exists for  $|1, -1\rangle$ .

To find the coefficients for  $|1, 0\rangle$  set  $J = 1$ ,  $M = 1$ ,  $m_1 = \frac{1}{2}$  and  $m_2 = -\frac{1}{2}$  in Eq. (7.8) with the lower sign, and then set  $J = 1$ ,  $M = -1$ ,  $m_1 = -\frac{1}{2}$  and  $m_2 = \frac{1}{2}$  in Eq. (7.8) with the upper sign.

- (2)** Using the definition  $\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2$  and using the dot product defition of the squared vector one obtains the equality

$$\begin{aligned}
 \hat{\mathbf{J}}^2 &= (\hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2)^2 \\
 &= (\hat{J}_{1x} + \hat{J}_{2x})^2 + (\hat{J}_{1y} + \hat{J}_{2y})^2 + (\hat{J}_{1z} + \hat{J}_{2z})^2 \\
 &= \hat{J}_{1x}^2 + \hat{J}_{1y}^2 + \hat{J}_{1z}^2 + \hat{J}_{2x}^2 + \hat{J}_{2y}^2 + \hat{J}_{2z}^2 + \\
 &\quad 2\hat{J}_{1x}\hat{J}_{2x} + 2\hat{J}_{1y}\hat{J}_{2y} + 2\hat{J}_{1z}\hat{J}_{2z} \\
 &= \hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + 2\hat{J}_{1x}\hat{J}_{2x} + 2\hat{J}_{1y}\hat{J}_{2y} + 2\hat{J}_{1z}\hat{J}_{2z}.
 \end{aligned}$$

It should be noted that the two terms in the squares are from two separate systems and hence these operators commute and can be treated like scalars for this particular expansion.

- (3)** It is known that  $\hat{J}_{1z}$  commutes will all operators that act exclusively on system 2, with  $\hat{\mathbf{J}}_1^2$  and itself. Therefore when substituting the expression from the previous problem and removing terms that evaluate to zero

$$\begin{aligned}
 [\hat{\mathbf{J}}^2, \hat{J}_{1z}] &= [2\hat{J}_{1x}\hat{J}_{2x}, \hat{J}_{1z}] + [2\hat{J}_{1y}\hat{J}_{2y}, \hat{J}_{1z}] \\
 &= 2[\hat{J}_{1x}, \hat{J}_{1z}]\hat{J}_{2x} + 2[\hat{J}_{1y}, \hat{J}_{1z}]\hat{J}_{2y} \\
 &= -2i\hbar(\hat{J}_{1y}\hat{J}_{2x} - \hat{J}_{1x}\hat{J}_{2y}) \\
 &= 2i\hbar(\hat{\mathbf{J}}_1 \times \hat{\mathbf{J}}_2)_z
 \end{aligned}$$

- (4)** The operators apply separately to each system and scalar values can be transported through operators as they are

linear. Applying these one at a time

$$\begin{aligned}
 \hat{J}_1 x \hat{J}_2 x \left| \frac{1}{2}, \frac{1}{2}; m_1, m_2 \right\rangle &= \frac{1}{2} \hbar \hat{J}_1 x \left| \frac{1}{2}, \frac{1}{2}; m_1, -m_2 \right\rangle \\
 &= \frac{1}{4} \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -m_1, -m_2 \right\rangle \\
 \hat{J}_1 y \hat{J}_2 y \left| \frac{1}{2}, \frac{1}{2}; m_1, m_2 \right\rangle &= i m_2 \hbar \hat{J}_1 y \left| \frac{1}{2}, \frac{1}{2}; m_1, -m_2 \right\rangle \\
 &= -m_1 m_2 \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -m_1, -m_2 \right\rangle \\
 \hat{J}_1 z \hat{J}_2 z \left| \frac{1}{2}, \frac{1}{2}; m_1, m_2 \right\rangle &= m_2 \hbar \hat{J}_1 z \left| \frac{1}{2}, \frac{1}{2}; m_1, m_2 \right\rangle \\
 &= m_1 m_2 \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; m_1, m_2 \right\rangle
 \end{aligned}$$

- (5) This is verified by taking each equality separately and computing the left and right hand side when applying the appropriate version of  $\hat{\mathbf{J}}^2$ .

For the first equation, applying  $\hat{\mathbf{J}}^2$  to the left hand side gives

$$\hat{\mathbf{J}}^2 |1, -1\rangle = 2\hbar^2 |1, -1\rangle$$

and the right hand side

$$\begin{aligned}
 \hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + 2\hat{J}_1 x \hat{J}_2 x + 2\hat{J}_1 y \hat{J}_2 y + 2\hat{J}_1 z \hat{J}_2 z &\left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} \right\rangle \\
 &= \frac{3}{4} \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} \right\rangle + \frac{3}{4} \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} \right\rangle \\
 &\quad + \frac{2}{4} \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \right\rangle - \frac{2}{4} \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \right\rangle + \frac{2}{4} \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} \right\rangle \\
 &= (\frac{3}{4} + \frac{3}{4} + \frac{2}{4}) \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} \right\rangle \\
 &= 2\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2} \right\rangle .
 \end{aligned}$$

So the eigenvalues are the same for both sides.

For the second equation, applying  $\hat{\mathbf{J}}^2$  to the left hand side gives

$$\hat{\mathbf{J}}^2 |1, 0\rangle = 2\hbar^2 |1, 0\rangle$$

and the right hand side (ignoring normalization as it does not affect the eigenvalue equation)

$$\begin{aligned}
& \hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + 2\hat{J}_1 x \hat{J}_2 x + 2\hat{J}_1 y \hat{J}_2 y + 2\hat{J}_1 z \hat{J}_2 z (\left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle + \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \right\rangle) \\
&= \frac{3}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle + \frac{3}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle + \frac{2}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \right\rangle \\
&\quad + \frac{2}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \right\rangle - \frac{2}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle + \frac{3}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \right\rangle \\
&\quad + \frac{3}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \right\rangle + \frac{2}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle + \frac{2}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle \\
&\quad - \frac{2}{4}\hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \right\rangle \\
&= \left( \frac{3}{4} + \frac{3}{4} - \frac{2}{4} + \frac{2}{4} + \frac{2}{4} \right) \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle \\
&\quad + \left( \frac{2}{4} + \frac{2}{4} + \frac{3}{4} + \frac{3}{4} - \frac{2}{4} \right) \hbar^2 \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \right\rangle \\
&= 2\hbar^2 \left( \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle + \left| \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2} \right\rangle \right).
\end{aligned}$$

So the eigenvalues are the same for both sides.

The other two equations follow in a similar way as above.

# 8. Irreducible tensor operators

## a. Wigner's 3-*j* symbols

It is often convenient to represent Clebsch-Gordan coefficients in terms of numbers known as **3-j symbols** which were introduced by Eugene Wigner. Each 3-*j* symbol is a single number that is a function of six variables, just as the Clebsch-Gordan coefficient is. It has certain symmetry properties that are best revealed by representing it as a  $2 \times 3$  array defined as follows:

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} \equiv \frac{(-1)^{j_1-j_2+M}}{\sqrt{2J+1}} \langle j_1, j_2; m_1, m_2 | J, M \rangle . \quad (8.1)$$

Here,  $j_1$  and  $j_2$  are the quantum numbers for the angular momenta of the two subsystems,  $J$  is the quantum number for the total angular momentum, and  $m_1$ ,  $m_2$  and  $M$  are the quantum numbers for the z components of, respectively, the first and second subsystems, and the combined system. The inner product  $\langle j_1, j_2; m_1, m_2 | J, M \rangle$  is a Clebsch-Gordan coefficient.

### 3-*j* symbol

It is important to remember that though the 3-*j* symbol is a  $2 \times 3$  array, each 3-*j* symbol is just a **single number**. You should think of the 3-*j* symbol as a function  $f(\cdot)$  as follows:

$$'' \begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} '' = f(j_1, j_2, m_1, m_2, J, M) \in \mathbb{C} .$$

where  $\mathbb{C}$  is the set of complex numbers.

Using Eq. (8.1) we can express Clebsch-Gordan coefficients in terms of 3-*j* symbols as follows:

$$\langle j_1, j_2; m_1, m_2 | J, M \rangle = (-1)^{j_2-j_1-M} \sqrt{2J+1} \begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} .$$

3-*j* symbols can be more convenient to use than Clebsch-Gordon coefficients due to the fact that they possess a number of symmetry-related properties. Specifically,

- 3-*j* symbols are invariant under **even permutations** of its columns.
- **Odd permutations** of the columns of the 3-*j* symbols result in multiplication by  $(-1)^{j_1+j_2+J}$ .
- Negating all  $m_1$ ,  $m_2$  and  $-M$  values multiplies a 3-*j* symbol by  $(-1)^{j_1+j_2+J}$ .

Here an ‘even permutation’ is an even number of column interchanges and an ‘odd permutation’ is an odd number of column interchanges, where a ‘column interchange’ means the swapping of two columns.

As with Clebsch-Gordan coefficients, 3- $j$  symbols are typically found via consulting tables or running a computer programme. An example of a program can be found at

<http://www-stone.ch.cam.ac.uk/wigner.html>.

• **Example (1).**

Evaluate the Clebsch-Gordon coefficient  $\langle 3, 1; 2, 0 | 2, 2 \rangle$  using the 3- $j$  symbols.

**Solution**

We find that

insert  
1

$$\begin{pmatrix} 3 & 1 & 2 \\ 2 & 0 & -2 \end{pmatrix} = -\frac{1}{\sqrt{21}}.$$

Using the relation

$$\langle 3, 1; 2, 0 | 2, 2 \rangle = (-1)^{1-3-2} \sqrt{2 \times 2 + 1} \begin{pmatrix} 3 & 1 & 2 \\ 2 & 0 & -2 \end{pmatrix}.$$

it follows that

$$\langle 3, 1; 2, 0 | 2, 2 \rangle = 1 \times \sqrt{5} \times \frac{-1}{\sqrt{21}} = -\sqrt{\frac{5}{21}}.$$

• • •

## b. Irreducible Tensor Operators

A number of operators we encounter in quantum mechanics are elements of a mathematical object called an **irreducible tensor operator**, or sometimes just an **irreducible tensor**.

Mathematically, an irreducible tensor operator is any set  $T_k$  containing  $2k+1$  elements  $T_k^{(q)}$  each satisfy the following commutation relations:

$$\left. \begin{aligned} [\hat{J}_z, T_k^{(q)}] &= q\hbar T_k^{(q)} \\ [\hat{J}_+, T_k^{(q)}] &= \sqrt{(k+q+1)(k-q)}\hbar T_k^{(q+1)} \\ [\hat{J}_-, T_k^{(q)}] &= \sqrt{(k-q+1)(k+q)}\hbar T_k^{(q-1)} \end{aligned} \right\} \quad (8.2)$$

where  $\hat{J}_z$ ,  $\hat{J}_+$  and  $\hat{J}_-$  are generic angular momentum operators. That is, they could represent  $\hat{L}_z$ ,  $\hat{L}_+$  and  $\hat{L}_-$  or, alternately,  $\hat{S}_z$ ,  $\hat{S}_+$  and  $\hat{S}_-$ . They might also represent total angular momentum operators.

**Problem**

Check that the set  $T_1 = \{T_1^{(-1)}, T_1^{(0)}, T_1^{(1)}\}$  with  $T_1^{(-1)} = \frac{1}{\sqrt{2}}\hat{J}_-$ ,  $T_1^{(0)} = \hat{J}_z$  and  $T_1^{(1)} = \frac{1}{\sqrt{2}}\hat{J}_+$  is an irreducible tensor operator by substitution into Eq. (8.2).

The **Wigner-Eckart theorem** allows us to calculate matrix elements governing transition rates between quantum-mechanical states:

**Wigner-Eckart theorem**

$$\langle J', M' | T_k^{(q)} | J, M \rangle = \frac{(-1)^{k-J+J'}}{\sqrt{(2J'+1)}} \langle J' | T_k | J \rangle \langle k, J; q, M | J', M' \rangle \quad (8.3)$$

where

$$\begin{aligned} & \langle J' | T_k | J \rangle \\ &= \sum_{q'} \sum_{\mu} \sum_{\mu'} (2J'+1)^{-1} \langle J, k; \mu, q' | J', \mu' \rangle \langle J, k; J', \mu' | T_k^{(q')} | J, \mu \rangle . \end{aligned}$$

The quantity  $\langle J | T_k | J \rangle$  is known as a **reduced matrix element**. Its **crucial property is that it is independent of  $M$  and  $M'$** .

The Wigner-Eckart theorem can also be expressed in terms of 3-j symbols as follows:

$$\langle J', M' | T_k^{(q)} | J, M \rangle = (-1)^{J'-M'} \langle J' | T_k | J \rangle \left( \begin{array}{ccc} k & J & J' \\ q & M & -M' \end{array} \right) .$$

**KEY POINT**

The theorem's utility lies in the fact that it replaces the more complex matrix element  $\langle J', M' | T_k^{(q)} | J, M \rangle$  by  $\langle J' | T_k | J \rangle$ , **which has no  $M$  or  $M'$  dependence**, and an easily-calculated Clebsch-Gordan coefficient. This means we can easily calculate the **relative strengths** of transitions between states of differing  $M$  values. The **absolute strength** of the transitions depend on the value of  $\langle J' | T_k | J \rangle$ . This is obtained from involved numerical calculations, theoretical approximations or, more often, directly from experiment.

• **Example (2).**

The probability of an electric dipole in an atom spontaneously emitting a photon of polarisation  $\mathbf{r}$  and making a transition from the state  $|J, M\rangle$  to the state  $|J', M'\rangle$  is

insert  
2

$$P_r = \frac{\omega^3}{2\pi\hbar c^3} |\langle J', M' | \mathbf{r} \cdot \mathbf{D} | J, M \rangle|^2,$$

where  $\mathbf{r}$  is the polarisation vector of the electric field and  $\mathbf{D}$  is the electric dipole moment operator of the atom. Let  $r^{(q)}$  and  $D_1^{(q)}$  be the spherical components of  $\mathbf{r}$  and  $\mathbf{D}$ , respectively, in the basis given by

$$\mathbf{e}^{(-1)} = \frac{\mathbf{x} - i\mathbf{y}}{\sqrt{2}}, \quad \mathbf{e}^{(0)} = \mathbf{z}, \quad \mathbf{e}^{(1)} = \frac{\mathbf{x} + i\mathbf{y}}{\sqrt{2}}$$

for orthogonal unit vectors  $\mathbf{x}$ ,  $\mathbf{y}$  and  $\mathbf{z}$ . That is,  $\mathbf{r} = r^{(-1)}\mathbf{e}^{(-1)} + r^{(0)}\mathbf{e}^{(0)} + r^{(1)}\mathbf{e}^{(1)}$  and  $\mathbf{D} = D_1^{(-1)}\mathbf{e}^{(-1)} + D_1^{(0)}\mathbf{e}^{(0)} + D_1^{(1)}\mathbf{e}^{(1)}$ . For linear polarised light  $\mathbf{r} = \mathbf{e}^{(0)}$  and for left and right circularly polarised light  $\mathbf{r} = \mathbf{e}^{(\pm 1)}$ . From this it follows that

$$P_r = \frac{\omega^3}{2\pi\hbar c^3} \left| \sum_q r^{(-q)} \langle J', M' | D_1^{(q)} | J, M \rangle \right|^2. \quad (8.4)$$

Using the Wigner-Eckart theorem Eq. (8.3),

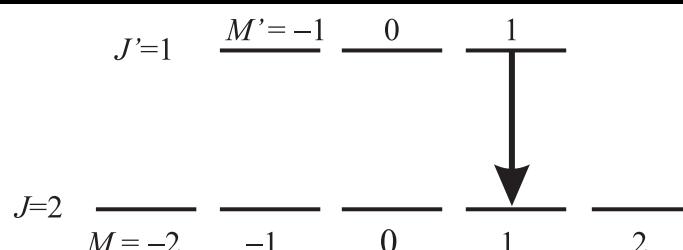
$$\langle J', M' | D_1^{(q)} | J, M \rangle = (-1)^{J'-M'} \langle J' \| D_1 \| J \rangle \begin{pmatrix} 1 & J & J' \\ q & M & -M' \end{pmatrix}. \quad (8.5)$$

As  $\langle J' \| D_1 \| J \rangle$  is independent of  $M$  and  $M'$  we are free to set them to any value we choose. To make the calculation the simplest, we choose  $M = M' = 0$ , and also consider  $q = 0$  (i.e. linear polarised light along  $\mathbf{z}$ ). Hence, from Eq. (8.5)

$$\langle J' \| D_1 \| J \rangle = \frac{\langle J', 0 | D_1^{(0)} | J, 0 \rangle}{(-1)^{J'-0} \times \begin{pmatrix} 1 & J & J' \\ 0 & 0 & 0 \end{pmatrix}}. \quad (8.6)$$

$\langle J', 0 | D_1^{(0)} | J, 0 \rangle$  can be obtained from the absolute strength of the  $M' = 0$  to  $M = 0$  transition (e.g. from spectroscopy). This allows  $\langle J' \| D_1 \| J \rangle$  to be evaluated from Eq. (8.6), and its value is then substituted into Eq. (8.5) to determine  $\langle J', M' | D_1^{(q)} | J, M \rangle$  for all desired  $M'$ ,  $M$  and  $q$  values.

• • •



**Fig 8.1:** Level structure for the  $J' = 2$  and  $J = 1$  states. The transition is from  $M' = 1$  to  $M = 1$ .

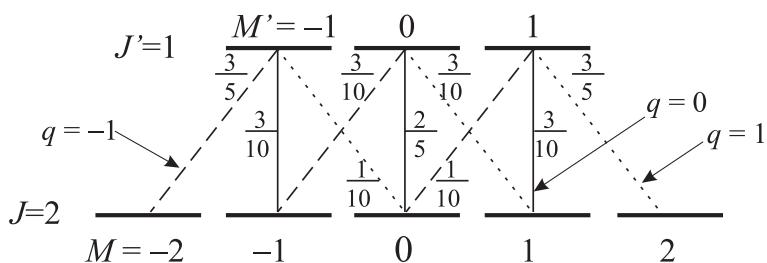
**Problem**

The value  $\langle 2, 0 | D_1^{(0)} | 1, 0 \rangle = \frac{1}{4}$  has been determined for Example (2). Given this, what is the probability of a transition occurring from the  $J' = 2, M' = 1$  state to the  $J = 1, M = 1$  state with polarisations  $\mathbf{r} = e^{(0)}\mathbf{z}$  (i.e.  $q = 0$ ) and, separately,  $\mathbf{r} = e^{(1)}(\mathbf{x} + i\mathbf{y})$  (i.e.  $q = 1$ ).

It is interesting to calculate the transition probability for the three types of polarisation,  $\mathbf{r} = e^{(q)}$  for  $q = -1, 0, 1$ . By substituting Eq. (8.5) into the expression for the transition probability Eq. (8.4) we eventually find

$$\begin{aligned} P_{e^{(q)}} &= \frac{\omega^3}{2\pi\hbar c^3} \left| \langle J' | D_1 | J \rangle \right|^2 \left| \begin{pmatrix} 1 & J & J' \\ -q & M & -M' \end{pmatrix} \right|^2 \\ &= C \left| \begin{pmatrix} 1 & J & J' \\ -q & M & -M' \end{pmatrix} \right|^2 = C' \left| \langle 1, J; -q, M | J', M' \rangle \right|^2 . \end{aligned}$$

where  $C$  and  $C' = (2J+1)C$  are constants independent of  $M$  and  $M'$ . The **relative** transition probability depends on the square of the modulus of a 3-j symbol, or equivalently of a Clebsch-Gordan coefficient, which is easily evaluated. The relative strengths of the  $J' = 1$  to  $J = 2$  transition are illustrated in the figure below.



**Fig 8.2:** Relative strengths of the  $J' = 2$  to  $J = 1$  transitions.

The polarisation is right circular ( $q = -1$ , dotted), linear ( $q = 0$ , solid) or left circular ( $q = 1$ , dashed).

## Problem Sheet

- (1) Check that the set  $T_1 = \{T_1^{(-1)}, T_1^{(0)}, T_1^{(1)}\}$  with  $T_1^{(-1)} = \frac{1}{\sqrt{2}}\hat{J}_-$ ,  $T_1^{(0)} = \hat{J}_z$  and  $T_1^{(1)} = \frac{1}{\sqrt{2}}\hat{J}_+$  is an irreducible tensor operator by substitution into Eq. (8.2).
- (2) The value  $\langle 2, 0 | D_1^{(0)} | 1, 0 \rangle = \frac{1}{4}$  has been determined for Example (2) on page 3. Given this, what is the probability of a transition occurring from the  $J' = 2, M' = 1$  state to the  $J = 1, M = 1$  state with polarisations  $\mathbf{r} = e^{(0)}\mathbf{z}$  (i.e.  $q = 0$ ) and, separately,  $\mathbf{r} = e^{(1)}(\mathbf{x} + i\mathbf{y})$  (i.e.  $q = 1$ ).

# Advanced Quantum Theory

## 8 Irreducible tensor operators

### Solutions to Problems

- (1)** Noting the following commutators from the definitions given in previous sections

$$\begin{aligned}
 \left[ \hat{J}_z, \frac{1}{\sqrt{2}} \hat{J}_- \right] &= -\hbar \frac{1}{\sqrt{2}} \hat{J}_- \\
 \left[ \hat{J}_z, \hat{J}_z \right] &= 0 \\
 \left[ \hat{J}_z, \frac{1}{\sqrt{2}} \hat{J}_+ \right] &= \hbar \frac{1}{\sqrt{2}} \hat{J}_+ \\
 \left[ \hat{J}_+, \frac{1}{\sqrt{2}} \hat{J}_- \right] &= \sqrt{2} \hbar \hat{J}_z \\
 \left[ \hat{J}_+, \hat{J}_z \right] &= -\sqrt{2} \hbar \frac{1}{\sqrt{2}} \hat{J}_+ \\
 \left[ \hat{J}_+, \frac{1}{\sqrt{2}} \hat{J}_+ \right] &= 0 \\
 \left[ \hat{J}_-, \frac{1}{\sqrt{2}} \hat{J}_- \right] &= 0 \\
 \left[ \hat{J}_-, \hat{J}_z \right] &= -\sqrt{2} \hbar \frac{1}{\sqrt{2}} \hat{J}_- \\
 \left[ \hat{J}_-, \frac{1}{\sqrt{2}} \hat{J}_+ \right] &= -\sqrt{2} \hbar \hat{J}_z
 \end{aligned}$$

it follows that the definitions hold in this particular scenario.

- (2)** From Eq. (8.5) it is possible to show that

$$\begin{aligned}
 \frac{\langle J' = 2, M' = 1 | D_1^{(0)} | J = 1, M = 1 \rangle}{\langle J' = 2, M' = 0 | D_1^{(0)} | J = 1, M = 0 \rangle} &= - \frac{\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \end{pmatrix}}{\begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}} \\
 &= \frac{\sqrt{1/10}}{\sqrt{2/15}} \\
 &= \sqrt{\frac{3}{4}},
 \end{aligned}$$

and also

$$\begin{aligned}
 \frac{\langle J' = 2, M' = 1 | D_1^{(-1)} | J = 1, M = 1 \rangle}{\langle J' = 2, M' = 0 | D_1^{(0)} | J = 1, M = 0 \rangle} &= - \frac{\begin{pmatrix} 1 & 1 & 2 \\ -1 & 1 & -1 \end{pmatrix}}{\begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}} \\
 &= \frac{0}{\sqrt{2/15}} \\
 &= 0.
 \end{aligned}$$

Therefore by substitution into Eq. (8.4) the probability of the  $J' = 2, M' = 1 \rightarrow J = 1, M = 1$  transition with linearly polarised light (i.e.  $q = 0$ ) is

$$\frac{\omega^3}{2\pi\hbar c^3} \frac{3}{64} \quad (8.7)$$

and the probability of the same transition with right circularly polarised light (i.e.  $q = 1$ ) is 0.

# 9. Time Evolution

## a. The time evolution operator

The Schrödinger equation for pure states is

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}_t |\psi(t)\rangle . \quad (9.1)$$

The behaviour of the system over time is governed by the details of the Hamiltonian  $\hat{H}_t$ . We begin with the simplest situation where the **Hamiltonian is independent of time**.

### Time-independent Hamiltonians

The Hamiltonian is an observable, and thus it is a self-adjoint (or Hermitian) operator. Hence its eigenstates form an orthonormal basis for the system's Hilbert space. Let the diagonal form of  $\hat{H}$  be:

$$\hat{H} = \sum_n E_n |E_n\rangle\langle E_n| \quad (9.2)$$

where  $E_n$  are the energy eigenvalues and  $\{|E_n\rangle\}$  is the set of orthonormal eigenstates. The expansion of the state of the system,  $|\psi(t)\rangle$ , in the energy eigenbasis is

$$|\psi(t)\rangle = \sum_n c_n(t) |E_n\rangle . \quad (9.3)$$

The time dependence of the state  $|\psi(t)\rangle$  is completely contained in the coefficients  $c_n(t)$  because the eigenstates  $\{|E_n\rangle\}$  are fixed in time. Substituting Eq. (9.2) into Eq. (9.1) and then using the expansion Eq. (9.3) yields

$$\begin{aligned} i\hbar \frac{d}{dt} \left( \sum_n c_n(t) |E_n\rangle \right) &= \left( \sum_n E_n |E_n\rangle\langle E_n| \right) \left( \sum_m c_m(t) |E_m\rangle \right) \\ \therefore i\hbar \sum_n \left( \frac{d}{dt} c_n(t) \right) |E_n\rangle &= \sum_n \sum_m E_n c_m(t) |E_n\rangle \langle E_n| E_m \rangle \\ &= \sum_n E_n c_n(t) |E_n\rangle . \end{aligned}$$

Taking the inner product with  $\langle E_m|$  gives a differential equation for the coefficients

$$i\hbar \frac{d}{dt} c_m(t) = E_m c_m(t)$$

whose solution is

$$c_m(t) = \exp[-\frac{i}{\hbar} E_m t] c_m(0) . \quad (9.4)$$

Thus the solution to Eq. (9.1) is thus

$$|\psi(t)\rangle = \sum_n \exp[-\frac{i}{\hbar}E_n t] c_n(0) |E_n\rangle .$$

The orthonormality of the basis  $\{|E_n\rangle\}$  allows us to write this as

$$|\psi(t)\rangle = \left( \sum_n \exp[-\frac{i}{\hbar}E_n t] |E_n\rangle \langle E_n| \right) \left( \sum_m c_m(0) |E_m\rangle \right) \quad (9.5)$$

$$= \exp \left( \sum_n -\frac{i}{\hbar}E_n t |E_n\rangle \langle E_n| \right) \left( \sum_m c_m(0) |E_m\rangle \right) \quad (9.6)$$

$$= \exp[-\frac{i}{\hbar}\hat{H}t] |\psi(0)\rangle \quad (9.7)$$

$$= \hat{U}(t, 0) |\psi(0)\rangle \quad (9.8)$$

where

### time evolution operator for time-independent $\hat{H}$

$$\hat{U}(t, t_0) \equiv \exp[-\frac{i}{\hbar}\hat{H} \times (t - t_0)] \quad (9.9)$$

### KEY POINT

A function of an operator, like  $\exp[-\frac{i}{\hbar}\hat{H} \times (t - t_0)]$ , is defined in terms of the eigenbasis of the operator. For example, consider the self-adjoint operator  $\hat{A}$  with orthonormal eigenbasis  $\{|\lambda_i\rangle\}$ , i.e.  $\hat{A} = \sum_i \lambda_i |\lambda_i\rangle \langle \lambda_i|$ . The general function  $f(\hat{A})$  is defined to be

$$f(\hat{A}) \equiv \sum_i f(\lambda_i) |\lambda_i\rangle \langle \lambda_i| .$$

This is a generalisation of powers of  $\hat{A}$  which we discussed in the **Introduction**, viz.  $\hat{A}^n = \sum_i \lambda_i^n |\lambda_i\rangle \langle \lambda_i|$ . The exponential function of  $\hat{H}$  is used in going from Eq. (9.5) to Eq. (9.7).

### Problem

Show that  $\hat{U}(t, t_0)$  in Eq. (9.9) is a unitary operator (i.e. that  $\hat{U}^\dagger(t, t_0) \hat{U}(t, t_0) = \hat{U}(t, t_0) \hat{U}^\dagger(t, t_0) = \mathbb{1}$ ).

## KEY POINT

The time evolution operator allows us to calculate the solution of the Schrödinger equation for all times. The general solution is usually written as

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle \quad (9.10)$$

where  $\hat{U}(t, t_0)$  takes the state from time  $t_0$  to time  $t$ . The time evolution operator factorises into a product of evolution operators for contiguous periods as

$$\begin{aligned} |\psi(t)\rangle &= \hat{U}(t, t_1)|\psi(t_1)\rangle = \hat{U}(t, t_1)\left(\hat{U}(t_1, t_0)|\psi(t_0)\rangle\right) \\ &= \left(\hat{U}(t, t_1)\hat{U}(t_1, t_0)\right)|\psi(t_0)\rangle. \end{aligned} \quad (9.11)$$

Comparing Eqs. (9.10) and (9.11) and noting that these equations hold for all states  $|\psi(t_0)\rangle$  we find

$$\hat{U}(t, t_0) = \hat{U}(t, t_1)\hat{U}(t_1, t_0).$$

For a time-**in**dependent Hamiltonian, it is simple to verify that this equation holds because  $\hat{U}(t, t_0)$  in Eq. (9.9) depends on just the time difference  $t - t_0$ . However, this is not the case for time-dependent Hamiltonians, which is why we have included the explicit beginning and ending times,  $t_0$  and  $t$ , of the evolution in the symbol for  $\hat{U}(t, t_0)$  for generality.

## Time-dependent Hamiltonians

We now consider the time-dependent case, where the Hamiltonian  $\hat{H}_t$  depends on time. In this case we Eqs. (9.1) and (9.10) still hold. Substituting the latter into the former implies

$$i\hbar \frac{d}{dt} \hat{U}(t, t_0)|\psi(t_0)\rangle = \hat{H}_t \hat{U}(t, t_0)|\psi(t_0)\rangle. \quad (9.12)$$

But  $|\psi(t_0)\rangle$  is an arbitrary vector, so we have the *operator differential equation*

$$i\hbar \frac{d}{dt} \hat{U}(t, t_0) = \hat{H}_t \hat{U}(t, t_0), \quad (9.13)$$

with the initial condition

$$\hat{U}(t_0, t_0) = \mathbf{1}. \quad (9.14)$$

It might be thought that this differential equation has the solution

$$\hat{U}(t, t_0) = \exp \left[ -\frac{i}{\hbar} \int_{t_0}^t \hat{H}_t dt \right]. \quad (?) \quad (9.15)$$

Indeed, this reduces to the above solution (9.9) for time-independent Hamiltonians. Equation (9.15) is also the correct solution for time-dependent Hamiltonians where

$$[\hat{H}_{t_a}, \hat{H}_{t_b}] = 0, \forall t_a, t_b \in [t_0, t]. \quad (9.16)$$

However, if this is not satisfied then Eq. (9.15) is incorrect. The

correct solution is

**time evolution operator for  $\hat{H}_t$**

$$\begin{aligned}\hat{U}(t, t_0) = & \mathbb{1} - \frac{i}{\hbar} \int_{t_1=t_0}^t dt_1 \hat{H}_{t_1} \\ & - \frac{1}{\hbar^2} \int_{t_1=t_0}^t dt_1 \hat{H}_{t_1} \int_{t_2=t_0}^{t_1} dt_2 \hat{H}_{t_2} \\ & + \left(\frac{-i}{\hbar}\right)^3 \int_{t_1=t_0}^t dt_1 \hat{H}_{t_1} \int_{t_2=t_0}^{t_1} dt_2 \hat{H}_{t_2} \int_{t_3=t_0}^{t_2} dt_3 \hat{H}_{t_3} \\ & + \dots .\end{aligned}\tag{9.17}$$

Notice that the ordering of later Hamiltonians to the left of earlier ones (e.g.  $\hat{H}_{t_1}$  is to the left of  $\hat{H}_{t_2}$ , and  $t_1 > t_2$ ).

**Problem**

Show that, for a time-independent Hamiltonian, Eq. (9.17) reduces to Eq. (9.9).

We can prove that  $\hat{U}(t, t_0)$  obeys Eq. (9.13) by substitution. Remembering the

**FUNDAMENTAL THEOREM OF CALCULUS**

$$\frac{d}{dt} \int_{s=a}^t ds f(s) \equiv f(t),\tag{9.18}$$

and that this holds even when  $f(s) = \int_{r=b}^s dr g(r)$  etc., we obtain:

insert

1

$$\begin{aligned}
 i\hbar \frac{d}{dt} \hat{U}(t, t_0) &= 0 + \hat{H}_t \\
 &\quad - \frac{i}{\hbar} \hat{H}_t \int_{t_2=t_0}^t dt_2 \hat{H}_{t_2} \\
 &\quad + \left( \frac{-i}{\hbar} \right)^2 \hat{H}_t \int_{t_2=t_0}^t dt_2 \hat{H}_{t_2} \int_{t_3=t_0}^{t_2} dt_3 \hat{H}_{t_3} \\
 &\quad + \dots . \tag{9.19}
 \end{aligned}$$

Relabeling  $t_2$  as  $t_1$  and  $t_3$  as  $t_2$  etc, we have:

$$\begin{aligned}
 i\hbar \frac{d}{dt} \hat{U}(t, t_0) &= \hat{H}_t \{ \mathbf{1} \\
 &\quad - \frac{i}{\hbar} \int_{t_1=t_0}^t dt_1 \hat{H}_{t_1} \\
 &\quad - \frac{1}{\hbar^2} \int_{t_1=t_0}^t dt_1 \hat{H}_{t_1} \int_{t_2=t_0}^{t_1} dt_2 \hat{H}_{t_2} \\
 &\quad + \dots \} \\
 &= \hat{H}_t \hat{U}(t, t_0) \tag{9.20}
 \end{aligned}$$

The result in Eq. (9.17) can be more compactly expressed using **Dyson time-ordering** which is often denoted by the symbol  $\mathcal{T}$ . This operation orders operators according to their time arguments such that

$$\mathcal{T}[\hat{A}(t_1)\hat{B}(t_2)] = \begin{cases} \hat{A}(t_1)\hat{B}(t_2), & \text{if } t_1 \geq t_2 \\ \hat{B}(t_2)\hat{A}(t_1), & \text{if } t_2 > t_1 \end{cases}$$

That is, it puts operators with greater time arguments to the left of those with lesser time argument. Using  $\mathcal{T}$  we can, after some algebraic manipulation, express the right-hand side of Eq. (9.17) as

$$\hat{U}(t, t_0) = \mathcal{T} \left\{ e^{-\frac{i}{\hbar} \int_{t'=t_0}^t \hat{H}_{t'} dt'} \right\} .$$

## b. “Pictures” of time evolution

Up to now we have only considered the time evolution in terms of changes to the **state** of a system. This “view” of the dynamics is called the **Schrödinger picture**. But this is not the only way to study the dynamics. In classical physics it is usually the observables, such as the position of a particle  $x$ , that change in time. This possibility also exists in the quantum case: the operators can change in time while the states remain fixed. This “view” is called the **Heisenberg picture**. Because of Heisenberg’s uncertainty principle, quantum operators that change in time are analogous to classical random variables that change in time. The classical analogue of the Schrödinger picture is an equation for a *probability distribution*  $P(x; t)$  that describes this randomness.

The two pictures can be summarised as follows:

Picture	What Evolves	What stays constant
Schrödinger	states	operators
Heisenberg	operators	states

### KEY POINT

Some books use a subscript, such as  $S$  or  $H$ , to denote a state or operator in the Schrödinger or Heisenberg picture. However, this notation quickly becomes cumbersome. Instead we will adopt the following convention. Let the time of the start of the Hamiltonian evolution be  $t_0$ . Then the time dependence of the state or operator automatically denotes if the object is in the Schrödinger or Heisenberg picture.

objects	picture	reason
$ \psi(t)\rangle$ , $\hat{A}(t_0)$	Schrödinger	the state is evolving
$ \psi(t_0)\rangle$ , $\hat{A}(t)$	Heisenberg	the operator is evolving

## c. Schrödinger picture

In the Schrödinger picture a pure state is represented by a time-dependent **state vector**, which evolves as

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle$$

where  $t_0$  and  $t$  are two arbitrary times and  $\hat{U}$  is the time-evolution operator for the system of interest. The evolution of mixed states is described by the time-dependent density operator or **state ma-**

**trix**, which evolves as

$$\begin{aligned}
\hat{\rho}(t) &= \sum_i p_i |\psi_i(t)\rangle\langle\psi_i(t)| \\
&= \sum_i p_i \hat{U}(t, t_0) |\psi_i(t_0)\rangle\langle\psi_i(t_0)| \hat{U}^\dagger(t, t_0) \\
&= \hat{U}(t, t_0) \left[ \sum_i p_i |\psi_i(t_0)\rangle\langle\psi_i(t_0)| \right] \hat{U}^\dagger(t, t_0) \\
&= \hat{U}(t, t_0) \hat{\rho}(t_0) \hat{U}^\dagger(t, t_0) .
\end{aligned}$$

Operators representing **observables** do not evolve in time under this transformation. Thus an arbitrary operator at time  $t$  will just be  $\hat{A}(t_0)$ . The time  $t_0$  here represents the beginning time of the Hamiltonian evolution. Its presence is not important for operators in the Schrödinger picture. Indeed, it is typically omitted if the analysis is done completely in the Schrödinger picture.

The expectation value of an arbitrary operator at time  $t$  is given by

$$\begin{aligned}
\langle \hat{A}(t) \rangle &= \langle \psi(t) | \hat{A}(t_0) | \psi(t) \rangle \\
&= \langle \psi(t_0) | \hat{U}^\dagger(t, t_0) \hat{A}(t_0) \hat{U}(t, t_0) | \psi(t_0) \rangle
\end{aligned} \quad (9.21)$$

for an initial pure state  $|\psi(t_0)\rangle$ , and

$$\begin{aligned}
\langle \hat{A}(t) \rangle &= \text{Tr} \left[ \hat{\rho}(t) \hat{A}(t_0) \right] \\
&= \text{Tr} \left[ \hat{U}(t, t_0) \hat{\rho}(t_0) \hat{U}^\dagger(t, t_0) \hat{A}(t_0) \right]
\end{aligned} \quad (9.22)$$

for the initial state matrix  $\hat{\rho}(t_0)$ .

### Schrödinger picture quantities

$\langle \hat{A}(t) \rangle = \langle \psi(t)   \hat{A}(t_0)   \psi(t) \rangle$ where $ \psi(t)\rangle = \hat{U}(t, t_0)  \psi(t_0)\rangle$ , (pure state) or $\langle \hat{A}(t) \rangle = \text{Tr} \left[ \hat{\rho}(t) \hat{A}(t_0) \right]$ where $\hat{\rho}(t) = \hat{U}(t, t_0) \hat{\rho}(t_0) \hat{U}^\dagger(t, t_0)$ . (mixed state)
---

## d. Heisenberg Picture

In the Heisenberg picture the states remain constant and the operators evolve. We can associate the time-evolution operator  $\hat{U}(t, t_0)$  in Eqs. (9.21) and (9.22) with the operator instead of

the states, that is, from Eq. (9.21)

$$\begin{aligned}\langle \hat{A}(t) \rangle &= \left\langle \psi(t_0) \left| \left[ \hat{U}^\dagger(t, t_0) \hat{A}(t_0) \hat{U}(t, t_0) \right] \right| \psi(t_0) \right\rangle \\ &= \langle \psi(t_0) | \hat{A}(t) | \psi(t_0) \rangle\end{aligned}$$

for pure states and, from Eq. (9.22) using the cyclic property of the trace,

$$\begin{aligned}\langle \hat{A}(t) \rangle &= \text{Tr} \left[ \hat{U}(t, t_0) \hat{\rho}(t_0) \hat{U}^\dagger(t, t_0) \hat{A}(t_0) \right] \\ &= \text{Tr} \left\{ \hat{\rho}(t_0) \left[ \hat{U}^\dagger(t, t_0) \hat{A}(t_0) \hat{U}(t, t_0) \right] \right\} \\ &= \text{Tr} \left[ \hat{\rho}(t_0) \hat{A}(t) \right]\end{aligned}$$

for mixed states, where we have defined the Heisenberg operator

$$\hat{A}(t) = \hat{U}^\dagger(t, t_0) \hat{A}(t_0) \hat{U}(t, t_0).$$

### Heisenberg picture quantities

or	$\langle \hat{A}(t) \rangle = \langle \psi(t_0)   \hat{A}(t)   \psi(t_0) \rangle,$	(pure state)
	$\langle \hat{A}(t) \rangle = \text{Tr} \left[ \hat{\rho}(t_0) \hat{A}(t) \right],$	(mixed state)
where	$\hat{A}(t) = \hat{U}^\dagger(t, t_0) \hat{A}(t_0) \hat{U}(t, t_0).$	

### Heisenberg equation of motion

The **Heisenberg equation of motion** is an equation that evolves operators in the Heisenberg picture, just the Schrödinger equation evolves states in the Schrödinger picture. We begin with the equation

$$\hat{A}(t) = \hat{U}^\dagger(t, t_0) \hat{A}(t_0) \hat{U}(t, t_0) \quad (9.23)$$

and differentiate both sides with respect to time  $t$  using the product rule on the right-hand side to yield

$$\frac{d\hat{A}(t)}{dt} = \frac{d\hat{U}^\dagger}{dt} \hat{A} \hat{U} + \hat{U}^\dagger \hat{A} \frac{d\hat{U}}{dt}. \quad (9.24)$$

where, for brevity, we have written  $\hat{U}(t, t_0)$  as  $\hat{U}$  and  $\hat{A}(t_0)$  as  $\hat{A}$ . Recall from Eq. (9.13) that

$$\frac{d}{dt} \hat{U}(t, t_0) = (i\hbar)^{-1} \hat{H}_t \hat{U}(t, t_0),$$

Taking the Hermitian adjoint of this equation yields

$$\frac{d\hat{U}^\dagger}{dt} = (-i\hbar)^{-1} \hat{U}^\dagger \hat{H}_t$$

where we have used the fact that  $(\hat{R}\hat{S})^\dagger = \hat{S}^\dagger \hat{R}^\dagger$  for any two

operators  $\hat{R}$  and  $\hat{S}$  and  $\hat{H}^\dagger = \hat{H}$ . Note that  $\hat{H}_t$  is not a Heisenberg-picture operator; it just may have some explicit time-dependence.

On substituting these expressions for  $d\hat{U}^\dagger/dt$  and  $d\hat{U}/dt$  into Eq. (9.24) we obtain

$$\frac{d\hat{A}(t)}{dt} = \frac{i}{\hbar} \hat{U}^\dagger \hat{H}_t \hat{A} \hat{U} + -\frac{i}{\hbar} \hat{U}^\dagger \hat{A} \hat{H}_t \hat{U}. \quad (9.25)$$

Now introducing  $\mathbf{1} = \hat{U}^\dagger \hat{U}$  between  $\hat{H}_t$  and  $\hat{A}$ , we get

$$\frac{d\hat{A}(t)}{dt} = \frac{i}{\hbar} \hat{U}^\dagger \hat{H}_t \hat{U} \hat{U}^\dagger \hat{A} \hat{U} - \frac{i}{\hbar} \hat{U}^\dagger \hat{A} \hat{U} \hat{U}^\dagger \hat{H}_t \hat{U}. \quad (9.26)$$

Further defining a Heisenberg-picture Hamiltonian operator, just as in Eq. (9.23),

$$\hat{H}(t) = \hat{U}^\dagger \hat{H}_t \hat{U}, \quad (9.27)$$

we have

$$\begin{aligned} \frac{d\hat{A}(t)}{dt} &= \frac{i}{\hbar} \hat{H}(t) \hat{A}(t) - \frac{i}{\hbar} \hat{A}(t) \hat{H}(t) \\ &= -\frac{i}{\hbar} [\hat{A}(t), \hat{H}(t)] \end{aligned} \quad (9.28)$$

To look more like Schrödinger's equation Eq. (9.1), this is often written as

### Heisenberg's equation of motion

insert  
2

$$i\hbar \frac{d\hat{A}(t)}{dt} = [\hat{A}(t), \hat{H}(t)].$$

### Problem

Use the same technique to show that the equation of motion for  $\rho(t)$  in the **Schrödinger picture** is

$$i\hbar \frac{d\rho(t)}{dt} = [\hat{H}_t, \rho(t)] \quad (9.29)$$

**Note the different order in the commutator** — this is **not** a Heisenberg equation of motion for the operator  $\rho(t)$ . Although  $\rho(t)$  is an operator, it represents the state, not an observable.

### Hamiltonians with no explicit time-dependence

If the Hamiltonian has no explicit time-dependence, so that  $\hat{H}_t = \hat{H}$ , then from Eq. (9.9),  $\hat{U}$  commutes with  $\hat{H}$  (it is just a function of  $\hat{H}$ ). Thus

$$\hat{H}(t) = \hat{U}^\dagger \hat{H} \hat{U} = \hat{H} \hat{U}^\dagger \hat{U} = \hat{H}. \quad (9.30)$$

That is, the Hamiltonian in the Heisenberg picture is constant.

For example, the Hamiltonian for the harmonic oscillator in the

Schrödinger picture is  $\hat{H}_{\text{HO}} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)$  which is clearly time independent.  $\hat{H}_{\text{HO}}$  is expressed in Heisenberg picture form as

$$\begin{aligned}\hat{H}_{\text{HO}}(t) &= U^\dagger \hat{H}_{\text{HO}} \hat{U} = \hbar\omega \hat{U}^\dagger \left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \hat{U} \\ &= \hbar\omega \left(U^\dagger \hat{a}^\dagger \hat{U} U^\dagger \hat{a} \hat{U} + \frac{1}{2}\right) \\ &= \hbar\omega \left[\hat{a}^\dagger(t)\hat{a}(t) + \frac{1}{2}\right].\end{aligned}$$

Despite  $\hat{a}(t)$  and  $\hat{a}^\dagger(t)$  being time dependent,  $\hat{H}_{\text{HO}}(t)$  is actually time **in**-dependent according to Eq. (9.30). Evidently  $\hat{a}(t)$  and  $\hat{a}^\dagger(t)$  co-evolve in such a way as to make the product  $\hat{a}^\dagger(t)\hat{a}(t)$  time independent.

Because of Eq. (9.30), it is common to write the Hamiltonian in the Heisenberg picture as simply  $\hat{H}$  and then replace  $\hat{H}$  with  $\hat{H}(t)$  when performing calculations such as commutators. This is demonstrated in Example below.

• **Example (1).** (Sakurai, Ch. 2, problem 11)

Consider a particle subject to a one-dimensional simple harmonic oscillator potential. Suppose at  $t = 0$  the state vector is given by

$$|\psi(0)\rangle = \exp[-i\hat{P}(0)\frac{L}{\hbar}]|0\rangle \quad (9.31)$$

where  $|0\rangle$  is the ground state (i.e. the energy eigenstate  $|n\rangle$  with  $n = 0$ ),  $\hat{P}(0)$  is the momentum operator at  $t = 0$  and  $L$  is some number with the dimensions of length. Using the Heisenberg picture, evaluate the expectation value of the position  $\langle \hat{X}(t) \rangle$  for  $t \geq 0$ .

### Solution

We first solve for the time evolution of the Heisenberg operator  $\hat{X}(t)$ . We know that

$$\hat{X} = \frac{\sqrt{\hbar}}{\sqrt{2m\omega}} \left(\hat{a}^\dagger + \hat{a}\right), \quad (9.32)$$

$$\hat{P} = i\frac{\sqrt{\hbar m\omega}}{\sqrt{2}} \left(\hat{a}^\dagger - \hat{a}\right),$$

$$\hat{a} = \frac{\sqrt{m\omega}}{\sqrt{2\hbar}} \left(\hat{X} + \frac{i}{m\omega}\hat{P}\right), \quad (9.33)$$

$$\hat{a}^\dagger = \frac{\sqrt{m\omega}}{\sqrt{2\hbar}} \left(\hat{X} - \frac{i}{m\omega}\hat{P}\right), \quad (9.34)$$

where  $m$  is the mass of the particle,  $\omega = \sqrt{k/m}$  is the oscillation frequency,  $k$  is the spring constant and  $\hat{a}$ ,  $\hat{a}^\dagger$  are the annihilation

and creation operators with  $[\hat{a}, \hat{a}^\dagger] = 1$ . These relations are definitions and hold for all times.

• • •

**Problem**

Show that the commutator between any two operators is preserved in the Heisenberg picture, e.g. if

$$[\hat{A}(t_0), \hat{B}(t_0)] = \hat{C}(t_0)$$

in the Schrödinger picture, then

$$[\hat{A}(t), \hat{B}(t)] = \hat{C}(t)$$

in the Heisenberg picture where  $\hat{A}(t) = \hat{U}^\dagger(t, t_0)\hat{A}(t_0)\hat{U}(t, t_0)$   
etc.

**Step 1.** The first step is to express  $\hat{X}(t)$  in terms of operators at  $t = 0$ . We begin by using the Heisenberg equation of motion for  $\hat{a}(t)$ :

$$\begin{aligned} i\hbar \frac{d\hat{a}(t)}{dt} &= [\hat{a}(t), \hat{H}] \\ &= \hbar\omega[\hat{a}(t), \hat{a}^\dagger(t)\hat{a}(t) + \frac{1}{2}], \end{aligned}$$

where we have written the Hamiltonian in the second expression in terms of Heisenberg picture operators explicitly. Thus Eq. (9.35) becomes

$$i\hbar \frac{d\hat{a}(t)}{dt} = \hbar\omega\hat{a}(t). \quad (9.35)$$

Solving Eq. (9.35) as a standard differential equation produces

$$\hat{a}(t) = \exp(-i\omega t)\hat{a}(0). \quad (9.36)$$

Taking the Hermitian conjugate of both sides of this equation, we arrive at an expression for  $\hat{a}^\dagger(t)$ , namely,

$$\hat{a}^\dagger(t) = \exp(i\omega t)\hat{a}^\dagger(0). \quad (9.37)$$

Substituting the expressions for  $\hat{a}(t)$  and  $\hat{a}^\dagger(t)$  in Eqs. (9.36) and (9.37) into the right-hand side of Eq. (9.32) produces

$$\begin{aligned} \hat{X}(t) &= \frac{\sqrt{\hbar}}{\sqrt{2m\omega}} \left( \hat{a}^\dagger(t) + \hat{a}(t) \right) \\ &= \frac{\sqrt{\hbar}}{\sqrt{2m\omega}} \left( e^{i\omega t}\hat{a}^\dagger(0) + e^{-i\omega t}\hat{a}(0) \right). \end{aligned} \quad (9.38)$$

Substituting expressions for  $\hat{a}(0)$  and  $\hat{a}^\dagger(0)$  in terms of  $\hat{X}(0)$  and

$\hat{P}(0)$  from Eqs. (9.33) and (9.34), viz.

$$\begin{aligned}\hat{a}(0) &= \frac{\sqrt{m\omega}}{\sqrt{2\hbar}} \left( \hat{X}(0) + \frac{i}{m\omega} \hat{P}(0) \right), \\ \hat{a}^\dagger(0) &= \frac{\sqrt{m\omega}}{\sqrt{2\hbar}} \left( \hat{X}(0) - \frac{i}{m\omega} \hat{P}(0) \right),\end{aligned}$$

into Eq. (9.38) gives

$$\hat{X}(t) = \cos(\omega t) \hat{X}(0) + \frac{1}{m\omega} \sin(\omega t) \hat{P}(0). \quad (9.39)$$

Thus we have expressed  $\hat{X}(t)$  in terms of operators at  $t = 0$ .

**Step 2.** Next we combine this result with the initial state  $|\psi(0)\rangle$  to evaluate  $\langle \hat{X}(t) \rangle$ :

$$\begin{aligned}\langle \hat{X}(t) \rangle &= \langle \psi(0) | \hat{X}(t) | \psi(0) \rangle \\ &= \langle 0 | \exp[i\hat{P}(0)\frac{L}{\hbar}] \left[ \cos(\omega t) \hat{X}(0) + \frac{1}{m\omega} \sin(\omega t) \hat{P}(0) \right] \exp[-i\hat{P}(0)\frac{L}{\hbar}] | 0 \rangle \\ &= \cos(\omega t) \langle 0 | \exp[i\hat{P}(0)\frac{L}{\hbar}] \hat{X}(0) \exp[-i\hat{P}(0)\frac{L}{\hbar}] | 0 \rangle \\ &\quad + \frac{1}{m\omega} \sin(\omega t) \langle 0 | \hat{P}(0) | 0 \rangle.\end{aligned} \quad (9.40)$$

The average value of momentum for the ground state is zero and so

$$\langle 0 | \hat{P}(0) | 0 \rangle = 0. \quad (9.41)$$

We now need to evaluate

$$\langle 0 | \exp[i\hat{P}(0)\frac{L}{\hbar}] \hat{X}(0) \exp[-i\hat{P}(0)\frac{L}{\hbar}] | 0 \rangle.$$

### Problem

Derive the operator identity

$$e^{\hat{A}} \hat{B} e^{-\hat{A}} = \hat{B} + [\hat{A}, \hat{B}] + \frac{1}{2!} [\hat{A}, [\hat{A}, \hat{B}]] + \frac{1}{3!} [\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \dots$$

(Hint: Expand the exponential functions in terms of the power series  $e^{\hat{A}} = \mathbf{1} + \hat{A} + \frac{1}{2!} \hat{A}^2 + \frac{1}{3!} \hat{A}^3 + \dots$  and collect terms starting with the lowest order. Be careful to preserve the operator ordering as  $\hat{A}$  and  $\hat{B}$  do not commute in general.)

Using the operator identity in the Problem and  $[\hat{X}, \hat{P}] = i\hbar$  we find

$$\exp(i\hat{P}\frac{L}{\hbar})\hat{X}\exp(-i\hat{P}\frac{L}{\hbar}) \quad (9.42)$$

insert  
3

$$\begin{aligned} &= \hat{X} + [i\hat{P}\frac{L}{\hbar}, \hat{X}] + \frac{1}{2}[i\hat{P}\frac{L}{\hbar}, [i\hat{P}\frac{L}{\hbar}, \hat{X}]] + \dots \\ &= \hat{X} + i\frac{L}{\hbar}[\hat{P}, \hat{X}] - \frac{1}{2}\frac{L^2}{\hbar^2}[\hat{P}, [\hat{P}, \hat{X}]] + \dots \\ &= \hat{X} + i\frac{L}{\hbar}(-i\hbar) - \frac{1}{2}\frac{L^2}{\hbar^2}[\hat{P}, -i\hbar] + \dots \\ &= \hat{X} + L + 0 . \end{aligned} \quad (9.43)$$

Thus

$$\begin{aligned} \left\langle 0 | \exp[i\hat{P}(0)\frac{L}{\hbar}] \hat{X}(0) \exp[-i\hat{P}(0)\frac{L}{\hbar}] | 0 \right\rangle &= \left\langle 0 | (\hat{X}(0) + L) | 0 \right\rangle \\ &= L \end{aligned} \quad (9.44)$$

as average value of position for the ground state is zero and so  $\langle 0 | \hat{X}(0) | 0 \rangle = 0$ . Substituting Eqs. (9.41) and (9.44) into Eq. (9.40) gives the final result:

$$\langle \hat{X}(t) \rangle = L \cos(\omega t) .$$

This sinusoidal behaviour is not unexpected for a harmonic oscillator! A more interesting exercise would be to calculate the time dependance of the variance in the position which is given by  $\Delta^2 X = \langle \hat{X}^2 \rangle - \langle \hat{X} \rangle^2$ .

### Problem

Use the solution  $\hat{X}(t)$  in the Heisenberg picture, Eq. (9.39), to calculate the time dependence of the variance  $\Delta^2 X$ .

[Hints: Use Eq. (9.39) to get an expression for  $\hat{X}^2(t)$ . For the expectation value calculation  $\langle \hat{X}^2 \rangle$  use the fact that  $e^{i\hat{P}c}$  is unitary and so  $e^{i\hat{P}c}\hat{X}^2e^{-i\hat{P}c} = (e^{i\hat{P}c}\hat{X}e^{-i\hat{P}c})(e^{i\hat{P}c}\hat{X}e^{-i\hat{P}c})$  where  $c = L/\hbar$ . Then use Eq. (9.43).]

**Problem**

The Hamiltonian for a single spin-1/2 particle precessing about the  $z$ -axis under the influence of an external magnetic field is

$$\hat{H} = -\left(\frac{eB}{mc}\right)\hat{S}_z.$$

Write down the Heisenberg equations of motion for  $\hat{S}_x$ ,  $\hat{S}_y$  and  $\hat{S}_z$ , the spin operators introduced in chapter 2. Solve them to find  $\langle \hat{S}_x(t) \rangle$ ,  $\langle \hat{S}_y(t) \rangle$  and  $\langle \hat{S}_z(t) \rangle$  given the initial state  $|\psi(0)\rangle = \alpha|\frac{1}{2}\rangle_z + \beta|-\frac{1}{2}\rangle_z$ .

[Hint: For the two coupled differential equations  $\dot{x} = -ay$  and  $\dot{y} = ax$  has the following solution:  $x(t) = \cos(at)x(0) - \sin(at)y(0)$  and  $y(t) = \sin(at)x(0) + \cos(at)y(0)$ .]

## e. Interaction Frame

In classical mechanics it is often useful to transform to a new reference frame to make the dynamics easier. For example, one might transform to the rotating frame of the Earth to study the dynamics of motion on its surface; the transformation removes large rotating terms of little interest at the expense of additional “fictitious” terms (such as the Coriolis “force” that gives rise to cyclones). In quantum mechanics the same thing can be done. It is most commonly done for Hamiltonians describing two interacting systems, where the “**free**” **Hamiltonian** describing the non-interacting systems give simple evolution (like the simple rotation of the earth) but the interaction between them gives more complicated. By moving to a frame where the free Hamiltonian is removed, one can concentrate on the more interesting interaction. This is called working in the **interaction frame**. \*

Consider the Hamiltonian for a system undergoing an interaction with an external device, or two systems interacting with each other. For example, the system could be an atom and the external device could be a laser field, or the two systems could be an atom and a laser cavity field. In these situations the Hamiltonian can be written as

$$\hat{H}_t = \hat{H}_0 + \hat{V}_t \quad (9.45)$$

where  $\hat{H}_0$  is the Hamiltonian of the system (or systems) without the interaction and  $\hat{V}_t$  represents the interaction, which often has an explicit time-dependence. Usually (as shown here)  $\hat{H}_0$  has no explicit time-dependence, and gives rise to relatively simple

\*In text books, one more often encounters talk of working in an interaction picture. This is related, but not the same, as working in the interaction frame. As we will see, an interaction frame can contain either a Heisenberg or a Schrödinger picture. Some text-books mistakenly use the phrase interaction picture when they mean interaction frame.

evolution via

$$\hat{U}_0(t, t_0) = \exp[-i\hat{H}_0 \times (t - t_0)/\hbar]. \quad (9.46)$$

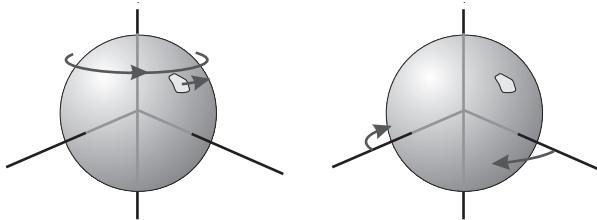
The full evolution via the Hamiltonian  $\hat{H}_t$  is given of course by Eq. (9.17), the solution to

$$i\hbar \frac{d}{dt} \hat{U}(t, t_0) = \hat{H}(t) \hat{U}(t, t_0) \quad (9.47)$$

The idea of the interaction frame is to remove the simple evolution  $\hat{U}_0(t, t_0)$  from the full evolution  $\hat{U}(t, t_0)$ . This is done very simply by “undoing” the free evolution at the end, to get the unitary evolution operator in the interaction frame  $I$ :

$$\hat{U}_I(t, t_0) = \hat{U}_0^{-1}(t, t_0) \hat{U}(t, t_0) \quad (9.48)$$

The inverse operator  $\hat{U}_0^{-1}(t, t_0) = \hat{U}_0^\dagger(t, t_0)$  for the free evolution is like transforming to the rotating frame of the Earth as illustrated in the figure. We take our full solution and then rotate it “backwards” to get the simpler evolution where (without the interaction) things would stand still on the earth.



**Fig 9.1:** Rotating frame. In order to make a point on the surface of the Earth stationary, we rotate the axes in the reverse direction.

What we want now is an **interaction frame Hamiltonian** that will generate  $\hat{U}_I(t, t_0)$ . We will show that the required Hamiltonian is

$$\hat{V}_I(t) = \hat{U}_0^\dagger(t, t_0) V_t \hat{U}_0(t, t_0). \quad (9.49)$$

That is, we will show that

$$i\hbar \frac{d}{dt} \hat{U}_I(t, t_0) = \hat{V}_I(t) \hat{U}(t, t_0), \quad (9.50)$$

so that the solution for  $\hat{U}_I(t, t_0)$  is just Eq. (9.17) with  $\hat{H}_t$  replaced by  $\hat{V}_I(t)$ .

## Proof:

insert  
4

$$\begin{aligned}
 & i\hbar \frac{d}{dt} \hat{U}_I(t, t_0) \\
 = & i\hbar \frac{d}{dt} \left\{ \hat{U}_0^\dagger(t, t_0) \hat{U}(t, t_0) \right\} \\
 = & \left\{ i\hbar \frac{d}{dt} \hat{U}_0^\dagger(t, t_0) \right\} \hat{U}(t, t_0) + \hat{U}_0^\dagger(t, t_0) \left\{ i\hbar \frac{d}{dt} \hat{U}(t, t_0) \right\} \\
 = & -\hat{U}_0^\dagger(t, t_0) \hat{H}_0 \hat{U}(t, t_0) + \hat{U}_0^\dagger(t, t_0) (\hat{H}_0 + V_t) \hat{U}(t, t_0) \quad (9.51)
 \end{aligned}$$

Here we have used the equation

$$i\hbar \frac{d}{dt} \hat{U}_0^\dagger(t, t_0) = -\hat{U}_0^\dagger(t, t_0) \hat{H}_0,$$

obvious from Eq. (9.46). Thus

$$\begin{aligned}
 i\hbar \frac{d}{dt} \hat{U}_I(t, t_0) &= \hat{U}_0^\dagger(t, t_0) V_t \hat{U}(t, t_0) \\
 &= \hat{U}_0^\dagger(t, t_0) V_t \hat{U}_0(t, t_0) \hat{U}_0^\dagger(t, t_0) \hat{U}(t, t_0) \\
 &= [\hat{U}_0^\dagger(t, t_0) V_t \hat{U}_0(t, t_0)] \hat{U}_I(t, t_0), \quad (9.52)
 \end{aligned}$$

which completes the proof.

Thus if we want to work in the interaction frame, we simply use the interaction-frame Hamiltonian (9.49) and treat it as we would any other Hamiltonian. If we want to get back to the usual frame, we simply add the free evolution  $\hat{U}_0(t, t_0)$  at the end. We can use the interaction-frame Hamiltonian in either the Schrödinger picture or the Heisenberg picture. That is:

### Frames versus pictures

The interaction frame can contain a Schrödinger picture:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{V}_I(t) |\psi(t)\rangle ; \quad \hat{A}(t) = \hat{A}(t_0).$$

or a Heisenberg picture:

$$|\psi(t)\rangle = |\psi(t_0)\rangle ; \quad i\hbar \frac{d}{dt} \hat{A}(t) = -[\hat{V}_I(t), \hat{A}(t)].$$

- Example (2).** Imagine an optical cavity supporting a mode of frequency  $\omega$ , driven resonantly by an oscillating field. That is to say, the frequency of the driving field is also  $\omega$ . The Hamiltonian for the cavity field mode is\*

$$\hat{H} = \hbar\omega(\hat{a}^\dagger \hat{a} + \frac{1}{2}) + \lambda(e^{i\omega t} \hat{a} + e^{-i\omega t} \hat{a}^\dagger)$$

where  $\lambda$  is a parameter relating to the amplitude of the driving field and its coupling to the cavity field. Let the cavity field mode

---

\*in the rotating wave and semi-classical approximations

begin in the ground (or vacuum) state  $|0\rangle$ . Calculate the state in the interaction frame.

### Solution

Notice that the Hamiltonian is **time-dependent** and so the corresponding time evolution operator is not a simple exponential as in the previous example. Nevertheless, by working in the interaction frame we can side-step this problem.

We split the Hamiltonian into two parts,

$$\hat{H} = \hat{H}_0 + \hat{V}_t$$

where

$$\hat{H}_0 = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2}) \quad (9.53)$$

represents the “free” cavity field mode, that is, without being driven, and where

$$\hat{V}_t = \lambda(e^{i\omega t}\hat{a} + e^{-i\omega t}\hat{a}^\dagger) \quad (9.54)$$

represents the interaction with the driving field. The Hamiltonian in the interaction frame is given by Eq. (9.49):

$$\hat{V}_I(t) = \hat{U}_0^\dagger(t, 0) \hat{V}_t \hat{U}_0(t, 0) \quad (9.55)$$

where

$$\hat{U}_0(t, 0) = \exp[-\frac{i}{\hbar}\hat{H}_0 t] = \exp[-i\omega t(\hat{a}^\dagger\hat{a} + \frac{1}{2})].$$

Evaluating  $\hat{V}_I(t)$  in Eq. (9.55) yields

$$\begin{aligned} \hat{U}_0^\dagger(t, 0) & \hat{V}_t \hat{U}_0(t, 0) \\ &= \hat{U}_0^\dagger(t, 0) \left[ \lambda(e^{i\omega t}\hat{a} + e^{-i\omega t}\hat{a}^\dagger) \right] \hat{U}_0(t, 0) \\ &= \lambda \left[ e^{i\omega t}\hat{U}_0^\dagger(t, 0)\hat{a}\hat{U}_0(t, 0) + e^{-i\omega t}\hat{U}_0^\dagger(t, 0)\hat{a}^\dagger\hat{U}_0(t, 0) \right]. \end{aligned}$$

Now  $\hat{U}_0(t, 0)$  is the evolution operator for the Harmonic oscillator. We showed in the previous example that this had the following action:

$$\begin{aligned} \hat{U}_0^\dagger(t, 0)\hat{a}\hat{U}_0(t, 0) &= e^{-i\omega t}\hat{a} \\ \hat{U}_0^\dagger(t, 0)\hat{a}^\dagger\hat{U}_0(t, 0) &= e^{i\omega t}\hat{a}^\dagger. \end{aligned}$$

Using these results we have

$$\hat{U}_0^\dagger(t, 0) \hat{V}_t \hat{U}_0(t, 0) = \lambda(\hat{a} + \hat{a}^\dagger).$$

Substituting this result into Eq. (9.55) then gives

$$\hat{V}_I(t) = \lambda(\hat{a} + \hat{a}^\dagger).$$

Hence, by working in the interaction frame we can reduce the problem to that for a simple Hamiltonian  $\hat{V}_I(t)$ , which is *time-independent!* Moving to the interaction frame has cancelled the

time-dependence of  $\hat{V}_t$ .

We can now solve the problem either using the Schrödinger picture or the Heisenberg picture. Here we choose to use the Schrödinger picture. The Schrödinger equation in the interaction frame is

$$i\hbar \frac{d}{dt} |\psi_I(t)\rangle = \hat{V}_I(t) |\psi_I(t)\rangle . \quad (9.56)$$

For a cavity field mode which is initially in the ground (or vacuum) state  $|0\rangle$ , the solution is

$$\begin{aligned} |\psi_I(t)\rangle &= \exp[-\frac{i}{\hbar} \hat{V}_I(t) \times t] |0\rangle = \exp[-\frac{i}{\hbar} \lambda t (\hat{a} + \hat{a}^\dagger)] |0\rangle \\ &\equiv |-\frac{i}{\hbar} \lambda t\rangle_{coh} . \end{aligned}$$

Here the final state is the **coherent state**  $|\alpha\rangle$  with amplitude  $\alpha = -\frac{i}{\hbar} \lambda t$  which increases linearly with time. A brief description of coherent states is given in the Appendix.

• • •

**ASIDE:** For comparison, from Eq. (9.48), if we don't want to work in the interaction frame, we simply have to add the "free" evolution  $\hat{U}_0(t, 0)$  at the end, so that

$$|\psi(t)\rangle = \hat{U}_0(t, 0)|\psi_I(t)\rangle = \exp[-i\omega t(\hat{a}^\dagger \hat{a} + \frac{1}{2})] \left| -\frac{i}{\hbar} \lambda t e^{-i\omega t} \right\rangle_{\text{coh}} \quad (9.57)$$

This can be shown to equal a coherent state with linearly increasing amplitude **and a rotating optical phase**:

$$|\psi(t)\rangle = e^{-i\omega t \frac{1}{2}} \left| -\frac{i}{\hbar} \lambda t e^{-i\omega t} \right\rangle_{\text{coh}}. \quad (9.58)$$

This is the quantum analogue of a classical oscillator with frequency  $\omega$  and amplitude that increases linearly in time. For example, if we work out the expectation value of the electric field  $\langle \hat{E}(t) \rangle$ , we find it oscillates sinusoidally with increasing amplitude.

The electric field amplitude is defined by [see e.g. R. Loudon, **The quantum theory of light** (Oxford Uni Press, 2000)]

$$\hat{E} = \sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} (\hat{a} + \hat{a}^\dagger)$$

where  $V$  is the volume of the cavity mode. Remembering that the **global phase factor for the wavefunction**,  $e^{-i\omega t \frac{1}{2}}$ , is irrelevant to all expectation values, the expectation value in the interaction frame is thus

$$\begin{aligned} \langle \hat{E}(t) \rangle_I &= \langle \psi_I(t) | \hat{E} | \psi_I(t) \rangle \\ &= \sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} \left\langle -\frac{i}{\hbar} \lambda t e^{-i\omega t} \right|_{\text{coh}} (\hat{a} + \hat{a}^\dagger) \left| -\frac{i}{\hbar} \lambda t e^{-i\omega t} \right\rangle_{\text{coh}} \end{aligned}$$

Using the results  $\langle \alpha | \hat{a} | \alpha \rangle = \alpha$  and  $\langle \alpha | \hat{a}^\dagger | \alpha \rangle = \alpha^*$  for coherent state  $|\alpha\rangle \equiv |\alpha\rangle_{\text{coh}}$  we find

$$\langle \hat{E}(t) \rangle = -\sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} \frac{2\lambda t}{\hbar} \sin(\omega t). \quad (9.59)$$

### Problem

Calculate the variance  $\Delta^2 E$  in the electric field as a function of time for the example on page 16.

[Hint: the cavity field mode is always a coherent state.]

## Appendix

We briefly outline the properties of coherent states. These states are defined by

$$\begin{aligned} |\alpha\rangle &= \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})|0\rangle \\ &= e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \end{aligned}$$

The coherent state is an eigenstate of the annihilation operator  $\hat{a}$ , i.e.

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle .$$

The photon number distribution of  $|\alpha\rangle$  is Poissonian:

$$P_n = |\langle n|\alpha\rangle|^2 = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}$$

where the mean is  $\langle \hat{a}^\dagger \hat{a} \rangle = |\alpha|^2$ . Coherent states are important because they closely approximate the output of a laser.

The operators

$$\hat{x} = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}}, \quad \hat{p} = \frac{\hat{a} - \hat{a}^\dagger}{i\sqrt{2}},$$

are analogous to the position and momentum operators in Example 1. It is not difficult to show that their variances for coherent states are

$$\Delta^2 x = \frac{1}{2}, \quad \Delta^2 p = \frac{1}{2}$$

where  $\Delta^2 A \equiv \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2$ , that their commutator is

$$[\hat{x}, \hat{p}] = i$$

and that the Heisenberg uncertainty relation is therefore

$$\Delta^2 x \Delta^2 p \geq \frac{1}{4} .$$

Thus coherent states are also important because they give the minimum uncertainty in  $\hat{x}$  and  $\hat{p}$  allowed by Heisenberg's uncertainty relation.

For more details see R. Loudon, **The quantum theory of light** (Oxford Uni Press, 2000).

# Problem Sheet

- (1) Show that  $\hat{U}(t, t_0)$  in Eq. (9.9) is a unitary operator (i.e. that  $\hat{U}^\dagger(t, t_0)\hat{U}(t, t_0) = \hat{U}(t, t_0)\hat{U}^\dagger(t, t_0) = \mathbf{1}$ ).
- (2) Show that, for a time-independent Hamiltonian, Eq. (9.17) reduces to Eq. (9.9).
- (3) Use the same technique that was used to derive the Heisenberg equation of motion on page 8 to show that the equation of motion for  $\rho(t)$  in the **Schrödinger picture** is

$$i\hbar \frac{d\rho(t)}{dt} = [\hat{H}_t, \rho(t)] .$$

- (4) Show that the commutator between any two operators is preserved in the Heisenberg picture, e.g. if

$$[\hat{A}(t_0), \hat{B}(t_0)] = \hat{C}(t_0)$$

in the Schrödinger picture, then

$$[\hat{A}(t), \hat{B}(t)] = \hat{C}(t)$$

in the Heisenberg picture where  $\hat{A}(t) = \hat{U}^\dagger(t, t_0)\hat{A}(t_0)\hat{U}(t, t_0)$  etc.

- (5) Derive the operator identity

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A}, \hat{B}] + \frac{1}{2!}[\hat{A}, [\hat{A}, \hat{B}]] + \frac{1}{3!}[\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \dots$$

(Hint: Expand the exponential functions in terms of the power series  $e^{\hat{A}} = \mathbf{1} + \hat{A} + \frac{1}{2}\hat{A}^2 + \frac{1}{3!}\hat{A}^3 + \dots$  and collect terms starting with the lowest order. Be careful to preserve the operator ordering as  $\hat{A}$  and  $\hat{B}$  do not commute in general.)

- (6) Use the solution  $\hat{X}(t)$  in the Heisenberg picture, Eq. (9.39), to calculate the time dependence of the variance  $\Delta^2 X$ .

[Hints: Use Eq. (9.39) to get an expression for  $\hat{X}^2(t)$ . For the expectation value calculation  $\langle \hat{X}^2 \rangle$  use the fact that  $e^{i\hat{P}_c}$  is unitary and so  $e^{i\hat{P}_c}\hat{X}^2e^{-i\hat{P}_c} = (e^{i\hat{P}_c}\hat{X}e^{-i\hat{P}_c})(e^{i\hat{P}_c}\hat{X}e^{-i\hat{P}_c})$  where  $c = L/\hbar$ . Then use Eq. (9.43).]

- (7) The Hamiltonian for a single spin-1/2 particle precessing about the  $z$ -axis under the influence of an external magnetic field is

$$\hat{H} = -\left(\frac{eB}{mc}\right)\hat{S}_z .$$

Write down the Heisenberg equations of motion for  $\hat{S}_x$ ,  $\hat{S}_y$  and  $\hat{S}_z$ , the spin operators introduced in chapter 2. Solve

them to find  $\langle \hat{S}_x(t) \rangle$ ,  $\langle \hat{S}_y(t) \rangle$  and  $\langle \hat{S}_z(t) \rangle$  given the initial state  $|\psi(0)\rangle = \alpha|\frac{1}{2}\rangle_z + \beta|-\frac{1}{2}\rangle_z$ .

[Hint: For the two coupled differential equations  $\dot{x} = -ay$  and  $\dot{y} = ax$  has the following solution:  $x(t) = \cos(at)x(0) - \sin(at)y(0)$  and  $y(t) = \sin(at)x(0) + \cos(at)y(0)$ .]

- (8) Calculate the variance  $\Delta^2 E$  in the electric field as a function of time for the example on page 16.

[Hint: the cavity field mode is always a coherent state.]

# 10. Dynamical Perturbation Theory

## a. Why use perturbation theory?

Say the evolution due to the system Hamiltonian  $\hat{H}$  is difficult to solve but  $\hat{H}$  is close to some other Hamiltonian  $\hat{H}_0$  whose evolution is easy to solve. Then we can regard  $\hat{H}$  as a perturbation of  $\hat{H}_0$  and use perturbation theory to find a solution to a suitable order of approximation.

Multiple-order time-**in**dependent (for calculating stationary states) and first-order time-**de**pendent (for calculating dynamics) perturbation theory was covered in a previous course. You may like to re-read those lectures. We begin by reviewing the time-dependent results from that course.

## b. 1<sup>st</sup> order time-dependent perturbation theory

Consider a system with a total Hamiltonian  $\hat{H} = \hat{H}_0 + \hat{V}_t$  consisting of a time-independent part  $\hat{H}_0$  (the unperturbed Hamiltonian) and a time-dependent part  $\hat{V}_t$  (the perturbation Hamiltonian). Assume that the eigenstates and eigenvalues of  $\hat{H}_0$  are known:

$$\hat{H}_0|\psi_j^{(0)}\rangle = E_j^{(0)}|\psi_j^{(0)}\rangle .$$

Examples of such situations include an electron in an atom interacting with (or perturbed by) time-varying and spatially varying electromagnetic fields or a harmonic oscillator being driven weakly.

### KEY POINT

The system Hamiltonian is time dependent  $\hat{H}(t) = \hat{H}_0 + \hat{V}_t$ . It can be split into:

- a time **in**dependent part  $\hat{H}_0$ , and
- a perturbation part  $\hat{V}_t$  which may be time **de**pendent.

Note that if  $\hat{V}_t = 0$  the solution is given by

$$\begin{aligned} |\psi(t)\rangle &= \hat{U}_0(t, 0)|\psi(0)\rangle = \exp(-\frac{i}{\hbar}\hat{H}_0 t)|\psi(0)\rangle \\ &= \sum_{j=0}^{\infty} a_j \exp(-\frac{i}{\hbar}E_j^{(0)} t)|\psi_j^{(0)}\rangle . \end{aligned}$$

where

$$a_j = \langle\psi_j^{(0)}|\psi(0)\rangle .$$

This is the solution we discussed in the chapter **9. Time Evolution**.

To include the effect of  $\hat{V}_t \neq 0$  we let the coefficients  $a_j$  depend on time. The problem then is to find their time evolution. Thus we begin with

$$|\psi(t)\rangle = \sum_{j=0}^{\infty} a_j(t) \exp\left(-\frac{i}{\hbar} E_j^{(0)} t\right) |\psi_j^{(0)}\rangle . \quad (10.1)$$

which we then substitute into the Schrödinger equation; this yields the differential equation for the coefficient  $a_j$ :

$$\dot{a}_k = \frac{1}{i\hbar} \sum_n V_{kn} a_n \exp(i\omega_{kn} t) \quad (10.2)$$

where  $V_{kn} = \langle \psi_k^{(0)} | \hat{V}_t | \psi_n^{(0)} \rangle$  and  $\omega_{kn} = \frac{1}{\hbar}(E_k^{(0)} - E_n^{(0)})$ .

These coupled equations have the general solution

$$a_k(t) = a_k(0) + \frac{1}{i\hbar} \sum_n \int_0^t V_{kn}(t') a_n(t') \exp(i\omega_{kn} t) . \quad (10.3)$$

This is only an *implicit* solution because the right-hand-side contains the coefficients  $a_n(t')$  — the very things we are trying to find. However, we can find an approximate solution, first-order in time, by replacing  $a_n(t')$  by  $a_n(0)$ . Then we have the *explicit* solution

$$a_k(t) \approx a_k(0) + \frac{1}{i\hbar} \sum_n a_n(0) \int_0^t V_{kn}(t') \exp(i\omega_{kn} t) . \quad (10.4)$$

For example, if the system begins in the eigenstate  $|\psi_j^{(0)}\rangle$ ,  $a_n(0) = \delta_{n,j}$  and the first-order solution is given by

insert  
1

$$\begin{aligned} a_j(t) &\approx 1 - \frac{i}{\hbar} \int_0^t dt' V_{jj}(t') \\ a_k(t) &\approx -\frac{i}{\hbar} \int_0^t dt' V_{kj}(t') \exp(i\omega_{kj} t') \quad \text{for } k \neq j . \end{aligned}$$

## c. Higher-order time-dependent perturbation theory

The above method can be extended to high-orders simply by substituting in the approximate solutions (10.4) to the right-hand-side of Eq. (10.2). Since these approximations are already a first-order solution (rather than the zeroth-order solution  $a_n(t') = a_n(0)$  we used the first time), they give now a *second-order* approximate solution.

Rather than proceed this way, we will use the more powerful tools we developed in the preceding chapter. These tools enable us to develop approximations to the unitary evolution operator  $\hat{U}(t, t_0)$ , which can then be applied in either the Schrödinger picture (as used above) or the Heisenberg picture. We derive these approximations in the *interaction frame* with respect to  $\hat{H}_0$ , which makes sense since  $\hat{V}_t$  is meant to be a perturbation on  $\hat{H}_0$ . Note that the approach used above is really working in the interaction frame already because the free evolution has been taken into account through the rotating terms in Eq. (10.1). That is, the coefficients  $a_k(t)$  are really coefficients of  $|\psi_I(t)\rangle$ , the Schrödinger picture state in the interaction frame.

Recall from chapter 8 that the total time-evolution unitary operator for a Hamiltonian  $\hat{H}_t = \hat{H}_0 + \hat{V}(t)$  can be written as

$$\hat{U}(t, t_0) = \hat{U}_0(t, t_0)\hat{U}_I(t, t_0). \quad (10.5)$$

Here

$$\hat{U}_0(t, t_0) = \exp[-(i/\hbar)\hat{H}_0 \times (t - t_0)], \quad (10.6)$$

(we assume that this is known — that is, that the eigenstates and eigenvalues of  $\hat{H}_0$  are known), and  $\hat{U}_I(t, t_0)$  is the evolution generated by

$$\hat{V}_I(t) = \hat{U}_0^\dagger(t, t_0)\hat{V}_t\hat{U}_0(t, t_0). \quad (10.7)$$

From the previous chapter, we have

insert  
2

$$\begin{aligned} \hat{U}_I(t, t_0) &= \mathbb{1} - \frac{i}{\hbar} \int_{t_1=t_0}^t dt_1 \hat{V}_I(t_1) \\ &\quad - \frac{1}{\hbar^2} \int_{t_1=t_0}^t dt_1 \hat{V}_I(t_1) \int_{t_2=t_0}^{t_1} dt_2 \hat{V}_I(t_2) \\ &\quad + \left(\frac{-i}{\hbar}\right)^3 \int_{t_1=t_0}^t dt_1 \hat{V}_I(t_1) \int_{t_2=t_0}^{t_1} dt_2 \hat{V}_I(t_2) \int_{t_3=t_0}^{t_2} dt_3 \hat{V}_I(t_3) \\ &\quad + \dots . \end{aligned} \quad (10.8)$$

This is an infinite series, and every term must be included to obtain the correct result. \* However, an  $n$ -th order approximation to  $\hat{U}_I(t, t_0)$ , and hence to  $\hat{U}(t, t_0)$ , may be obtained by taking the first  $n$  terms (after the zeroth-order  $\mathbb{1}$  term in Eq. (10.8)). Technically, this is interpreting Eq. (10.8) as an **asymptotic expansion**: the larger  $n$  is, the better approximation this will be to the full solution, but only if the perturbation  $\hat{V}_t$  is weak enough, or  $t - t_0$  is short enough. The first-order approximation is the same as that given in Sec. II above.

- **Example (1).**

Consider an optical cavity driven resonantly by a field at angular frequency  $\omega$  as we did for Example 2 in section **9. Time Evolution**. We describe this by the Hamiltonian

$$\hat{H}_t = \hbar\omega\hat{a}^\dagger\hat{a} + \lambda(t)(e^{i\omega t}\hat{a} + e^{-i\omega t}\hat{a}^\dagger).$$

Here  $\lambda(t)$  is proportional to the strength of the driving field, which may vary in time as e.g.  $\lambda(t) = \gamma \exp(-t/\tau)$  for decay constant  $\tau$ . (We've also ignored the  $\hbar\omega/2$  term in the free Hamiltonian, as it has no physical effect.) Calculate the first order approximation to the time evolution operator in the Schrödinger picture starting at  $t_0 = 0$ . Assuming the cavity field mode begins in the ground (or vacuum) state  $|0\rangle$ , what is the state at time  $t \gg \tau$  in this approximation?

### Solution

We let  $\hat{H}_t = \hat{H}_0 + \hat{V}_t$  where

$$\begin{aligned}\hat{H}_0 &= \hbar\omega\hat{a}^\dagger\hat{a} \\ \hat{V}_t &= \lambda(t)(e^{i\omega t}\hat{a} + e^{-i\omega t}\hat{a}^\dagger).\end{aligned}$$

The time evolution operator is given by

$$\hat{U}(t, 0) = \hat{U}_0(t, 0)\hat{U}_I(t, 0) \quad (10.9)$$

where

$$\hat{U}_0(t, 0) = \exp[-i\omega t\hat{a}^\dagger\hat{a}]$$

and where the first order approximation to  $\hat{U}_I(t, t_0)$  is given by

$$\hat{U}_I(t, 0) \approx \mathbb{1} + \frac{1}{i\hbar} \int_0^t dt' \hat{V}_I(t') \quad (10.10)$$

where

$$\begin{aligned}\hat{V}_I(t) &= \hat{U}_0^\dagger(t, 0) \hat{V}_t \hat{U}_0(t, 0) \\ &= \hat{U}_0^\dagger(t, 0) \lambda(t)(e^{i\omega t}\hat{a} + e^{-i\omega t}\hat{a}^\dagger) \hat{U}_0(t, 0).\end{aligned}$$

insert  
3

---

\*There is in fact no guarantee that the series will converge, if the Hamiltonians are unbounded. This is a fact we have glossed over, as it is not relevant to any Hamiltonians we will deal with, but is an important question in quantum field theory.

We know from Example 2 in section **9. Time Evolution** that

$$\begin{aligned}\hat{U}_0^\dagger(t, 0)\hat{a}\hat{U}_0(t, 0) &= e^{-i\omega t}\hat{a} \\ \hat{U}_0^\dagger(t, 0)\hat{a}^\dagger\hat{U}_0(t, 0) &= e^{i\omega t}\hat{a}^\dagger,\end{aligned}$$

and so

$$\hat{V}_I(t) = \lambda(t)(\hat{a} + \hat{a}^\dagger). \quad (10.11)$$

Evaluating Eq. (10.10) using Eq. (10.11) yields

$$\begin{aligned}\hat{U}_I(t, t_0) &\approx \mathbb{1} + \frac{1}{i\hbar} \int_0^t dt' \lambda(t')(\hat{a} + \hat{a}^\dagger) \\ &= \mathbb{1} + \frac{\gamma\tau(1 - e^{-t/\tau})}{i\hbar}(\hat{a} + \hat{a}^\dagger)\end{aligned}$$

Substituting this result into Eq. (10.9) yields the first order approximation as required:

$$\hat{U}(t, 0) = \left[ \exp[-i\omega t\hat{a}^\dagger\hat{a}] \right] \left[ \mathbb{1} + \frac{\gamma\tau(1 - e^{-t/\tau})}{i\hbar}(\hat{a} + \hat{a}^\dagger) \right].$$

If  $t \gg \tau$ , the state is

$$|\psi(t)\rangle = \exp[-i\omega t\hat{a}^\dagger\hat{a}] \left[ \mathbb{1} + \frac{\gamma\tau}{i\hbar}(\hat{a} + \hat{a}^\dagger) \right] |0\rangle \quad (10.12)$$

$$= \exp[-i\omega t\hat{a}^\dagger\hat{a}] \left[ |0\rangle + \frac{\gamma\tau}{i\hbar}|1\rangle \right] \quad (10.13)$$

$$= |0\rangle + \alpha(t)|1\rangle, \quad (10.14)$$

where  $\alpha(t) = e^{-i\omega t}\gamma\tau/i\hbar$ . Note that this parameter must be small for the perturbative solution to make sense.

• • •

### Problem

Obtain expressions for the second-order approximations to  $\hat{U}_I$ , and find the state for times  $t \gg \tau$ , to second-order in  $\alpha(t)$ . Show that this state is equal, to second-order in  $\alpha(t)$ , to the *coherent state*  $|\alpha(t)\rangle$ . Explain why, with reference to results in the previous section.

### Problem

Consider the evolution of a cavity mode initially in state  $|0\rangle$ , under the Hamiltonian

$$\hat{H}_t = \hbar\omega\hat{a}^\dagger\hat{a} + \lambda(t)(e^{i2\omega t}\hat{a}^2 + e^{-i2\omega t}\hat{a}^{\dagger 2}), \quad (10.15)$$

where  $\lambda(t)$  is as previously. Show that the first-order approximation to the long-time ( $t \gg \tau$ ) state is

$$|\psi(t)\rangle = |0\rangle + \beta(t)|2\rangle, \quad (10.16)$$

where  $\beta(t) = \sqrt{2}e^{-i2\omega t}\gamma\tau/i\hbar$ , and we require  $|\beta| \ll 1$  for the perturbation to be valid.

This describes a degenerate parametric amplifier. The second term in Eq. (10.15) comes from driving the cavity with a field at *twice* the resonant frequency. If the cavity contains a crystal with a so-called  $\chi^{(2)}$  nonlinearity, some of that high frequency light can be *down-converted* to photons in the cavity. As our perturbative solution shows, it actually produces *pairs* of photons in the mode.

### Problem

Consider the evolution of two modes initially in state  $|0\rangle|0\rangle$ , under the Hamiltonian

$$\hat{H}_t = \hbar\omega(\hat{a}^\dagger\hat{a} + \hat{b}^\dagger\hat{b}) + \lambda(t)(e^{i2\omega t}\hat{a}\hat{b} + e^{-i2\omega t}\hat{a}^\dagger\hat{b}^\dagger), \quad (10.17)$$

where  $\lambda(t)$  is as previously. Show that the first-order approximation to the long-time ( $t \gg \tau$ ) state is

$$|\psi(t)\rangle = |0\rangle|0\rangle + \beta(t)|1\rangle|1\rangle, \quad (10.18)$$

where we require  $\beta(t) = e^{-i2\omega t}\gamma\tau/i\hbar$  to be  $\ll 1$  in magnitude.

This describes a **non**-degenerate parametric amplifier, so-called because modes  $a$  and  $b$  are different — for example they may propagate in different directions, even though they have the same frequency. This is one method by which single photons are produced for quantum information tasks such as quantum cryptography. This might seem odd, because the state Eq. (10.18) almost certainly contains zero photons, with only a small amplitude to have two photons, one in each mode. The way single photons are produced is that if one photon is detected in mode  $b$  (which destroys that photon), the state collapses so that there is exactly one photon in mode  $a$ .

### Problem

Show this. The measurement operator  $\hat{M}$  for detecting a photon in mode  $b$  can be taken to be  $|0\rangle\langle 1|$ , where these are mode  $b$  number states.

# Problem Sheet

- (1) Obtain expressions for the second-order approximations to  $\hat{U}_I$ , and find the state for times  $t \gg \tau$ , to second-order in  $\alpha(t)$ . Show that this state is equal, to second-order in  $\alpha(t)$ , to the *coherent state*  $|\alpha(t)\rangle$ . Explain why, with reference to results in the previous section.
- (2) Consider the evolution of a cavity mode initially in state  $|0\rangle$ , under the Hamiltonian

$$\hat{H}_t = \hbar\omega\hat{a}^\dagger\hat{a} + \lambda(t)(e^{i2\omega t}\hat{a}^2 + e^{-i2\omega t}\hat{a}^{\dagger 2}) ,$$

where  $\lambda(t)$  is as previously. Show that the first-order approximation to the long-time ( $t \gg \tau$ ) state is

$$|\psi(t)\rangle = |0\rangle + \beta(t)|2\rangle ,$$

where  $\beta(t) = \sqrt{2}e^{-i2\omega t}\gamma\tau/i\hbar$ , and we require  $|\beta| \ll 1$  for the perturbation to be valid.

- (3) Consider the evolution of two modes initially in state  $|0\rangle|0\rangle$ , under the Hamiltonian

$$\hat{H}_t = \hbar\omega(\hat{a}^\dagger\hat{a} + \hat{b}^\dagger\hat{b}) + \lambda(t)(e^{i2\omega t}\hat{a}\hat{b} + e^{-i2\omega t}\hat{a}^\dagger\hat{b}^\dagger) ,$$

where  $\lambda(t)$  is as previously. Show that the first-order approximation to the long-time ( $t \gg \tau$ ) state is

$$|\psi(t)\rangle = |0\rangle|0\rangle + \beta(t)|1\rangle|1\rangle ,$$

where we require  $\beta(t) = e^{-i2\omega t}\gamma\tau/i\hbar$  to be  $\ll 1$  in magnitude.

- (4) In the previous problem, show that if one photon is detected in mode  $b$  (which destroys that photon), the state collapses so that there is exactly one photon in mode  $a$ . The measurement operator  $\hat{M}$  for detecting a photon in mode  $b$  can be taken to be  $|0\rangle\langle 1|$ , where these are mode  $b$  number states.