Module F14ZD1: Quantum Computing

Bernd Schroers **2007/08**

"Information is physical" (Rolf Landauer)

Definition of the subject in two lines: "Quantum computing is the study of information processing which may be realised by physical systems obeying the laws of quantum mechanics."

Table of Contents

1 Introduction

The following subsections are modified excerpts form articles on quantum mechanics and quantum computing in the the on-line encyclopedia Wikipedia, `http://www.wikipedia.org`.

1.1 Quantum mechanics

Quantum mechanics is the framework in which most fundamental physical theories are formulated. There exist quantum versions of most classical theories, including mechanics and electromagnetism (but not general relativity), which provide accurate descriptions for many previously unexplained phenomena such as black body radiation and stable electron orbits. The effects of quantum mechanics are typically not observable on macroscopic scales, but become evident at the atomic and subatomic level. The term quantum (Latin, "how much") refers to the discrete units that the theory assigns to certain physical quantities, such as the energy of an atom at rest.

Quantum mechanics has had enormous success in explaining many of the features of our world. The individual behaviour of the microscopic particles that make up all forms of matter, such as electrons, protons or neutrons, can often only be satisfactorily described using quantum mechanics. The application of quantum mechanics to chemistry - known as quantum chemistry - can provide quantitative insight into chemical bonding processes by explicitly showing which molecules are energetically favourable to which others, and by approximately how much. Most of the calculations performed in computational chemistry rely on quantum mechanics.

Much of modern technology operates at a scale where quantum effects are significant. Examples include the laser, the transistor, the electron microscope, and magnetic resonance imaging. The study of semiconductors led to the invention of the diode and the transistor, which are indispensable for modern electronics.

In the formalism of quantum mechanics, the state of a system at a given time is described by an element of a complex vector space. This abstract mathematical object allows for the calculation of probabilities of outcomes of concrete experiments. For example, it allows one to compute the probability of finding an electron in a particular region around the nucleus at a particular time. Contrary to classical mechanics, one can never make simultaneous predictions of conjugate quantities, such as position and momentum, with arbitrary accuracy. Heisenberg's uncertainty principle quantifies the inability to precisely specify conjugate quantities.

Quantum mechanics remains the subject of intense research, both concerning applications and the foundations of the subject. One important challenge is to find robust methods for directly manipulating quantum states. Efforts are being made to develop quantum cryptography, which will allow guaranteed secure transmission of information. A long-term goal is the development of quantum computers, which are expected to perform certain computational tasks exponentially faster than classical computers. Another active research topic is quantum teleportation, which deals with techniques to transmit quantum states over arbitrary distances.

1.2 A brief history of quantum mechanics

The foundations of quantum mechanics were established during the first half of the 20th century by Max Planck (1858-1947), Albert Einstein (1879-1955), Niels Bohr 1885-1962), Werner Heisenberg (1901-1976), Erwin Schrödinger (1887-1961), Max Born (1882-1970), John von Neumann (1903-1957), Paul Dirac (1902-1984), Wolfgang Pauli (1900-1958) and others.

In 1900, Max Planck introduced the idea that energy is quantised, in order to derive a formula for the observed frequency dependence of the energy emitted by a black body. In 1905, Einstein explained the photoelectric effect by postulating that light energy comes in quanta called photons. In 1913, Bohr explained the spectral lines of the hydrogen atom, again by using quantisation. In 1924, Louis de Broglie put forward his theory of matter waves.

These theories, though successful, were strictly phenomenological: there was no rigorous justification for quantisation. They are collectively known as the old quantum theory.

Modern quantum mechanics was born in 1925, when Heisenberg developed matrix mechanics and Schrödinger invented wave mechanics and the Schrödinger equation. Schrödinger subsequently showed that the two approaches were equivalent.

Heisenberg formulated his uncertainty principle in 1927, and the Copenhagen interpretation took shape at about the same time. Starting around 1927, Paul Dirac unified quantum mechanics with special relativity. He also pioneered the use of operator theory, including the influential bra-ket notation, as described in his famous 1930 textbook. During the same period, John von Neumann formulated the rigorous mathematical basis for quantum mechanics as the theory of linear operators on Hilbert spaces, as described in his likewise famous 1932 textbook. These, like many other works from the founding period still stand, and remain widely used.

1.3 Quantum computing

A quantum computer is any device for computation that makes direct use of distinctively quantum mechanical phenomena, such as superposition and entanglement, to perform operations on data. In a classical (or conventional) computer, the amount of data is measured by bits; in a quantum computer, it is measured by qubits. The basic principle of quantum computation is that the quantum properties of particles can be used to represent and structure data, and that devised quantum mechanisms can be used to perform operations with these data.

Experiments have already been carried out in which quantum computational operations were executed on a very small number of qubits. Research in both theoretical and practical areas continues at a frantic pace. Many national government and military funding agencies support quantum computing research, to develop quantum computers for both civilian and national security purposes, such as cryptanalysis.

It is widely believed that if large-scale quantum computers can be built, they will be able to solve certain problems faster than any classical computer. Quantum computers are different from classical computers based on transistors, even though these may ultimately use some kind of quantum mechanical effect. Some computing architectures such as optical computers may use classical superposition of electromagnetic waves, but without some specifically quantum mechanical resource such as entanglement, they do not share the potential for computational speed-up of quantum computers.

2 Algebraic Structures

In this section we review some algebraic structures which you studied in the second year module on linear algebra, and introduce some algebraic concepts which you have not yet come across. I will not give formal definitions and proofs of concepts and results which you studied in second year, but will remind you of the basic ideas. Please refer to your notes on linear algebra for further details. All new concepts will be carefully defined and I will give plenty of examples for both old and new material.

2.1 Vector spaces

2.1.1 Basic concepts and notation

A vector space is a set whose elements one can add together and multiply by a number, often called a scalar, and which contains a special element 0, the zero vector. The scalar will generally be a complex number in this course. Vector spaces with complex numbers as scalars are called **complex vector spaces**. In quantum mechanics, the vectorial nature of a quantity v is usually expressed by enclosing it between a vertical line and a right bracket $|v\rangle$. We will adopt this convention here, which goes back to Paul Dirac, who also introduced the name "ket" for a vector. As we shall see later, this name is motivated by thinking of a vector as "half a brac-ket".

Example 2.1.1 *The set $\mathbb{C}^2$ of column vectors made up of two complex numbers is a complex vector space. Find the vector obtained by adding the vectors*

$$|v_1\rangle = \begin{pmatrix} i \\ -4 \end{pmatrix}, \qquad |v_2\rangle = \begin{pmatrix} 6-i \\ 5+i \end{pmatrix},$$

and multiplying the result by the scalar $\alpha = 3e^{i\frac{\pi}{2}}$.

Since $3e^{i\frac{\pi}{2}} = 3i$ we have

$$\alpha(|v_1\rangle + |v_2\rangle) = 3i \begin{pmatrix} 6 \\ 1+i \end{pmatrix} = \begin{pmatrix} 18i \\ -3+3i \end{pmatrix}.$$

$\square$

Recall that a vector $|v\rangle$ is called a **linear combination** of vectors $|v_1\rangle$ and $|v_2\rangle$ if it can be written

$$|v\rangle = \alpha_1|v_1\rangle + \alpha_2|v_2\rangle$$

for two complex numbers α_1 and α_2. The **span** of a subset $S = \{|v_1\rangle, \ldots, |v_n\rangle\}$ is the set of all linear combinations of the vectors $|v_1\rangle, \ldots, |v_n\rangle$ and denoted $[|v_1\rangle, \ldots, |v_n\rangle]$. We say that the subset $S = \{|v_1\rangle, \ldots, |v_n\rangle\}$ of a vector space V is a **spanning set** if any vector can be written as a linear combination of the vectors $|v_1\rangle, \ldots, |v_n\rangle$ i.e. if $[|v_1\rangle, \ldots, |v_n\rangle] = V$. The vectors $|v_1\rangle, \ldots, |v_n\rangle$ are called **linearly independent** if

$$\sum_{i=1}^{n} \alpha_i |v_i\rangle = 0 \Rightarrow \alpha_i = 0, \qquad i = 1, \ldots, n. \tag{2.1}$$

Conversely, the vectors $|v_1\rangle, \ldots, |v_n\rangle$ are **linearly dependent** if we can find complex numbers $\alpha_1, \ldots, \alpha_n$, not all zero, so that

$$\sum_{i=1}^{n} \alpha_i |v_i\rangle = 0 \tag{2.2}$$

Example 2.1.2 *Show that the vectors* $|v_1\rangle = \begin{pmatrix} 1-i \\ 1 \end{pmatrix}$ *and* $|v_2\rangle = \begin{pmatrix} 1 \\ \frac{1}{2} + \frac{i}{2} \end{pmatrix}$ *in* $\mathbb{C}^2$ *are linearly dependent.*

Since $(1+i)|v_1\rangle = 2|v_2\rangle$ we have $(1+i)|v_1\rangle + (-2)|v_2\rangle = 0$. $\square$

Example 2.1.3 *Suppose that the vectors* $|v_1\rangle, \ldots, |v_n\rangle$ *are linearly independent. Show that a vector* $|v\rangle$ *in* V *can be written as linear combinations of* $|v_1\rangle, \ldots, |v_n\rangle$ *in at most one way.*

Suppose that there are two ways of writing $|v\rangle$ as a linear combination, i.e.

$$v = \sum_{i=1}^{n} \alpha_i |v_i\rangle \tag{2.3}$$

and

$$v = \sum_{i=1}^{n} \beta_i |v_i\rangle. \tag{2.4}$$

Then, taking the difference, we deduce that

$$\sum_{i=1}^{n} (\alpha_i - \beta_i)|v_i\rangle = 0.$$

But since the $|v_i\rangle$ are linearly independent we deduce that $\alpha_i = \beta_i$ for $i = 1, \ldots, n$, so that the two linear combinations (2.3) and (2.4) are in fact the same. $\square$.

A set $S = \{|v_1\rangle, \ldots |v_n\rangle\}$ is called a **basis** of the vector space V if S is both spanning and linearly independent. One can show that every vector space has a basis. The basis is not unique - in fact there are infinitely many different bases as we shall see below - but the number of elements in any basis is the same; that number is called the **dimension** of the vector space. The dimension may be finite or infinite. In this course we only deal with finite dimensional vector spaces. For a vector space of finite dimension n one can show that any set of n linearly independent vectors is automatically spanning, i.e. a basis. In order to check if a given set containing n vectors constitutes a basis we therefore only need to check for linear independence. There are simple tests for this, one of which we give below.

The vector space $\mathbb{C}^n$ has a **canonical basis** consisting of the column vectors

$$|b_1\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \\ 0 \end{pmatrix}, |b_2\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ \\ 0 \end{pmatrix}, \ldots, |b_n\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \\ 1 \end{pmatrix}. \tag{2.5}$$

The space $\mathbb{C}^2$ plays a particularly important role in quantum computing and it is conventional to denote the canonical basis as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.6}$$

The notation anticipates the role of the space $\mathbb{C}^2$ as a quantum bit or qubit. Whereas a classical bit can be in one of two states "0" or "1", quantum bit can be in the basis states $|0\rangle$ or $|1\rangle$ or in **any linear combination** of the basis states. Any two vectors

$$|x\rangle = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad |y\rangle = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

in $\mathbb{C}^2$ are independent (and hence constitute a basis) if the matrix made from the the column vectors has a non-vanishing determinant:

$$\det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \neq 0. \tag{2.7}$$

2.1.2 Coordinates and basis change

Suppose that V is a complex vector space of dimension n and that $B = \{|b_1\rangle, \ldots |b_n\rangle\}$ is a basis of V. Then a vector $|x\rangle$ has a unique expansion

$$|x\rangle = \sum_{i=1}^{n} x_i |b_i\rangle \tag{2.8}$$

in terms of this basis. The complex numbers $x_1, \ldots, x_n$ are called the **coordinates** of the vector $|x\rangle$ with respect to the basis B.

A vector can be expanded in any basis, and its coordinates with respect to different bases differ. We are interested in the change of coordinates under a change of basis. Suppose the basis $B' = \{|b'_1\rangle, \ldots, |b'_n\rangle\}$ of an n-dimensional vector space is obtained from the basis $B = \{|b_1\rangle, \ldots, |b_n\rangle\}$ via

$$|b'_i\rangle = \sum_{j=1}^{n} M_{ji} |b_j\rangle, \qquad \text{for} \quad i = 1, \ldots n, \tag{2.9}$$

where M_{ij} are the matrix elements of an invertible $n \times n$-matrix of complex numbers. Now we have the two expansions

$$|x\rangle = \sum_{j=1}^{n} x_j |b_j\rangle \tag{2.10}$$

and

$$|x\rangle = \sum_{i=1}^{n} x'_i |b'_i\rangle. \tag{2.11}$$

Inserting the relation (2.9) into the expansion (2.11) we have

$$|x\rangle = \sum_{i,j=1}^{n} x'_i M_{ji} |b_j\rangle. \tag{2.12}$$

Comparing with (2.10) and using the uniqueness of expansions in a basis we deduce

$$x_j = \sum_{i=1}^{n} M_{ji} x'_i. \tag{2.13}$$

Collecting the coordinates x_i and x'_j into column vectors this can be written

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & \ldots & M_{1n} \\ M_{21} & M_{22} & \ldots & M_{2n} \\ \vdots & & & \\ & & & \\ M_{n1} & M_{n2} & \ldots & M_{nn} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ \\ x'_n \end{pmatrix} \tag{2.14}$$

or, denoting the matrix with matrix entries M_{ij} by M,

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{pmatrix} = M \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ \\ x'_n \end{pmatrix} \tag{2.15}$$

so that

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ \\ x'_n \end{pmatrix} = M^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{pmatrix}. \tag{2.16}$$

Performing the inversion explicitly in the case $n = 2$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \tag{2.17}$$

we find

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \frac{1}{M_{11}M_{22} - M_{12}M_{21}} \begin{pmatrix} M_{22} & -M_{12} \\ -M_{21} & M_{11} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{2.18}$$

Example 2.1.4 *Give the coordinates of the vector $|x\rangle = i|0\rangle - |1\rangle$ in $\mathbb{C}^2$ in the basis with basis vectors $|v_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|v_2\rangle = \frac{1}{\sqrt{2}}(-|0\rangle + |1\rangle)$*

With

$$M = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

and

$$M^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

we have the coordinates

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} i \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} i-1 \\ -i-1 \end{pmatrix}$$

so that $|x\rangle = \frac{i-1}{\sqrt{2}}|v_1\rangle - \frac{i+1}{\sqrt{2}}|v_2\rangle$.

2.2 Linear maps

Recall that a **linear map** from a vector space V to a vector space W is a map $A : V \to W$ which satisfies $A(\alpha|u\rangle + \beta|v\rangle) = \alpha A(|u\rangle) + \beta A(|v\rangle)$ for any complex numbers α and β and any two elements $|u\rangle$ and $|v\rangle$ in V. In quantum mechanics it is customary to call linear maps **linear operators**, though mathematicians tend to reserve this term for situations where both V and W are infinite dimensional. We will mostly be concerned with the situation $V = W$ in the following. It is not difficult to show (check you linear algebra notes) that a linear map is completely determined by its action on basis of V. This leads to the **matrix representation of a linear map** as follows.

Consider a linear map $A : V \to V$ in a complex vector space of dimension n, and let $B = \{|b_1, \ldots, |b_n\rangle\}$ be a basis of V. We consider the action of A on each of the basis elements, and expand the images in the basis B:

$$A(|b_i\rangle) = \sum_{j=1}^{n} A_{ji}|b_j\rangle. \tag{2.19}$$

The matrix made up of the $n \times n$ numbers $A_{ij}, i, j = 1, \ldots, n$ is the **matrix representation of A** with respect to the basis B. The action of A on a general element $|x\rangle \in V$ can be written conveniently in terms of the matrix representation. With the expansion $|x\rangle = \sum_{i=1}^{n} x_i|b_i\rangle$ we have

$$A(|x\rangle) = \sum_{j=1}^{n} A_{ji}x_i|b_j\rangle. \tag{2.20}$$

so that the coordinates of the image $A(|x\rangle)$ with respect to the basis B are obtained from the coordinates of $|x\rangle$ with respect to B by putting them into a column vector and multiplying them with the matrix representation of A.

Important notational convention: In the following we will often fix one basis for a given vector space V and work with coordinates and matrix representations relative to that basis.

In particular when working with $V = \mathbb{C}^n$ we use the canonical basis (2.5). In that case we do not distinguish notationally between the operator A and its matrix representation relative to the canonical basis.

Example 2.2.1 *The linear map* $A : \mathbb{C}^2 \to \mathbb{C}^2$ *satisfies* $A(|0\rangle) = 3i|0\rangle + 4|1\rangle$ *and* $A(|1\rangle) = 3|0\rangle - 4i|1\rangle$. *Give its matrix representation with respect to the canonical basis, and give the image of the vector* $|x\rangle = |0\rangle - |1\rangle$ *under the action of* A.

The matrix representation is

$$A = \begin{pmatrix} 3i & 3 \\ 4 & -4i \end{pmatrix}$$

so the image of the vector with coordinates $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ has coordinates

$$\begin{pmatrix} 3i & 3 \\ 4 & -4i \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -3+3i \\ 4+4i \end{pmatrix}.$$

□

Before leaving linear operators we need to understand how the matrix representation of an operator A changes when we change the basis of V. Consider again a complex vector space V with two distinct bases B and B'. The basis $B' = \{|b_1'\rangle, \dots, |b_n'\rangle\}$ is obtained from the basis $B = \{|b_1\rangle, \dots, |b_n\rangle\}$ via

$$|b_i'\rangle = \sum_{j=1}^{n} M_{ji}|b_j\rangle, \qquad \text{for} \quad i = 1, \dots n. \tag{2.21}$$

Suppose we are given the matrix representation of A relative to the basis B via (2.19) and would like to know its matrix representation with respect to the basis B'. Defining

$$A(|b_i'\rangle) = \sum_{j=1}^{n} A'_{ji}|b_j'\rangle. \tag{2.22}$$

we replace $|b_i'\rangle$ by the expression in (2.21) and use the linearity of A to deduce

$$\sum_{k=1}^{n} M_{ki}A(|b_k\rangle) = \sum_{j,l=1}^{n} A'_{ji}M_{lj}|b_l\rangle. \tag{2.23}$$

Expanding the left-hand side according to (2.19) we have

$$\sum_{k,l=1}^{n} M_{ki}A_{lk}(|b_l\rangle) = \sum_{j,l=1}^{n} A'_{ji}M_{lj}|b_l\rangle \tag{2.24}$$

Comparing coefficients of basis elements $|b_l\rangle$ we deduce that the matrices M, A, A' satisfy

$$AM = MA' \tag{2.25}$$

or

$$A' = M^{-1}AM \tag{2.26}$$

Example 2.2.2 *See problem sheet 1!*

In the linear algebra course in second year you came across the definition of the **determinant** and the **trace** of a matrix. Using the relation (2.26) we can now define the determinant and trace of a linear map. Although one requires a matrix representation to compute both, the result is independent of the basis to which the matrix representation refers. To see this, recall that for any two $n \times n$ matrices A and B

$$\det(AB) = \det(A)\det(B), \tag{2.27}$$

which implies in particular $\det(A^{-1}) = (\det A)^{-1}$. Recall also that the definition

$$\mathrm{tr}(A) = \sum_{i=1}^{n} A_{ii}, \tag{2.28}$$

which implies

$$\mathrm{tr}(AB) = \sum_{i,j=1}^{n} A_{ij}B_{ji} = \mathrm{tr}(BA). \tag{2.29}$$

It follows that for the two matrices A' and A related by conjugation with M as in (2.26) that

$$\det(A') = (\det(M))^{-1}\det(A)\det(M) = \det(A) \tag{2.30}$$

and

$$\mathrm{tr}(A') = \mathrm{tr}(M^{-1}AM) = \mathrm{tr}(MM^{-1}A) = \mathrm{tr}(A). \tag{2.31}$$

2.3 Inner product spaces

For vector spaces to be of use in quantum mechanics they need to be equipped with an addition structural feature: an inner product or scalar product. For complex vector spaces this is defined as follows.

Definition 2.3.1 *(**Inner product**) An inner product on a complex vector space V is a map*

$$(\ ,\ ) : V \times V \quad \rightarrow \quad \mathbb{C} \tag{2.32}$$

which satisfies

1. $(|v\rangle, \alpha_1|w_1\rangle + \alpha_2|w_2\rangle) = \alpha_1(|v\rangle, |w_1\rangle) + \alpha_2(|v\rangle, |w_2\rangle)$
   (Linearity in the second argument)
2. $(|v\rangle, |w\rangle) = \overline{(|w\rangle, |v\rangle)}$, *(Symmetry)*
3. $|v\rangle \neq 0 \Rightarrow (|v\rangle, |v\rangle) > 0$ *(Positivity)*

Note that the last condition makes sense since $(|v\rangle, |v\rangle)$ is real, which follows directly from condition 2.

Before we study examples we note an important property.

Lemma 2.3.2 *(***Conjugate linearity***) The inner product $(\cdot,\cdot)$ is conjugate-linear in the first argument, i.e.*

$$(\alpha_1|v_1\rangle + \alpha_2|v_2\rangle, |w\rangle) = \bar{\alpha}_1(|v\rangle_1, |w\rangle) + \bar{\alpha}_2(|v_2\rangle, w\rangle) \tag{2.33}$$

Proof: Using the properties of the inner product we compute

$$\begin{aligned}
(\alpha_1|v_1\rangle + \alpha_2|v_2\rangle, |w\rangle) &= \overline{(|w\rangle, \alpha_1|v_1\rangle + \alpha_2|v_2\rangle)} \quad \text{(Property 2)} \\
&= \bar{\alpha}_1\overline{(|w\rangle, |v_1\rangle)} + \bar{\alpha}_2\overline{(|w\rangle, |v_2\rangle)} \quad \text{(Property 1)} \\
&= \bar{\alpha}_1(|v_1\rangle, |w\rangle) + \bar{\alpha}_2(|v_2\rangle, |w\rangle) \quad \text{(Property 2)}.
\end{aligned} \tag{2.34}$$

□

Example 2.3.3 *Define an inner product on $\mathbb{C}^2$ via*

$$(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}) = \bar{x}_1 y_1 + \bar{x}_2 y_2. \tag{2.35}$$

Show that it satisfies all the properties of the definition 2.3.1.

Checking linearity and symmetry of (2.35) is left as a simple exercise. For positivity note that

$$(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}) = |z_1|^2 + |z_2|^2,$$

which is a sum of positive terms and non-vanishing if z_1 and z_2 are not both zero. □

In quantum mechanics it is customary to write

$$(|v\rangle, |w\rangle) = \langle v|w\rangle \tag{2.36}$$

The mathematical motivation for this notation is that in an inner product space every vector $|v\rangle$ defines a linear map

$$\begin{aligned}
\langle v| : V &\rightarrow \mathbb{C} \quad \text{via} \\
|w\rangle &\mapsto \langle v|w\rangle.
\end{aligned} \tag{2.37}$$

The inner product $\langle v|w\rangle$ can thus be thought of as the the map $\langle v|$ evaluated on the vector $|w\rangle$. In quantum mechanics the map $\langle v|$ is called a **bra**: the "left half" of the "bra-ket".

Example 2.3.4 *Suppose that $B = \{|b_1\rangle, \ldots, |b_2\rangle\}$ is an orthogonal basis of V and $|x\rangle = \sum_{i=1}^{n} x_i|b_i\rangle$. Find the matrix representation of the linear map $\langle x|$.*

We have only considered matrix representations of maps $V \to V$ in this course so far, but it is not difficult to extend this notion to the situation $\langle\, x| : V \to \mathbb{C}$. The idea is again to apply the map to each of the basis vectors $|b_i\,\rangle$. We find

$$\langle\, x|b_i\,\rangle = \overline{\langle\, b_i|x\,\rangle} = \bar{x}_i \tag{2.38}$$

There is no need to expand the result in a basis since the target space $\mathbb{C}$ is one-dimensional. Comparing with (2.19) and noting that the index i labels the columns of the matrix representation we conclude that the matrix representation of the map $\langle\, x|$ is the row vector $(\bar{x}_1, \ldots, \bar{x}_n)$. □

The inner product allows one to define the **norm** of a vector and the notion of orthogonality.

Definition 2.3.5 *Let V be a vector space with inner product.*

1. *The norm of a vector $|v\rangle$ is*

$$||v\rangle| = \sqrt{\langle v|v\rangle}. \tag{2.39}$$

2. *Two vectors $|v\rangle$ and $|w\rangle$ are orthogonal if $\langle v|w\rangle = 0$.*
3. *A basis $B = \{|b_1\rangle, \ldots, |b_n\rangle\}$ of V is called orthonormal if*

$$\langle b_i|b_j\rangle = \delta_{ij}, \quad i, j = 1, \ldots, n \tag{2.40}$$

In the last part of the definition we use the **Kronecker delta** symbol: δ_{ij} is 1 when $i = j$ and zero otherwise. Any basis of a vector space V with inner product can be turned into an orthonormal basis by the **Gram-Schmidt** process, which you studied in second year and which I will not review here. Since every vector space has a basis it follows from the Gram-Schmidt procedure that every vector space with an inner product has an orthonormal basis.

Example 2.3.6 *Show that $|b_1\rangle = (\cos\theta|0\rangle + \sin\theta|1\rangle)$ and $|b_2\rangle = i(\cos\theta|1\rangle - \sin\theta|0\rangle)$ form an orthonormal basis of $\mathbb{C}^2$ with the canonical inner product defined in* (2.35) *for any value of the parameter $\theta \in [0, 2\pi)$.*

It is easy to check that $\{|0\rangle, |1\rangle\}$ form an orthonormal basis. Hence $\langle b_1|b_1\rangle = \cos^2\theta + \sin^2\theta = 1$ and similarly $\langle b_2|b_2\rangle = 1$. Moreover $\langle b_1|b_2\rangle = -i\cos\theta\sin\theta + i\cos\theta\sin\theta = 0$. □.

Example 2.3.7 *For the case $V = \mathbb{C}^n$, a canonical inner product is defined via*

$$\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \\ y_n \end{pmatrix} \right) = \sum_{i=1}^{n} \bar{x}_i y_i. \tag{2.41}$$

Check that the canonical basis (2.5) *is an orthonormal basis with respect to this inner product.*

Inserting the coordinate given in (2.5) one finds $\langle b_i | b_j \rangle = \delta_{ij}$ $\square$

The inner product allows one to define the orthogonality not only of vectors but of entire subspaces. For later use we note

Definition 2.3.8 *(**Orthogonal complement**) If W is a subspace of a vector space V with inner product we define the orthogonal complement to be the space*

$$W^{\perp} = \{|v\rangle \in V | \langle v|w\rangle = 0 \quad \textit{for all} \quad |w\rangle \in W\} \tag{2.42}$$

It is not difficult to check that $W^{\perp}$ is indeed a vector space (see problem sheet).

Example 2.3.9 *Let $V = \mathbb{C}^3$ and W be the linear span of $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Find the orthogonal complement of W.*

Elements $|v\rangle = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$ in v are orthogonal to $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ iff $\bar{z}_1 = 0$. Thus

$$W^{\perp} = \{\begin{pmatrix} 0 \\ z_2 \\ z_3 \end{pmatrix} | z_2, z_3 \in \mathbb{C}\}.$$

$\square$

We have already seen in (2.8) that any element $|x\rangle$ of a vector space can be expanded in a given basis. However, in the previous subsection we did not give an algorithm for computing the expansion coefficients x_i. If the vector space V is equipped with an inner product, the computation of the expansion coefficients is considerably simplified. Suppose that $B = \{|b_1\rangle, \ldots, |b_n\rangle\}$ is an orthonormal basis of V and we want to find the coordinates of $|x\rangle$ in this basis:

$$|x\rangle = \sum_{i=1}^{n} x_i |b_i\rangle. \tag{2.43}$$

Acting on both sides of the equation with the bra's $\langle b_j|$, $j = 1, \ldots, n$ we find

$$\langle b_j | x\rangle = \sum_{i=1}^{n} x_i \delta_{ij} = x_j, \tag{2.44}$$

thus giving us an explicit formula for the coordinates x_j.

We can similarly give an explicit formula for the matrix representation of a linear operator A on the vector space V with inner product. We consider the action of A on each of the basis elements in B:

$$A|b_i\rangle = \sum_{k=1}^{n} A_{ki} |b_j\rangle. \tag{2.45}$$

Acting on both sides of the equation with the bra's $\langle b_j|$, $k = 1, \dots, n$ we find

$$\langle b_j|A|b_i\rangle = \sum_{k=1}^{n} A_{ji}\delta_{jk} = A_{ji}, \tag{2.46}$$

The inner product structure even helps in explicitly reconstructing the linear operator A from its matrix representation. For this purpose we introduce the maps

$$\begin{aligned} |b_i\rangle\langle b_j| : V &\rightarrow V \\ |x\rangle &\mapsto |b_i\rangle\langle b_j|x\rangle \end{aligned} \tag{2.47}$$

associated to the elementary bras and kets $\langle b_j|$ and $|b_i\rangle$. We claim

Lemma 2.3.10 *For any linear operator A in a vector space V with inner product and orthonormal basis B we have the representation*

$$A = \sum_{i,j=1}^{n} A_{ij}|b_i\rangle\langle b_j|, \tag{2.48}$$

where $A_{ij} = \langle\, b_i|A|b_j\,\rangle$.

To prove this claim we show the left and the right hand side have the same action on each of the basis vectors $|b_k\rangle$:

$$A|b_k\rangle = \sum_{i,j=1}^{n} A_{ij}|b_i\rangle\langle b_j|b_k\rangle = \sum_{i=1}^{n} A_{ik}|b_i\rangle, \tag{2.49}$$

which is true by the definition of the matrix elements A_{ik}. $\square$

We note in particular

Corollary 2.3.11 *(***Resolution of the identity***) The identity operator $I : V \mapsto V$ has the representation*

$$I = \sum_{i=1} |b_i\rangle\langle b_i| \tag{2.50}$$

This representation of identity is often useful in calculations. As an example we give a quick proof of the

Theorem 2.3.12 *(***Cauchy-Schwarz inequality***) For any two vectors $|\varphi\rangle$ and $|\psi\rangle$ in the vector space V with inner product we have*

$$\langle\varphi|\psi\rangle\langle\psi|\varphi\rangle \leq \langle\varphi|\varphi\rangle\langle\psi|\psi\rangle \tag{2.51}$$

Proof: We may assume without loss of generality that the vector $|\psi\rangle$ is normalised i.e. $\langle\psi|\psi\rangle = 1$; otherwise we divide left and right-hand side of the inequality by the real, positive number $\langle\psi|\psi\rangle$. We need to show that

$$\langle\varphi|\psi\rangle\langle\psi|\varphi\rangle \leq \langle\varphi|\varphi\rangle. \tag{2.52}$$

To see this, complete $|\psi\rangle$ to an orthonormal basis $B = \{|\psi\rangle, |b_2\rangle, \ldots, |b_n\rangle\}$ and write the identity as

$$I = |\psi\rangle\langle\psi| + \sum_{i=2}^{n} |b_i\rangle\langle b_i|. \tag{2.53}$$

Now consider the inner product $\langle\varphi|\varphi\rangle$ and insert the identity:

$$\begin{aligned}\langle\varphi|\varphi\rangle &= \langle\varphi|I|\varphi\rangle = \langle\varphi|\psi\rangle\langle\psi|\varphi\rangle + \sum_{i=2}^{n}\langle\varphi|b_i\rangle\langle b_i|\varphi\rangle \\ &\geq \langle\varphi|\psi\rangle\langle\psi|\varphi\rangle\end{aligned} \tag{2.54}$$

where we used that $\langle\varphi|b_i\rangle\langle b_i|\varphi\rangle = \langle\varphi|b_i\rangle\overline{\langle\varphi|b_i\rangle} = |\langle\varphi|b_i\rangle|^2 \geq 0$. $\square$

2.4 Hermitian and Unitary operators, Projectors

Having defined inner product spaces, we now consider operators in such spaces in some detail. We begin with the fundamental

Definition 2.4.1 *(***Adjoint operator***) Let A be a linear operator in a complex vector space V with inner product $(\cdot,\cdot)$. Then we define the adjoint operator $A^\dagger$ by the condition*

$$(|\varphi\,\rangle, A|\psi\,\rangle) = (A^\dagger|\varphi\,\rangle, |\psi\,\rangle) \quad \textit{for all} \quad |\varphi\,\rangle, |\psi\,\rangle \in V \tag{2.55}$$

or, using bra-ket notation,

$$\langle\,\varphi|A|\psi\,\rangle = \overline{\langle\,\psi|A^\dagger|\varphi\,\rangle}. \tag{2.56}$$

Let $B = \{|b_1\,\rangle, \ldots, |b_n\,\rangle\}$ be an orthonormal basis of V and A_{ij} be the matrix elements of the matrix representation of A i.e.

$$\langle\,b_i|A|b_j\,\rangle = A_{ij} \tag{2.57}$$

Then, we can read off the matrix representation of $A^\dagger$ with respect to the same basis from (2.56):

$$\langle\,b_i|A^\dagger|b_j\,\rangle = \overline{\langle\,b_j|A|b_i\,\rangle} = \bar{A}_{ji}. \tag{2.58}$$

Thus the matrix representing $A^\dagger$ is obtained from the matrix representing A by transposition and complex conjugation. Using the same symbols for the matrices as for the operators which they represent, we write

$$A^\dagger = \bar{A}^t. \tag{2.59}$$

Example 2.4.2 *The matrix representing the operator $A : \mathbb{C}^2 \to \mathbb{C}^2$ relative to a fixed orthonormal basis of $\mathbb{C}^2$ is*

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

Find the matrix representing the adjoint $A^\dagger$.

Transposing and complex conjugating we obtain

$$A^\dagger = \begin{pmatrix} 2+i & 1+i \\ 3-2i & 1-i \end{pmatrix}.$$

We note the following general properties of adjoints:

Lemma 2.4.3 *Let A and B be linear operators in a vector space V with inner product and $\alpha, \beta \in \mathbb{C}$. Then*

1. $(A^\dagger)^\dagger = A$
2. $(\alpha A + \beta B)^\dagger = \bar{\alpha} A^\dagger + \bar{\beta} B^\dagger$
3. $(AB)^\dagger = B^\dagger A^\dagger$

The proof is straightforward - and left as an exercise.

Example 2.4.4 *Let $B = \{|b_1\rangle, \ldots, |b_n\rangle\}$ be an orthonormal basis of the inner product space V. Find the adjoint of the map*

$$|b_i\rangle\langle b_j| : V \to V \qquad |x\rangle \mapsto |b_i\rangle\langle b_j|x\rangle$$

considered in (2.47)

For arbitrary elements $|\varphi\rangle, |\psi\rangle \in V$ we have

$$\begin{aligned} (|\varphi\rangle, |b_i\rangle\langle b_j|\psi\rangle) &= \langle \varphi, |b_i\rangle\langle b_j, |\psi\rangle \\ &= \overline{\langle b_i|\varphi\rangle}\langle b_j, |\psi\rangle \\ &= (|b_j\rangle\langle b_i|\varphi\rangle, |\psi\rangle) \end{aligned} \tag{2.60}$$

Comparing with the definition (2.55) we conclude

$$(|b_i\rangle\langle b_j|)^\dagger = |b_j\rangle\langle b_i|. \tag{2.61}$$

□

One can extend the definition of an adjoint to maps $A : V \to W$, where V and W are two different inner product spaces. In that case $A^\dagger$ is a map $W \to V$. The matrix representation of $A^\dagger$ is still obtained from the matrix representation of A be transposition (turning rows into columns) and complex conjugation. We will not need this definition in full generality, but note the special case where $W = \mathbb{C}$. We saw in example 2.3.4 that any bra $\langle x|$, thought of as a map $V \to \mathbb{C}$, can be represented by the row vector $(\bar{x}_1, \ldots, \bar{x}_2)$ with respect to a basis $\{|b_1\rangle, \ldots, |b_2\rangle\}$ of V. The transposition and complex conjugation of this row vector gives

$$\overline{(\bar{x}_1, \bar{x}_2 \ldots, \bar{x}_n)^t} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{pmatrix}, \tag{2.62}$$

which is just the coordinate representation of $|x\rangle$. It is therefore consistent to extend our definition of the adjoint to

$$\langle x|^\dagger = |x\rangle \tag{2.63}$$

so that, by Lemma 2.4.3

$$|x\rangle^\dagger = \langle x|. \tag{2.64}$$

Note that these facts, together with the second part of Lemma 2.4.3 gives a quick proof of (2.61).

The two classes of linear operators which are important in quantum mechanics are defined by relations between the operator and its adjoint.

Definition 2.4.5 *(***Unitary operators** *) Let V be a vector space with inner product and $U : V \to V$ be a linear operator. We say that U is unitary if*

$$U^\dagger = U^{-1}. \tag{2.65}$$

An important property of unitary operators is that they preserve the inner product

Lemma 2.4.6 *If U is a unitary operator in the vector space V with inner product $\langle \cdot|\cdot\rangle$ then*

$$(U|\varphi\rangle, U|\psi\rangle) = (|\varphi\rangle, |\psi\rangle) = \langle \varphi|\psi\rangle. \tag{2.66}$$

This follows directly from the definition of the adjoint and the definition of a unitary operator:

$$(U|\varphi\rangle, U|\psi\rangle) = (U^\dagger U|\varphi\rangle, |\psi\rangle) = (|\varphi\rangle, |\psi\rangle). \tag{2.67}$$

$\Box$.

Specialise now to the case $V = \mathbb{C}^n$ with the canonical inner product (2.41) and the canonical orthonormal basis (2.5). Identifying, as before, the matrix representation of $U : \mathbb{C}^n \to \mathbb{C}^n$ relative to the canonical basis (2.5) with U, we can write the condition for unitarity in matrix form as

$$\bar{U}^t U = I. \tag{2.68}$$

Example 2.4.7 *Show that the matrix*

$$U = \begin{pmatrix} e^{i\phi}\cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & e^{-i\phi}\cos(\frac{\theta}{2}) \end{pmatrix}$$

is unitary for $\theta \in (0, 2\pi)$ and $\phi \in [0, 2\pi)$

Using $\cos^2(\frac{\theta}{2}) + \sin^2(\frac{\theta}{2}) = 1$ we find

$$\bar{U}^t U = \begin{pmatrix} e^{-i\phi}\cos(\frac{\theta}{2}) & \sin(\frac{\theta}{2}) \\ -\sin(\frac{\theta}{2}) & e^{i\phi}\cos(\frac{\theta}{2}) \end{pmatrix} \begin{pmatrix} e^{i\phi}\cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & e^{-i\phi}\cos(\frac{\theta}{2}) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$\square$.

Definition 2.4.8 *(**Hermitian operators**) Let V be a vector space with inner product and $A : V \to V$ be a linear operator. We say that A is Hermitian if*

$$A^\dagger = A \tag{2.69}$$

Example 2.4.9 *(**Pauli matrices** The following three matrices are called the Pauli matrices*

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.70}$$

Show that they are both Hermitian and unitary.

This is an elementary calculation $\square$

We collect some properties of unitary and Hermitian operators in the following lemma.

Lemma 2.4.10 *In the following V is a complex vector space with inner product. Then*

1. *$A^\dagger A$ is Hermitian for any operator $A : V \to V$.*
2. *If $B : V \to V$ is Hermitian, then so is B^{-1}.*
3. *If $W : V \to V$ is unitary, then so is W^{-1}.*
4. *If B is Hermitian and U unitary, then $U^{-1}BU$ is Hermitian*
5. *If W and U are unitary, then $U^{-1}WU$ is unitary*

Proof:

1. Applying the rules 1 and 3 in Lemma 2.4.3, we have $(A^\dagger A)^\dagger = A^\dagger (A^\dagger)^\dagger = A^\dagger A$, thus establishing the first claim.

2. Taking the adjoint of the equation $B^{-1}B = I$ we find $B^\dagger (B^{-1})^\dagger = I$ since the identity is Hermitian. Now use Hermiticity of B to deduce $B(B^{-1})^\dagger = I$ so that $(B^{-1})^\dagger = B^{-1}$, establishing the Hermiticity of B^{-1}.

3. Taking the adjoint of the equation $W^{-1} = W^\dagger$ we find $(W^{-1})^\dagger = W$. Hence $(W^{-1})^\dagger = (W^{-1})^{-1}$, showing the W^{-1} is unitary.
4. $(U^{-1}BU)^\dagger = U^\dagger(U^{-1}B)^\dagger = U^\dagger B^\dagger (U^{-1})^\dagger = U^{-1}BU$.
5. $(U^{-1}WU)^\dagger = U^\dagger W^\dagger (U^{-1})^\dagger = U^{-1}W^{-1}U = (U^{-1}WU)^{-1}$ $\square$.

Finally we turn to a class of operators called projectors or projection operators

Definition 2.4.11 *(***Projection operator***) An operator $P : V \to V$ is called projection operator if $P^2 = P$. If V is equipped with an inner product and P is Hermitian with respect to that inner product, P is called an orthogonal projection operator*

Example 2.4.12 *Consider the vector space $\mathbb{R}^2$ ("the xy-plane") with its canonical inner product and canonical basis $|0\rangle, |1\rangle$. Write down the matrix representation, with respect to the canonical basis, of*

1. *The projection along the y-axis onto the x-axis.*
2. *The projection along the line $x + y = 0$ onto the x-axis*

You can visualise the examples in terms of shining light along the y-axis for 1. and along the line $x + y = 0$ for 2. Working out the projection operator is equivalent to determining the shadow cast on the x-axis. Which of the projections is (are) orthogonal?

In order to determine any linear map, it is enough to determine its action on a basis. In the first example we have

$$P|0\rangle = |0\rangle, \qquad P|1\rangle = 0.$$

Hence the matrix representing P is

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

In the second example we have

$$P|0\rangle = |0\rangle, \qquad P|1\rangle = |0\rangle$$

leading to the matrix representation

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

It is clear geometrically that the first projection operator is orthogonal and the second is not. This is also reflected in the matrix representation: the first projection is represented by a Hermitian matrix, but the matrix representing the second projection is not Hermitian.

Generally, given an m-dimensional subspace W of an inner product space V, we can construct an orthogonal project operator onto W by picking an orthonormal basis $\{|b_1\rangle, \ldots |b_m\rangle\}$ of W. We claim

Lemma 2.4.13 *The operator P_W defined via*

$$P_W = \sum_{i=1}^{m} |b_i\,\rangle\langle\, b_i|. \tag{2.71}$$

is an orthogonal projection operator

Proof: In order to check that P_W is a projection we compute

$$\begin{aligned} P_W^2 &= \sum_{i=1}^{m} |b_i\,\rangle\langle\, b_i| \sum_{j=1}^{m} |b_j\,\rangle\langle\, b_j| \\ &= \sum_{i,j=1}^{m} |b_i\,\rangle\langle b_i|b_j\,\rangle\langle\, b_j| \\ &= \sum_{i,j=1}^{m} \delta_{ij}|b_i\,\rangle\langle\, b_j| \\ &= \sum_{i}^{m} |b_i\,\rangle\langle\, b_i| = P_W \end{aligned} \tag{2.72}$$

The orthogonality

$$P_W^\dagger = P_W \tag{2.73}$$

follows from $(|b_i\,\rangle\langle\, b_i|)^\dagger = |b_i\,\rangle\langle\, b_i|$, which is a special case of (2.61). $\square$

Note that if P is a projection operator, then so is $I-P$ since $(I-P)^2 = I-2P+P = I-P$. Similarly, if P is an orthogonal projection operator, then so is $I-P$. Geometrically, if P is the orthogonal projection onto a subspace W, then $I-P$ is the projection onto the orthogonal complement $W^\perp$ defined in (2.3.8) .

2.5 Eigenvalues and commutators

An important part of solving problems in quantum mechanics involves finding eigenvalues and eigenvectors of linear operators. Recall that if $A : V \to V$ is a linear operator, we call $\lambda \in \mathbb{C}$ an eigenvalue of A if there exists a non-zero vector $|v\,\rangle \in V$ such that

$$A|v\,\rangle = \lambda|v\,\rangle. \tag{2.74}$$

Any such vector $|v\,\rangle$ is called an eigenvector of A with eigenvalue λ. More generally, there may be several linearly dependent eigenvectors for a given eigenvalue λ. The space of all eigenvectors is called the **eigenspace** for the eigenvalue λ and denoted

$$\mathrm{Eig}_\lambda = \{|v\,\rangle \in V | A|v\,\rangle = \lambda|v\,\rangle\}. \tag{2.75}$$

It is not difficult to check that Eig_λ is indeed a vector space (do it!)

The eigenvalues of A are most easily determined by solving the characteristic equation

$$\det(A - \lambda I) = 0. \tag{2.76}$$

This is a polynomial equation in λ of degree $n = \dim V$. By the fundamental theorem of algebra such an equation has at least one solution ("root") in the complex numbers, and this fact considerably simplifies the eigenvalue problem in complex vector spaces compared to real vector spaces. It follows that every operator in a complex vector spaces has at least one eigenvalue. For some operators one can find an entire basis of V consisting of eigenvectors. Such operators are called **diagonalisable**. Remarkably, the Hermitian and unitary operators which are important in quantum mechanics are always diagonalisable. The key reason for their diagonalisability lies in the following

Lemma 2.5.1 *Suppose $|v\rangle \in V$ is an eigenvector of the Hermitian operator A with eigenvalue λ. Then A maps the orthogonal complement of $[|v\rangle]$ into itself, i.e. if $|w\rangle \perp |v\rangle$ then also $A|w\rangle \perp |v\rangle$.*

Proof: Suppose $\langle v|w\rangle = 0$. Then $\langle v|A|w\rangle = \overline{\langle w|A|v\rangle} = \bar{\lambda}\langle v|w\rangle = 0$ $\square$.

Theorem 2.5.2 *Suppose V is a (complex) vector space with inner product. If $A : V \to V$ is a Hermitian operator, all eigenvalues are real and eigenvectors for different eigenvalues are necessarily orthogonal. Moreover, there exists an orthonormal basis of eigenvectors of A.*

Proof: To see that any eigenvalue of a Hermitian operator has to be real, suppose λ is an eigenvalue of the Hermitian operator A, with associated eigenvector $|v\rangle$, which we assume to be normalised. Then

$$\langle v|A|v\rangle = \lambda. \tag{2.77}$$

On the other hand

$$\langle v|A|v\rangle = \overline{\langle v|A^\dagger|v\rangle} = \overline{\langle v|A|v\rangle} = \bar{\lambda}. \tag{2.78}$$

Comparing (2.77) with (2.78) we conclude that

$$\bar{\lambda} = \lambda \tag{2.79}$$

so that λ is real. Now suppose that $|v_1\rangle$ and $|v_2\rangle$ are eigenvectors associated to distinct eigenvalues λ_1 and λ_2. Then $\langle v_1|A|v_2\rangle = \lambda_2\langle v_1|v_2\rangle$ but also, by Hermiticity, $\langle v_1|A|v_2\rangle = \lambda_1\langle v_1|v_2\rangle$. Hence $(\lambda_1 - \lambda_2)\langle v_1|v_2\rangle = 0$. Since $\lambda_1 \neq \lambda_2$ this implies $\langle v_1|v_2\rangle = 0$.

In order to prove the existence of an orthonormal basis we proceed by induction over the dimension of V. If the dimension is 1 there is nothing to prove. Suppose we have proved the theorem for vector spaces of dimension $n-1$, and let V be a vector space of dimension n. A is a Hermitian operator in V and has at least one eigenvalue with eigenspace W. Pick one eigenvector $|v\rangle$ and consider the orthogonal complement $[|v\rangle]^\perp$. It has dimension $n-1$ and by Lemma 2.5.1 is mapped into itself by A. Hence the restriction of A to $[|v\rangle]^\perp$ is a Hermitian operator in a vector space of dimension $n-1$. By the induction assumption it is diagonalisable and has an orthonormal basis $\{|v_1\rangle, \ldots, |v_{n-1}\rangle\}$ of eigenvectors. Then $B = \{|v\rangle, |v_1\rangle, \ldots, |v_{n-1}\rangle\}$ is an orthonormal basis of eigenvectors for A.

□

We can rephrase the results of this theorem by collecting all eigenvectors which have the same eigenvalue into eigenspaces, thus obtaining the following

Corollary 2.5.3 *Suppose V is a n-dimensional (complex) vector space with inner product. and $A : V \to V$ is a Hermitian operator with $m \leq n$ distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. Then there is a unique decomposition of V into mutually orthogonal eigenspaces of V, i.e.*

$$V = Eig_{\lambda_1} \oplus \ldots \oplus Eig_{\lambda_m} \tag{2.80}$$

Example 2.5.4 *A Hermitian operator $A : \mathbb{C}^2 \to \mathbb{C}^2$ has the matrix representation*

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.81}$$

with respect to the canonical basis $\{|0\rangle, |1\rangle\}$. Find the eigenvalues λ_1 and λ_2 and corresponding orthonormal eigenvectors $|v_1\rangle, |v_2\rangle$ of A. Give the matrix representation A' of A relative to the basis $\{|v_1\rangle, |v_2\rangle\}$ and find the 2×2 matrix M so that

$$A' = M^{-1} A M \tag{2.82}$$

The characteristic equation

$$\det(A - \lambda) = 0 \Leftrightarrow \lambda^2 - 1 = 0$$

has solutions $\lambda_1 = 1$ and $\lambda_2 = -1$. To find an eigenvector $\begin{pmatrix} x \\ y \end{pmatrix}$ for the eigenvalue -1 we need to solve

$$y = x, \quad x = y,$$

yielding the (normalised) eigenvector

$$|v_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Similarly one finds the eigenvector for the eigenvalue $\lambda_2 = -1$ to be

$$|v_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Thus, the matrix representation of A relative to the basis $\{|v_1\rangle, |v_2\rangle\}$ is

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

We read off the transformation matrix M from the expansion

$$|v_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |v_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

according to (2.21) and find

$$M = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

it is now easy to verify that (2.82) holds. □.

Theorem 2.5.5 *Suppose V is a (complex) vector space with inner product. If $U : V \to V$ is a unitary operator, there exists an orthonormal basis of eigenvectors of U. Moreover, all eigenvalues λ of U have modulus 1, i.e. can be written in the form $e^{i\alpha}$ for some $\alpha \in [0, 2\pi)$. Eigenvectors corresponding to different eigenvalues are necessarily orthogonal.*

We will not prove this here, since the proof is analogous to the that of the corresponding statement for Hermitian operators. We only show that any eigenvalue of a unitary operator has to have modulus 1. Suppose λ is an eigenvalue of the unitary operator U, with associated normalised eigenvector $|v\rangle$. Then

$$\langle v|U|v\rangle = \lambda. \tag{2.83}$$

On the other hand

$$\langle v|U|v\rangle = \overline{\langle v|U^\dagger|v\rangle} = \overline{\langle v|U^{-1}|v\rangle} = \frac{1}{\bar{\lambda}}. \tag{2.84}$$

Comparing (2.83) with (2.84) we conclude that

$$\bar{\lambda}\lambda = 1 \tag{2.85}$$

so that $|\lambda| = 1$. $\square$.

Example 2.5.6 *Find the eigenvalues and normalised eigenvectors of the unitary matrix*

$$A = \begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix} \tag{2.86}$$

The method is as for example 2.5.4. This time we find eigenvalues $\lambda_1 = e^{i\gamma}$ and $\lambda_2 = e^{-i\gamma}$ with eigenvectors

$$|v_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} \qquad |v_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

$\square$.

Example 2.5.7 *Show that any eigenvalue of a projection operator is either* 0 *or* 1.

Suppose λ is an eigenvalues of a projection operator P i.e., there exists a non-zero $|v\rangle$ so that

$$P|v\rangle = \lambda|v\rangle$$

Applying P again to both sides of the equation and using $P^2 = P$ we find

$$\lambda|v\rangle = \lambda^2|v\rangle$$

Since $|v\rangle$ is non-zero by assumption we have $\lambda = \lambda^2$ which is solved by $\lambda = 0$ and $\lambda = 1$. $\square$

In quantum mechanics it is often necessary to consider several operators and to find a basis of eigenvectors for both. It is not always possible to find such a basis, even if each of the operators is diagonalisable. However, there is a simple test for simultaneous diagonalisation. In order to state it succinctly, we define

Definition 2.5.8 (**Commutator**) *The commutator of two operators $A, B : V \to V$ is defined as*

$$[A, B] = AB - BA \tag{2.87}$$

Theorem 2.5.9 *Let A, B be two Hermitian or unitary operators in a vector space V. Then A and B can be diagonalised simultaneously if and only if their commutator vanishes i.e. if $[A, B] = 0$.*

Proof: In the proof we assume for definiteness that A and B are Hermitian. The proof for unitary operators is analogous.

Suppose there is a basis with respect to which both A and B are both diagonal, say

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & & \\ & & & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \qquad B = \begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \vdots & & & \\ & & & \\ 0 & 0 & \dots & \mu_n \end{pmatrix}, \tag{2.88}$$

then clearly $AB = BA$, so the commutator of A and B vanishes.

Now suppose that the commutator $[A, B]$ is zero. The operator A, being Hermitian, can be diagonalised, producing the decomposition of V into $m \le n$ eigenspaces given in corollary 2.5.3:

$$V = \mathrm{Eig}_{\lambda_1} \oplus \ldots \oplus \mathrm{Eig}_{\lambda_m} \tag{2.89}$$

Now pick one of the eigenvalues λ_i and let $|v\,\rangle$ be in the eigenspace Eig_{λ_i}. Then

$$A(B|v\,\rangle) = BA|v\,\rangle = \lambda_i(B|v\,\rangle)$$

so that $B|v\,\rangle \in \mathrm{Eig}_{\lambda_i}$ for all $|v\,\rangle \in \mathrm{Eig}_{\lambda_i}$. Hence we can restrict B to Eig_{λ_i} and obtain a Hermitian operator

$$B_{|\mathrm{Eig}_{\lambda_i}} : \mathrm{Eig}_{\lambda_i} \to \mathrm{Eig}_{\lambda_i}.$$

Since this operator is Hermitian, there exists an orthonormal basis B_i of eigenvectors which are eigenvectors of A by construction. Repeating this process for every eigenvalue λ_i of A we obtain the basis

$$\bigcup_{i=1}^{m} B_i \tag{2.90}$$

consisting of simultaneous eigenvectors of A and B. $\square$.

3 Quantum Mechanics

3.1 General remarks: the postulates of quantum mechanics

In this section we state the basic postulates of quantum mechanics and illustrate with simple examples. The postulates summarise how physics is mathematically described in quantum mechanics. Like all good theories of physics, quantum mechanics allows one to make predictions about the outcomes of physical experiments. However, unlike the laws of classical physics, which predict outcomes with certainty, quantum mechanics only singles out the possible outcomes and predicts the probabilities with which they happen.

The quantum mechanical postulates emerged as a succinct summary of the quantum mechanical rules in the second half of the 1920's. In contrast to other famous physical laws, for example Newton's laws in classical mechanics, they were not historically written down in definitive form by one person. Instead they emerged from research activity lasting several years and involving many physicists. As a result there is not one definitive version of the postulates. Different books give slightly different versions - even the number and numbering of the postulates is not standardised.

Inner product spaces play a key role in quantum mechanics, and for many applications of quantum mechanics it is essential to consider infinite-dimensional vector spaces. We do not need infinite dimensional vector spaces in this course, but nonetheless use notation and names which are customary in the infinite dimensional context. An example of such terminology is the word "linear operator" for linear maps. Another, very important term is "Hilbert space" to describe an inner product space which is complete with respect to the norm derived from the inner product. In finite dimensions all inner product spaces are complete, i.e. Hilbert spaces and inner product spaces are the same thing in finite dimensions.

In this course all Hilbert spaces are assumed to be finite-dimensional!

3.2 States

The first postulate says how we describe the state of a physical system mathematically in quantum mechanics.

Postulate 1: State space

Associated to every isolated physical system is a complex vector space V with inner product (Hilbert space) called the state space of the system. At any give time the physical state of the system is completely described by a state vector, which is a vector $|v\rangle$ in V with norm 1.

Example 3.2.1 *The vector $|v\rangle = \frac{i}{\sqrt{2}}(-|0\rangle + |1\rangle)$ describes a state of the system with Hilbert space $\mathbb{C}^2$. We say that it is a superposition of the state vectors $|0\rangle$ and $|1\rangle$, and the coefficients $-\frac{i}{\sqrt{2}}$ and $\frac{i}{\sqrt{2}}$ are sometimes called* amplitudes.

Note that, while the state of the system is completely characterised by giving a state vector, the postulate leaves open the possibility that different unit vectors may describe the same state. In fact we shall see that in calculations of physical quantities it does not matter if we use the state vector $|v\rangle$ or $|v'\rangle = e^{i\alpha}|v\rangle$, $\alpha \in [0, 2\pi)$. The state vectors $|v\rangle$ and $|v'\rangle$ may thus be regarded as equivalent descriptions of the same physical state. There is a mathematical formulation (using "projective Hilbert space") which takes this equivalence into account, but it is a little more complicated to handle, and we will not use it in this course. Strictly speaking we should therefore distinguish between a state of a system and the state vector used to describe this. However, since the phrase "the state vector describing the state.." is much longer than "the state .." we shall often use the latter as a shorthand.

3.3 Observables and measurement

The second postulate deals with possible outcomes of measurements and specifies how to compute their probabilities.

Postulate 2: Observables and measurements

The physically observable quantities of a physical system, also called the observables, are mathematically described by Hermitian operators acting on the state space V of the system. The possible outcomes of measurements of an observable A are given by the eigenvalues $\lambda_1, \ldots \lambda_m$ of A. If the system is in the state with state vector $|\psi\rangle$ at the time of the measurement, the probability of obtaining the outcome λ_i is

$$p_\psi(\lambda_i) = \langle \psi|P_i|\psi\rangle, \tag{3.1}$$

where P_i is the orthogonal projection operator on the eigenspace of λ_i. Given that this outcome occurred, the state of the system immediately after the measurement is described by

$$|\tilde{\psi}\rangle = \frac{P_i|\psi\rangle}{\sqrt{p_\psi(\lambda_i)}}. \tag{3.2}$$

(This is sometimes called the collapse of the wavefunction)

We need to check that the prescriptions given in Postulate 2 make sense:

1. Do the the numbers (3.1) lie between 0 and 1 and add up to 1, so that they can indeed be interpreted as probabilities?

2. Is (3.2) really a state vector, i.e. does it have norm 1?

We postpone the discussion of both these question until a little later in this section. In order to build up an understanding of the second postulate we first apply it in the following example.

Example 3.3.1 *Consider a system with Hilbert space $V = \mathbb{C}^3$, equipped with the canonical inner product. The system is in the state described by $|\psi\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ when the observable*

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{3.3}$$

is measured. Show that the possible outcomes of the measurement are 0 *and* 2 *and compute the probablity of each. For each of the possible outcomes, give the state of the system immediately after the measurement.*

From the characteristic equation $\det(A - \lambda I) = 0$ we find

$$(1-\lambda)^2(2-\lambda) - (2-\lambda) = 0 \Leftrightarrow (2-\lambda)(\lambda^2 - 2\lambda) = 0$$

which has solutions $\lambda_1 = 0$ and $\lambda_2 = 2$. The normalised eigenvector with eigenvalue $\lambda_1 = 0$ is

$$|b_{1,1}\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix} \tag{3.4}$$

but the eigenvalue $\lambda_2 = 2$ has a two dimensional eigenspace with orthonormal basis given by

$$|b_{2,1}\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}, \quad |b_{2,2}\rangle = \begin{pmatrix}0\\0\\1\end{pmatrix}. \tag{3.5}$$

Hence the projectors onto the eigenspaces are

$$P_1 = |b_{1,1}\rangle\langle b_{1,1}| \quad \text{and} \quad P_1 = |b_{2,1}\rangle\langle b_{2,1}| + |b_{2,2}\rangle\langle b_{2,2}|. \tag{3.6}$$

The probability of measuring $\lambda_1 = 0$ is

$$p_\psi(0) = \langle \psi|P_1|\psi\rangle = \langle\psi|b_{1,1}\rangle\langle b_{1,1}|\psi\rangle = |\langle b_{1,1}|\psi\rangle|^2 = \frac{1}{2} \tag{3.7}$$

and the probability of measuring $\lambda_2 = 2$ is

$$p_\psi(2) = \langle \psi|P_2|\psi\rangle = \langle\psi|b_{2,1}\rangle\langle b_{2,1}|\psi\rangle + \langle\psi|b_{2,2}\rangle\langle b_{2,2}|\psi\rangle = |\langle b_{2,1}|\psi\rangle|^2 + |\langle b_{2,2}|\psi\rangle|^2 = \frac{1}{2} \tag{3.8}$$

If the measurement produces the result $\lambda_1 = 0$, the state after the measurement is

$$|\varphi\rangle = \frac{P_1|\psi\rangle}{\sqrt{p_\psi(\lambda_1)}} = \sqrt{2} \times \langle b_{1,1}|\psi\rangle|b_{1,1}\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix} \tag{3.9}$$

If the measurement produces the result $\lambda_2 = 2$, the state after the measurement is

$$|\varphi\rangle = \frac{P_2|\psi\rangle}{\sqrt{p_\psi(\lambda_1)}} = \sqrt{2} \times (\langle b_{2,1}|\psi\rangle|b_{2,1}\rangle + \langle b_{2,2}|\psi\rangle|b_{2,2}\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix} \tag{3.10}$$

$\square$.

Note that the projection operators only play an intermediate role in the calculation. They are useful in stating the measurement postulate, but in specific calculations we can go straight from the calculation of the eigenvalues and eigenfunctions to the evaluation of probabilities and final states. In particular, note that the state of the system after the measurement of the non-degenerate eigenvalue $\lambda_1 = 0$ is the eigenstate $|b_{1,1}\rangle$ associated to that eigenvalue. This fact generalises: **if a measurement outcome is an eigenvalue with one-dimensional eigenspace spanned by the normalised eigenvector $|v\rangle$, the state of the system after the measurement is given by $|v\rangle$.**

Example 3.3.2 *(***"Measurement of a state"***) Consider the single qubit system with Hilbert space $\mathbb{C}^2$. Consider the orthogonal projection operators associated to the canonical basis states*

$$P = |0\rangle\langle 0|, \qquad Q = |1\rangle\langle 1| \tag{3.11}$$

If the system is in the state $|\psi\rangle = \frac{1}{2}(\sqrt{3}|0\rangle + |1\rangle)$, what is the probability of obtaining the eigenvalue 1 in a measurement of P. What is the probability of obtaining the eigenvalue 0? What is the probability of obtaining the eigenvalue 0 in a measurement of Q?

The projection operator P has the eigenstate $|0\rangle$ with eigenvalue 1 and the eigenstate $|1\rangle$ with eigenvalue 0. For Q the situation is the reverse: $|0\rangle$ is eigenstate with eigenvalue 0 and $|1\rangle$ is eigenstate with eigenvalue 1. Hence the probability of measuring 1 in a measurement of P is $|\langle\psi|0\rangle|^2 = \frac{3}{4}$. The probability of measuring 0 in a measurement of P is $|\langle\psi|1\rangle|^2 = \frac{1}{4}$. The probability of measuring 0 in a measurement of Q is $|\langle\psi|0\rangle|^2 = \frac{3}{4}$. □

The example shows that measuring projection operators $|\varphi\rangle\langle\varphi|$ associated to states $|\varphi\rangle$ amounts to asking for the probability of the system to be in the state $|\varphi\rangle$. It is therefore common practice in discussions of quantum mechanical systems to replace the long question "What is the probability of obtaining the eigenvalue 1 in a measurement of the projection operator $|\varphi\rangle\langle\varphi|$ given that the system is in the state $|\psi\rangle$?" with the shorter question "what is the probability of finding the system in the state $|\varphi\rangle$, given that it is in the state $|\psi\rangle$? ". As we have seen, the answer to that question is

$$|\langle\varphi|\psi\rangle|^2 \tag{3.12}$$

The complex number $\langle\varphi|\psi\rangle$ is often called the **overlap of the states $|\varphi\rangle$ and $|\psi\rangle$**. Note that the probablity (3.12) can be non-zero even when the system's state $|\psi\rangle$ is different from $|\varphi\rangle$. It is zero if and only if $|\varphi\rangle$ and $|\psi\rangle$ are orthogonal.

We have yet to prove that the probabilities defined in (3.1) can consistently be interpreted as probabilities. To show this we need the following lemma, which will be useful in other applications as well.

Lemma 3.3.3 *V is a Hilbert space and A a Hermitian operator in V with eigenvalues λ_i, $i = 1, \ldots, m$ and eigenspaces Eig_{λ_i}. Let P_i be the orthogonal projector onto Eig_{λ_i}. Then*

1. *The* **orthogonality relations**

$$P_i P_j = \delta_{ij} P_i \tag{3.13}$$

   hold.

2. *The* **completeness relations**

$$\sum_{i=1}^{m} P_i = I \tag{3.14}$$

   hold.

3. **Spectral decomposition of** A: *we can write* A *in terms of the orthogonal projection operators* P_i *onto the eigenspaces* Eig_{λ_i} *as*

$$A = \sum_{i=1}^{m} \lambda_i P_i \tag{3.15}$$

Proof: 1. If $i = j$, the claim reduces to $P_i^2 = P_i$, which is the defining property of any projection operator. If $i \neq j$ we need to show that $P_i P_j = 0$. To show this, consider arbitrary states $|\varphi\rangle, |\psi\rangle \in V$. Then, by the definition of the projection operators P_i, $P_i|\psi\rangle \in \mathrm{Eig}_{\lambda_i}$. Since Eig_{λ_i} and Eig_{λ_j} are orthogonal for $i \neq j$, we conclude

$$0 = (P_i|\varphi\rangle, P_j|\psi\rangle) = \langle \varphi|P_i P_j|\psi\rangle.$$

However, if the matrix element $\langle \varphi|P_i P_j|\psi\rangle$ vanishes for all $|\varphi\rangle, |\psi\rangle \in V$, then we have the operator identity $P_i P_j = 0$.

2. Suppose the dimension of Eig_{λ_i} is k_i and $B^i = \{|b_{i,1}\rangle, \ldots, |b_{i,k_i}\rangle$ is an orthonormal basis of Eig_{λ_i} so that $B = \cup_{i=1}^m B^i$ is an orthonormal basis of eigenvectors of A. Then

$$P_i = \sum_{l=1}^{k_i} |b_{i,l}\rangle\langle b_{i,l}| \tag{3.16}$$

and hence

$$\sum_{i=1}^{m} P_i = \sum_{i=1}^{m}\sum_{l=1}^{k_i} |b_{i,l}\rangle\langle b_{i,l}| = I \tag{3.17}$$

by the general formula (3.14) for the identity in terms of an orthonormal basis.

3. To show the equality of operators (3.15) we show their equality when acting on a basis of V. Using

$$P_i|b_{j,l}\rangle = \delta_{ij}|b_{j,l}\rangle, \quad l = 1, \ldots, k_j \tag{3.18}$$

we have

$$\sum_{i=1}^{m} \lambda_i P_i|b_{j,l}\rangle = \lambda_j|b_{j,l}\rangle \tag{3.19}$$

which agrees with the action of A on $|b_{k,j}\rangle$, as was to be shown. $\square$.

Before we study examples we note

Corollary 3.3.4 *With the assumptions of the previous theorem*

$$(\sum_{i=1}^{m} \lambda_i P_i)^n = \sum_{i=1}^{m} \lambda_i^n P_i \tag{3.20}$$

Proof: We prove the corollary by induction. Clearly the claim holds for $n = 1$. Suppose it holds for $n-1$ i.e.

$$(\sum_{i=1}^{m} \lambda_i P_i)^{n-1} = \sum_{i=1}^{m} \lambda_i^{n-1} P_i \tag{3.21}$$

Using this identity, and applying (3.13) and (3.14) we compute

$$\begin{aligned}
(\sum_{i=1}^{m} \lambda_i P_i)^{n} &= (\sum_{i=1}^{m} \lambda_i P_i)(\sum_{j=1}^{m} \lambda_i P_i)^{n-1} \\
&= (\sum_{i=1}^{m} \lambda_i P_i)(\sum_{j=1}^{m} \lambda_j^{n-1} P_j) \\
&= \sum_{i,j=1}^{m} \lambda_i \lambda_j^{n-1} P_i P_j \\
&= \sum_{i=1}^{m} \lambda_i^{n} P_i
\end{aligned} \tag{3.22}$$

as was to be shown. □

Example 3.3.5 *Consider again the Hermitian operator studied in example 2.5.4, whose matrix representation relative to the canonical basis of* $\mathbb{C}^2$ *is*

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.23}$$

Using the results of 2.5.4 write A in the form (3.15)*.*

The eigenspaces for the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$ are both one dimensional, and the projectors onto these eigenspaces can be written in terms of the eigenvectors found in example 2.5.4:

$$P_1 = |v_1\rangle\langle v_1|, \qquad P_2 = |v_2\rangle\langle v_2|$$

Hence (3.15) takes the form

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

It is instructive to check that this reproduces the matrix (3.23) when we insert the coordinates of the eigenvectors $|v_1\rangle$ and $|v_2\rangle$ relative to the canonical basis

$$P_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

and

$$P_2 = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\begin{pmatrix} 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

so that

$$P_1 - P_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

as required. □

We now come to the promised proof that the quantities $p_\psi(\lambda_i)$ defined in Postulate 2 can consistently be interpreted as probabilities.

Lemma 3.3.6 *The probabilities defined in* (3.1) *satisfy*

1. $0 \leq p_\psi(\lambda_i) \leq 1$
2. $\sum_{i=1}^{m} p_\psi(\lambda_i) = 1$

Proof: 1. Starting from the definition $p_\psi(\lambda_i) = \langle\, \psi | P_i | \psi\, \rangle$ we use the projection property $P_i^2 = P_i$ and the Hermiticity of P_i to write

$$p_\psi(\lambda_i) = (|\psi\,\rangle, P_i^2|\psi\,\rangle) = (P_i|\psi\,\rangle, P_i|\psi\,\rangle) = |P_i|\psi\,\rangle|^2 \tag{3.24}$$

showning that $p_\psi(\lambda_i)$ is real and positive. To see that it is less than one note

$$(\langle\, \psi|P_i|\psi\,\rangle)^2 \leq ||\psi\,\rangle|^2 |P_i|\psi\,\rangle|^2$$

by the Cauchy-Schwarz inequality. Since $||\psi\,\rangle| = 1$ we deduce

$$p_\psi(\lambda_i)^2 \leq p_\psi(\lambda_i)$$

or

$$p_\psi(\lambda_i) \leq 1$$

2. Inserting the definition (3.1) and using the identity (3.14) we have

$$\sum_{i=1}^{m} p_\psi(\lambda_i) = \langle\, \psi| \sum_{i=1}^{m} P_i|\psi\,\rangle = \langle\, \psi|I|\psi\,\rangle = 1.$$

□

Corollary 3.3.7 *The ket* (3.2) *is a state vector, i.e. has norm 1.*

Proof: This follows from the calculation (3.24), which shows that the norm of $P_i|\psi\,\rangle$ is $\sqrt{p_\psi(\lambda_i)}$, so that $P_i|\psi\,\rangle/\sqrt{p_\psi(\lambda_i)}$ has norm 1 □

The Postulate 2 discussed in this subsection selects the possible outcomes of measurements of an observable A of a physical system and, given a state $|\psi\,\rangle$ of the system, assigns probabilities to each of these outcomes. Given such data we can compute the expectation value and standard deviation for repeated measurements of the observable A, assuming that the system is always prepared in the same state $|\psi\,\rangle$ before the measurement. Using the usual definition of expectation value as the average of the possible outcomes, weighted with their probabilities we have

$$\begin{aligned} E_\psi(A) &= \sum_{i=1}^{m} \lambda_i p_\psi(\lambda_i) \\ &= \sum_{i=1}^{m} \lambda_i \langle \psi | P_i | \psi \rangle \\ &= \langle \psi | \sum_{i=1}^{m} \lambda_i P_i | \psi \rangle \\ &= \langle \psi | A | \psi \rangle \end{aligned} \tag{3.25}$$

Motivated by this calculation we define:

Definition 3.3.8 *(***Expectation value and standard deviation***) Consider a system with Hilbert space V. The quantum mechancial expectation value of an observable A in the state $|\psi\rangle$ is defined as*

$$E_\psi(A) = \langle \psi | A | \psi \rangle. \tag{3.26}$$

The standard deviation of A is defined via

$$\Delta_\psi(A) = \sqrt{E_\psi(A^2) - (E_\psi(A))^2} \tag{3.27}$$

Note that

$$E_\psi\left((A - E_\psi(A)I)^2\right) = E_\psi\left((A^2 - 2E_\psi(A)A + (E_\psi(A))^2 I\right) = E_\psi(A^2) - (E_\psi(A))^2$$

so that the standard deviation is also given by

$$\Delta_\psi(A) = \sqrt{E_\psi\left((A - E_\psi(A)I)^2\right)} \tag{3.28}$$

Example 3.3.9 *Suppose that $|\psi\rangle$ is an eigenstate of the observable A with eigenvalue λ. Show that then $\Delta_\psi(A) = 0$.*

If $A|\psi\rangle = \lambda|\psi\rangle$ we have $\langle \psi | A | \psi \rangle = \lambda$ and $\langle \psi | A^2 | \psi \rangle = \lambda^2$. Hence

$$\Delta_\psi^2(A) = E_\psi(A^2) - (E_\psi(A))^2 = 0.$$

□

Physical interpretation: The expectation value and standard deviation of an observable play a crucial role in linking the formalism of quantum mechanics with experiment. The expectation value $\langle \psi | A | \psi \rangle$ of an observable is the prediction quantum mechanics makes for the average over the results of a repeated measurement of the observable A, assuming that the system is the state ψ at the time of the measurements. The standard deviation $\Delta_\psi(A)$

is the prediction quantum mechanics makes for the standard deviation of the experimental measurements. Note the contrast with classical physics, where an ideal experimental confirmation of a theory would produce the predicted result every time, with vanishing standard deviation. A non-vanishing standard deviation in experimental results is interpreted as a consequence of random errors and inaccurate measurements. In quantum mechanics even an experiment free of errors and inaccuracies is predicted to produce results with a non-vanishing standard deviation, except when the state of the system happens to be an eigenstate of the observable to be measured.

Although we have motivated the definitions of expectation value and standard deviation by the analogy with classiscal probablity theory, we will find some important differences between quantum mechanical expectation values and expectation values in classical probability theory in later sections, particularly in the discussion of Bell inequalities.

Example 3.3.10 *Compute the expectation value and standard deviation of the observable A in the state $|\psi\rangle$ of example 3.3.1*

$$\langle\psi|A|\psi\rangle = (1,0,0)\begin{pmatrix}1 & 1 & 0\\ 1 & 1 & 0\\ 0 & 0 & 2\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} = 1.$$

Since

$$A^2 = \begin{pmatrix}2 & 2 & 0\\ 2 & 2 & 0\\ 0 & 0 & 4\end{pmatrix}$$

we have

$$\langle\psi|A^2|\psi\rangle = (1,0,0)\begin{pmatrix}2 & 2 & 0\\ 2 & 2 & 0\\ 0 & 0 & 4\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} = 2$$

and therefore

$$\Delta_\psi(A) = \sqrt{2-1} = 1. \tag{3.29}$$

3.4 Time evolution

An important part of any physical model is mathematical description of how the system changes in time. In Newtonian mechanics this is achieved by Newton's second law, which states that the rate of change of the momentum of a particle is proportional to the force exerted on it. Newton's law does not specify the force but it postulates that there always is a force responsible for a change in momentum. The time evolution postulate in quantum mechanics is similar in this respect. It restricts the way in which the state of a quantum mechanical system changes with time.

Postulate 3: Time evolution is unitary

The time evolution of a closed system is described by a unitary transformation. If the state of the system is $|\psi\rangle$ at time t and $|\psi'\rangle$ at time t' then there is a unitary operator U so that

$$|\psi'\rangle = U|\psi\rangle \tag{3.30}$$

Before studying an example we note an important property of time evolution

Lemma 3.4.1 *Quantum mechanical time evolution preserves the norm of a state. In particular, in the terminology of Postulate 1, it maps a state vector into a state vector*

Proof: The preservation of the norm follows directly from the unitarity of U:

$$|U|\psi\,\rangle|^2 = (U|\psi\,\rangle, U|\psi\,\rangle) = (U^\dagger U|\psi\,\rangle, |\psi\,\rangle) = (|\psi\,\rangle, |\psi\,\rangle) = ||\psi\,\rangle|^2.$$

According to Postulate 1, state vectors are vectors of norm one. Since U perserves the norm, it maps state vectors to state vectors. $\square$.

Example 3.4.2 *Suppose a single qubit system with Hilbert space $V = \mathbb{C}^2$ is in the state $|0\,\rangle$ at time $t = 0$ seconds. The time evolution operator from time $t = 0$ seconds to time $t = 1$ second has the matrix representation*

$$U = \frac{1}{2}\begin{pmatrix} i\sqrt{3} & -1 \\ 1 & -i\sqrt{3} \end{pmatrix} \tag{3.31}$$

relative to the canonical basis. Check that U is unitary and find the state of the system at time $t = 1$ second. If a measurement in the canonical basis is carried out what is the probability of finding the system in the state $|0\,\rangle$ at time $t = 1$ seconds? What is the probability of finding in the state $|1\,\rangle$?

Checking unitary amounts to checking if $\bar{U}^t U = I$. This is a straightforward matrix calculation. According to the time evolution postulate, the state of the system at time $t = 1$ seconds is

$$|\psi'\,\rangle = \frac{1}{2}\begin{pmatrix} i\sqrt{3} & -1 \\ 1 & -i\sqrt{3} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} i\sqrt{3} \\ 1 \end{pmatrix} = \frac{i\sqrt{3}}{2}|0\,\rangle + \frac{1}{2}|1\,\rangle \tag{3.32}$$

According to the discussion preceding (3.12) the probability of finding the system in the the state $|0\,\rangle$ at time $t = 1$ seconds is therefore $|\langle\psi'|0\,\rangle|^2 = \frac{3}{4}$ and the probability of finding it in the state $|1\,\rangle$ at time $t = 1$ seconds is $|\langle\psi'|1\,\rangle|^2 = \frac{1}{4}$ $\square$.

The time evolution postulate of quantum mechanics is often stated in terms of a differential equation for the state vector. We give this alternative version here, and then show that it implies our earlier version of the time evolution postulate.

Postulate 3': Schrödinger equation *The time evolution of a closed system with associated Hilbert space V is governed by a differential equation for state vectors, called the Schrödinger equation. It takes the form*

$$i\hbar\frac{d}{dt}|\psi\,\rangle = H|\psi\,\rangle, \tag{3.33}$$

where $H : V \to V$ is a Hermitian operator, called the **Hamiltonian** *and $2\pi\hbar$ is a constant called Planck's constant.*

It is instructive to consider the "trivial" case where $V = \mathbb{C}$, so the time-dependent state vector is just a map $\psi : \mathbb{R} \to \mathbb{C}$, and a Hermitian operator H is a Hermitian 1×1 matrix, i.e. a real number. Then the Schrödinger equation becomes

$$\frac{d\psi}{dt} = -\frac{iH}{\hbar}\psi, \tag{3.34}$$

which is a first-order linear differential equation. The unique solution satisfying the initial condition $\psi(0) = \psi_0$ is

$$\psi(t) = e^{-\frac{i}{\hbar}tH}\psi_0. \tag{3.35}$$

Thus we see that the state at time t is obtained from the state at time $t = 0$ by multiplication with the phase $\exp(-\frac{i}{\hbar}tH)$ - which is a unitary operator $\mathbb{C} \to \mathbb{C}$, as required by Postulate 3.

In order to generalise the derivation of Postulate 3 from Postulate 3' to Hilbert spaces of arbitrary (finite) dimension, we need to study the exponentiation of Hermitian operators. We begin with the more general notion of a function of a Hermitian operator. The basic idea is to use the spectral decomposition given in (3.15):

Definition 3.4.3 *Let $A : V \to V$ be a Hermitian operator in the Hilbert space V, and suppose the spectral decomposition of A is*

$$A = \sum_{i=1}^{m} \lambda_i P_i \tag{3.36}$$

For a given function $f : \mathbb{R} \to \mathbb{R}$ we define the Hermitian operator $f(A)$ via

$$f(A) = \sum_{i=1}^{m} f(\lambda_i) P_i \tag{3.37}$$

The evaluation of the operator $f(A)$ is cumbersome if we have to find the spectral decomposition of A first. We can avoid this it the function f is analytic i.e. has a convergent power series in some neighbourhood of 0.

$$f(\lambda) = \sum_{n=0}^{\infty} a_n \lambda^n, \tag{3.38}$$

for real numbers a_n. In that case we use the result (3.20) to compute

$$\begin{aligned} f(A) &= \sum_{i=1}^{m} f(\lambda_i) P_i \\ &= \sum_{n=0}^{\infty} a_n \sum_{i=1}^{m} \lambda_i^n P_i \\ &= \sum_{n=0}^{\infty} a_n (\sum_{i=1}^{m} \lambda_i P_i)^n \\ &= \sum_{n=0}^{\infty} a_n A^n. \end{aligned} \tag{3.39}$$

Thus we see that we can compute $f(A)$ by formally inserting the operator A into the power series for f.

The following example shows that such power series of operators can sometimes be evaluated explicitly.

Example 3.4.4 *If* $H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *compute the matrix* $\exp(itH)$ *for* $t \in \mathbb{R}$.

We need to compute

$$\exp(itH) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!}(H)^n. \tag{3.40}$$

Noting that

$$H^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

and

$$H^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = H$$

etc. we have

$$\exp(itH) = \sum_{n\,\mathrm{even}} \frac{(it)^n}{n!} I + \sum_{n\,\mathrm{odd}} \frac{(it)^n}{n!} H.$$

But

$$\sum_{n\,\mathrm{even}} \frac{(it)^n}{n!} = 1 - \frac{t^2}{2} + \frac{t^4}{4!} \ldots = \cos(t)$$

and

$$\sum_{n\,\mathrm{odd}} \frac{(it)^n}{n!} = it - i\frac{t^3}{3!} + i\frac{t^5}{5!} \ldots = i\sin(t)$$

and therefore

$$\exp(itH) = \cos(t)I + i\sin(t)H = \begin{pmatrix} \cos t & i\sin t \\ i\sin t & \cos t \end{pmatrix}. \tag{3.41}$$

□

In the example we could evaluate the power series explicitly and thereby show that it converges. For a general operator A and a general analytic function f, the convergence of the power series for $f(A)$ needs to be checked. In general, the series will only have a finite radius of convergence. However, it follows from the convergence of the power series

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

for all x that the operator $\exp(A)$ has a convergent power series for any operator A. We combine this result with a result on the differentiation of power series in the following

Theorem 3.4.5 *Let H be a Hermitian operator in a Hilbert space V. Then the power series for* $\exp(itH)$ *converges for all $t \in \mathbb{R}$. Moreover,*

$$\frac{d}{dt}\exp(itH) = iH\exp(itH) = i\exp(itH)H. \tag{3.42}$$

Proof: The power series (3.40) for $\exp(itH)$ is absolutely and uniformly convergent and can therefore be differentiated term by term. Thus we find

$$\begin{aligned}
\frac{d}{dt}\exp(itH) &= \sum_{n=0}^{\infty} in\frac{(it)^{n-1}}{n!}(H)^n \\
&= iH\sum_{n=1}^{\infty}\frac{(it)^{n-1}}{(n-1)!}(H)^{n-1} \\
&= iH\exp(itH)
\end{aligned} \tag{3.43}$$

From the power series it is obvious that H commutes with $\exp(itH)$, so we also have

$$\frac{d}{dt}\exp(itH) = i\exp(itH)H$$

$\square$.

This theorem is very useful for writing down solutions of the Schrödinger equation with given initial conditions.

Corollary 3.4.6 *(***Time evolution operator***) The unique solution of the Schrödinger equation* (3.33) *satisfying the initial condition $|\psi(t=0)\rangle = |\psi_0\rangle$ is given by*

$$|\psi(t)\rangle = U(t)|\psi_0\rangle \tag{3.44}$$

where $U(t)$ is the time evolution operator

$$U(t) = \exp(-i\frac{t}{\hbar}H) \tag{3.45}$$

Proof: Using the theorem 3.4.5 and the chain rule to differentiate (3.44) we find

$$\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}H\exp(-i\frac{t}{\hbar}H)|\psi_0\rangle = -\frac{i}{\hbar}H|\psi(t)\rangle$$

so that

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle$$

and the Schrödinger equation is indeed satisfied. Moreover $U(0) = 1$ so $|\psi(t)\rangle = |\psi_0\rangle$ as required. $\square$

In order to make contact with our first version of the time evolution postulate we have to show that the time evolution operator defined by (3.45) is unitary. To do this we need the following lemma.

Lemma 3.4.7 *If A and B are Hermitian operators in a Hilbert space V with vanishing commutator $[A, B] = 0$ then*

$$\exp(A + B) = \exp(A)\exp(B) \tag{3.46}$$

Proof: According to the theorem 2.5.9 there exists a basis of V such that both A and B are diagonal with respect to that basis. Thus we can give spectral decompositions

$$A = \sum_{i=1}^{m} \lambda_i P_i \qquad B = \sum_{i=1}^{m} \mu_i P_i \tag{3.47}$$

with the *same* complete set of orthogonal projectors P_i. Hence

$$A + B = \sum_{i=1}^{m} (\lambda_i + \mu_i) P_i \tag{3.48}$$

and

$$\exp(A + B) = \sum_{i=1}^{m} e^{\lambda_i + \mu_i} P_i = \sum_{i=1}^{m} e^{\lambda_i} e^{\mu_i} P_i. \tag{3.49}$$

But by the same calculation as we carried out in the proof of (3.20) we find

$$\begin{aligned} \exp(A)\exp(B) &= (\sum_{i=1}^{m} e^{\lambda_i} P_i)(\sum_{j=1}^{m} e^{\mu_j} P_j) \\ &= \sum_{i=1}^{m} e^{\lambda_i} e^{\mu_i} P_i. \end{aligned} \tag{3.50}$$

□

We deduce

Theorem 3.4.8 *If H is a Hermitian operator in the Hilbert space V, the time evolution operator*

$$U(t) = \exp(-i\frac{t}{\hbar}H) \tag{3.51}$$

is unitary for all $t \in \mathbb{R}$.

Proof: It follows from the power series expression for $U(t)$ that

$$U^\dagger(t) = \exp(i\frac{t}{\hbar}H) \tag{3.52}$$

since H is Hermitian, i.e. $H^\dagger = H$. Since H commutes with $-H$ we can apply lemma 3.4.7 to conclude

$$U^\dagger U(t) = \exp(i\frac{t}{\hbar}H - i\frac{t}{\hbar}H) = \exp(0) = I, \tag{3.53}$$

thus establishing the unitarity of $U(t)$. □

Example 3.4.9 *Consider the Hilbert space $V = \mathbb{C}^2$ with its canonical inner product and the Hamiltonian with matrix representation*

$$H = b\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.54}$$

relative to the canonical basis.

1. *Find the time evolution operator and use it to solve the Schrödinger equation with initial condition $|\psi(t=0)\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.*
2. *What is the probability of finding the system in the orthogonal state $|\varphi\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ at time t?*
3. *Compute the expectation value at time t of the observable*

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

1. Since the matrix representing the Hamiltonian is diagonal the time evolution operator is

$$U(t) = \exp(-\frac{it}{\hbar}H) = \begin{pmatrix} e^{-\frac{itb}{\hbar}} & 0 \\ 0 & e^{\frac{itb}{\hbar}} \end{pmatrix} \tag{3.55}$$

Hence the state of the system at time t is

$$\begin{aligned} |\psi(t)\rangle &= U(t)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-\frac{itb}{\hbar}} \\ e^{\frac{itb}{\hbar}} \end{pmatrix} \\ &= \frac{1}{\sqrt{2}}e^{-\frac{itb}{\hbar}}|0\rangle + \frac{1}{\sqrt{2}}e^{\frac{itb}{\hbar}}|1\rangle. \end{aligned} \tag{3.56}$$

2. The probability of finding the system in the state $|\varphi\rangle$ is

$$|\langle\varphi||\psi(t)\rangle|^2 = \frac{1}{2}|e^{-\frac{itb}{\hbar}} - e^{\frac{itb}{\hbar}}|^2 = \sin^2\left(\frac{tb}{\hbar}\right). \tag{3.57}$$

Note that the probability oscillates between 0 and 1.

3. To compute the expectation value of the observable A at time t we note

$$A|\psi(t)\rangle = \frac{1}{\sqrt{2}}e^{-\frac{itb}{\hbar}}|1\rangle + \frac{1}{\sqrt{2}}e^{\frac{itb}{\hbar}}|0\rangle$$

and hence

$$\begin{aligned} \langle\psi(t)|A|\psi(t)\rangle &= \left(\frac{1}{\sqrt{2}}e^{-\frac{itb}{\hbar}}|0\rangle + \frac{1}{\sqrt{2}}e^{\frac{itb}{\hbar}}|1\rangle, \frac{1}{\sqrt{2}}e^{-\frac{itb}{\hbar}}|1\rangle + \frac{1}{\sqrt{2}}e^{\frac{itb}{\hbar}}|0\rangle\right) \\ &= \frac{1}{2}(e^{\frac{2itb}{\hbar}} + e^{\frac{-2itb}{\hbar}}) = \cos(\frac{2tb}{\hbar}). \end{aligned} \tag{3.58}$$

□

Generally, in order to compute the expectation value of an observable in the state $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ we need to evaluate

$$\langle\psi(t)|A|\psi(t)\rangle = (\psi(t), A\psi(t)) = (U(t)\psi(0), AU\psi(0)) = \big(\psi(0), U^\dagger AU\psi(0)\big) \qquad (3.59)$$

Writing the last expression in bra-ket notation and using the unitarity of U we have the equality

$$\langle\psi(t)|A|\psi(t)\rangle = \langle\psi(0)|U^{-1}(t)AU(t)|\psi(t)\rangle. \qquad (3.60)$$

This shows that the expectation value of the (time-independent) observable A in the time dependent state $|\psi(t)\rangle$ is the same as the expectation value of the time-dependent observable

$$A(t) = U^{-1}(t)AU(t) \qquad (3.61)$$

in the time-independent state $|\psi(0)\rangle$. The point of view where the observables obey the time evolution law (3.61) and the states are time-independent is called the **Heisenberg picture** of quantum mechancis. The point of view where states evolve according to the fundamental equation (3.30) is called the **Schrödinger picture**. We will mostly stick to the Schrödinger picture in this course.

3.5 The Heisenberg uncertainty relation

Heisenberg's uncertainty relation is one of the best known results in quantum mechanics. It sets an upper limit on the accuracy with which non-commuting observables can be measured. More precisely, for a given state of a system it gives a lower bound on the product of the standard deviations of two observables in terms of the expectation value of their commutator. We already saw in the example 3.3.9 that the standard deviation of an observable A in a state $|\psi\rangle$ vanishes if the state $|\psi\rangle$ is an eigenstate of A. On the other hand, we know from theorem 2.5.9 that, given two observables, there is a basis of simultaneous eigenvectors if and only if the observables commute. It therefore not surprising that the commutator of two observables controls the extent to which the standard deviation of both can be minimised. Mathematically, the uncertainty relation is not a very surprising or difficult result.

The fame of the uncertainty relation (also: uncertainty principle) is related to the role it played in the disucssion about the physical interpretation of quantum mechanics. It clearly points out a fundamental difference between quantum mechanics and classical physics, where any two quantities can, in principle, be measured to arbitrary accuracy. It is named after its discoverer, Werner Heisenberg, who, among the the inventors of quantum mechanics, is one of the most colourful and certainly the most controversial. Heisenberg belonged to the young generation of physicists who created quantum mechanics from the "old" quantum theory of Einstein, Planck and Bohr. He was awarded the Nobel Prize in 1932, at the mere age of 31. During the second world war he lead the unsuccessful German nuclear bomb project. His role there remains the source of much historical controversy. After the war, he played an important role in rebuilding German physics as the head of the Max-Planck-Institute of Physics in Göttingen (which later moved to Munich). If you want to read up about Heisenberg and his life, you could have a look at the Wikipedia article about Heisenberg, or read the wonderful play "Copenhagen" by Michael Frayn.

Theorem 3.5.1 (**Heisenberg uncertainty relation**) *Let A and B be two Hermitian operators in a Hilbert space V. Then, for any state $|\psi\rangle \in V$, the product of the standard deviations of A and B is bounded below by half the modulus of the expectation value of the commutator $[A, B]$; in symbols*

$$\Delta_\psi(A)\Delta_\psi(B) \geq \frac{1}{2}|\langle\,\psi|[A,B]|\psi\,\rangle| \tag{3.62}$$

Proof: Define the ($|\psi\,\rangle$-dependent) operators

$$C = A - E_\psi(A)I, \qquad D = B - E_\psi(B)I \tag{3.63}$$

so that, according to (3.28)

$$\Delta_\psi(A) = \sqrt{\langle\,\psi|C^2|\psi\,\rangle}, \qquad \Delta_\psi(B) = \sqrt{\langle\,\psi|D^2|\psi\,\rangle}. \tag{3.64}$$

Now apply the Cauchy-Schwarz inequality to the expectation value of the product CD, using the Hermiticity of C and D:

$$|\langle\,\psi|CD|\psi\,\rangle| \leq |C|\psi\,\rangle||D|\psi\,\rangle| = \sqrt{\langle\,\psi|C^2|\psi\,\rangle}\sqrt{\langle\,\psi|D^2|\psi\,\rangle} \tag{3.65}$$

Noting that

$$\begin{aligned}\langle\,\psi|[C,D]|\psi\,\rangle &= \langle\,\psi|CD|\psi\,\rangle - \langle\,\psi|DC|\psi\,\rangle \\ &= \langle\,\psi|CD|\psi\,\rangle - \langle\,\psi|(CD)^\dagger|\psi\,\rangle \\ &= \langle\,\psi|CD|\psi\,\rangle - \overline{\langle\,\psi|CD|\psi\,\rangle} \\ &= 2i\mathrm{Im}(\langle\,\psi|CD|\psi\,\rangle)\end{aligned} \tag{3.66}$$

and that for any complex number $w = a + ib$ we have $|w| = \sqrt{a^2+b^2} \geq |b| = |\mathrm{Im}(w)|$ we deduce

$$\frac{1}{2}|\langle\,\psi|[C,D]|\psi\,\rangle| \leq |\langle\,\psi|CD|\psi\,\rangle| \tag{3.67}$$

so that, together with (3.65) we have

$$\frac{1}{2}|\langle\,\psi|[C,D]|\psi\,\rangle| \leq \sqrt{\langle\,\psi|C^2|\psi\,\rangle}\sqrt{\langle\,\psi|D^2|\psi\,\rangle} \tag{3.68}$$

Now we note that $[A, B] = [C, D]$ so that (3.68) is equivalent to the claimed inequality (3.62). $\square$

Example 3.5.2 *Recall the definition of the Pauli matrices:*

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.69}$$

Show that $[\sigma_1, \sigma_2] = 2i\sigma_3$. Hence evaluate both sides of the Heisenberg uncertainty relation (3.62) for $A = \sigma_1$, $B = \sigma_2$ and for a general state $|\psi\,\rangle = \alpha|0\,\rangle + \beta|1\,\rangle$ in $\mathbb{C}^2$ (i.e. $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$). Find the condition on α and β for the equality to hold in (3.62).

Checking the commutation relation $[\sigma_1, \sigma_2] = 2i\sigma_3$ is a simple matrix calculation. Now note that $\sigma_1^2 = \sigma_2^2 = I$ so that

$$\Delta_\psi^2(\sigma_1) = \langle\psi|\psi\,\rangle - (\langle\,\psi|\sigma_1|\psi\,\rangle)^2 = |\alpha|^2 + |\beta|^2 - (\bar{\alpha}\beta + \bar{\beta}\alpha)^2 = 1 - (\bar{\alpha}\beta + \bar{\beta}\alpha)^2 \tag{3.70}$$

and

$$\Delta_\psi^2(\sigma_2) = \langle\psi|\psi\,\rangle - (\langle\,\psi|\sigma_2|\psi\,\rangle)^2 = |\alpha|^2 + |\beta|^2 + (\bar{\alpha}\beta - \bar{\beta}\alpha)^2 = 1 + (\bar{\alpha}\beta - \bar{\beta}\alpha)^2 \tag{3.71}$$

On the other hand

$$\langle\,\psi|\sigma_3|\psi\,\rangle = |\alpha|^2 - |\beta|^2. \tag{3.72}$$

Now define the real numbers

$$\begin{aligned} x &= \bar{\alpha}\beta + \bar{\beta}\alpha = 2\mathrm{Re}(\bar{\alpha}\beta) \\ y &= -i(\bar{\alpha}\beta - \bar{\beta}\alpha) = 2\mathrm{Im}(\bar{\alpha}\beta) \\ z &= |\alpha|^2 - |\beta|^2 \end{aligned} \tag{3.73}$$

and note that

$$\begin{aligned} x^2 + y^2 + z^2 &= (\bar{\alpha}\beta + \bar{\beta}\alpha)^2 - (\bar{\alpha}\beta - \bar{\beta}\alpha)^2 + (|\alpha|^2 - |\beta|^2)^2 \\ &= (|\alpha|^2 + |\beta|^2)^2 = 1 \end{aligned} \tag{3.74}$$

Therefore

$$\begin{aligned} \Delta_\psi^2(\sigma_1)\Delta_\psi^2(\sigma_2) &= (1 - x^2)(1 - y^2) \\ &= 1 - x^2 - y^2 + x^2y^2 = z^2 + x^2y^2 \end{aligned} \tag{3.75}$$

Since $z = \langle\,\psi|\sigma_3|\psi\,\rangle = \frac{1}{2i}\langle\,\psi|[\sigma_1, \sigma_2]|\psi\,\rangle$ we have

$$\Delta_\psi^2(\sigma_1)\Delta_\psi^2(\sigma_2) = \frac{1}{4}|\langle\,\psi|[\sigma_1, \sigma_2]|\psi\,\rangle|^2 + x^2y^2 \tag{3.76}$$

so that the equality

$$\Delta_\psi^2(\sigma_1)\Delta_\psi^2(\sigma_2) = \frac{1}{4}|\langle\,\psi|[\sigma_1, \sigma_2]|\psi\,\rangle|^2 \tag{3.77}$$

holds iff $x = 0$ or $y = 0$. Comparing with (3.73) we conclude that this is equivalent to $\mathrm{Re}(\bar{\alpha}\beta) = 0$ or $\mathrm{Im}(\bar{\alpha}\beta) = 0$ $\square$

4 Spin 1/2

We have often used the Hilbert space $V = \mathbb{C}^2$ in example calculations in this course. Historically, the use of this Hilbert space in physics goes back to 1924 when Wolfgang Pauli introduced what he called a "two-valued quantum degree of freedom" associated with the electron in the outermost shell of an atom. Pauli introduced these degrees of freedom to account for certain properties of atomic spectra, and for the behaviour of atoms in magnetic fields. It was subsequently pointed out by Uhlenbeck and Goudsmit that Pauli's degrees of freedom could be interpreted as describing a self-rotation or "spin" of the electron. Pauli formalised the theory of spin in 1927, introducing the Hilbert space $V = \mathbb{C}^2$ for his "two-valued quantum degree of freedom" and also giving Hermitian operators which describe the spin of the particles. As we shall explain, the spin of a particle with Hilbert space $\mathbb{C}^2$ is $\frac{\hbar}{2}$. Today we know that all experimentally observed elementary particles (electrons, muons, quarks etc.) have spin $\frac{\hbar}{2}$. It is common to drop the $\hbar$ and talk about "spin 1/2" particles.

In quantum computing the the Hilbert space $V = \mathbb{C}^2$ is the state space of a single qubit. This is the fundamental constituent of any quantum computer, just like a bit is the fundamental consituent of any classical computer. However, whereas there is little one can say about a single bit, a surprising amount of theory is necessary fully to understand a single qubit.

Mathematically, the Hilbert space $V = \mathbb{C}^2$ is the simplest space in which to illustrate the postulates of quantum mechanics. As we shall see, we can explicitly describe all Hermitian and all unitary operators acting in this space, thus giving us a complete picture of all observables and all possible time evolution operators. Moreover, we can interpret every state in $\mathbb{C}^2$ as an eigenstate of a physically interesting observable, thus giving us a physical interpretation of every state.

4.1 Spin operators

We begin by recalling the definition of the Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.1}$$

and noting the multiplication table

$$\begin{aligned} \sigma_1^2 = \sigma_2^2 = \sigma_3^2 &= I \\ \sigma_1\sigma_2 = -\sigma_2\sigma_1 &= i\sigma_3 \\ \sigma_2\sigma_3 = -\sigma_3\sigma_2 &= i\sigma_1 \\ \sigma_3\sigma_1 = -\sigma_1\sigma_3 &= i\sigma_2 \end{aligned} \tag{4.2}$$

which you should check as an exercise for sheet 5. The multiplication table (4.2) can be summarised succinctly using the epsilon symbol, defined as follows

$$\epsilon_{abc} = \begin{cases} 1 & \text{if} \quad a,b,c \quad \text{are a cyclical permutation of 1,2,3} \\ -1 & \text{if} \quad a,b,c \quad \text{are an anti-cyclical permutation of 1,2,3} \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

Thus, for example, $\epsilon_{121} = 0$ and $\epsilon_{213} = -1$. The required multiplication law takes the form

$$\sigma_a \sigma_b = \delta_{ab} I + i \sum_{c=1}^{3} \epsilon_{abc} \sigma_c \tag{4.4}$$

Definition 4.1.1 *(***Spin operators***) The Hermitian operators*

$$S_1 = \frac{\hbar}{2}\sigma_1, \quad S_2 = \frac{\hbar}{2}\sigma_2, \quad S_3 = \frac{\hbar}{2}\sigma_3 \tag{4.5}$$

are called the spin operators.

The characteristic mathematical property of spin operators is expressed in the following

Theorem 4.1.2 *(***Commutation relations of spin operators***)*

$$[S_a, S_b] = \sum_{c=1}^{3} i\hbar \epsilon_{abc} S_c. \tag{4.6}$$

Proof: This follows directly from the rule (4.4). For example

$$S_1 S_2 - S_2 S_1 = \frac{\hbar}{4}(\sigma_1\sigma_2 - \sigma_2\sigma_1) = \frac{2i\hbar}{4}\sigma_3 = i\hbar S_3 \tag{4.7}$$

etc. □

When the Hilbert space $\mathbb{C}^2$ describes the spin degrees of freedom of a particle, the Hermitian operators S_1,S_2 and S_3 represent the particle's spin about the 1, 2 and 3 axis. Here spin simply means angular momentum about an axis through the particle's centre of mass. As anticipated in the introductory remarks above, spin is therefore a measure of "self-rotation" of the particle. It is obvious from the matrix representation

$$S_3 = \hbar \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \tag{4.8}$$

that the spin operator S_3 has eigenvalues $\pm\frac{\hbar}{2}$. According to the quantum theory of spin these are the only possible outcomes in a measurement of spin along the 3-axis. Further below we shall give a simple argument why the eigenvalues of S_1 and S_2 are also $\pm\frac{\hbar}{2}$ (you are welcome to check this by a direct calculation). This fact is the reason for associating the internal Hilbert space $\mathbb{C}^2$ with "spin $\hbar/2$". It is worth comparing the quantum mechanical notion of spin with the description of spin in classical physics. When a top is spinning about a fixed axis with an angular momentum j classical mechanics (and our intuition) predicts that the projection of the angular momentum onto another axis can take any value in the interval $[-j, j] \subset \mathbb{R}$. According to quantum mechanics the measurement of the spin of a spin $s = 1/2$ particle along any axis only every produces the result $-\frac{\hbar}{2}$ or $\frac{\hbar}{2}$ - never any of the real numbers inbetween those values. More generally, the allowed values for the total spin in quantum mechanics are $s = \frac{n\hbar}{2}$ where n is an integer, and the allowed values for spin along any axis are $-\frac{n\hbar}{2}, -\frac{n\hbar}{2} + \hbar, \ldots, \frac{n\hbar}{2} - \hbar, -\frac{n\hbar}{2}$. Atomic and subatomic particles display precisely this kind of behaviour. Their spin is quantised, and the difference between any two allowed values of spin is an integer multiple of $\hbar$. In this sense, $\hbar$ is the "quantum of spin".

4.2 Hermitian operators in $\mathbb{C}^2$

The spin operators are examples of Hermitian operators in $\mathbb{C}^2$, and the identity operator is another obvious example. The next Lemma shows that all other Hermitian operators in $\mathbb{C}^2$ can be expressed as a linear combination of the identity matrix and the Pauli matrices.

Lemma 4.2.1 *Any Hermitian* 2×2 *matrix can be written as*

$$A = a_0 I + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3, \tag{4.9}$$

where a_0, a_1, a_2 *and* a_3 *are real numbers.*

Proof: First we check that the matrix (4.9) is indeed Hermitian. However, this follows from the fact that identity matrix I and the Pauli matrices σ_1, σ_2 and σ_3 are all Hermitian, so that

$$\begin{aligned}(a_0 I + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3)^\dagger &= a_0 I^\dagger + a_1\sigma_1^\dagger + a_2\sigma_2^\dagger + a_3\sigma_3^\dagger \\ &= a_0 I + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3. \end{aligned} \tag{4.10}$$

Alternatively, we can check the Hermiticity by writing out the matrix

$$A = \begin{pmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3 \end{pmatrix}. \tag{4.11}$$

Next we show that any Hermitian matrix can be written in the form (4.11). Thus consider a general 2×2 matrix with complex entries

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{4.12}$$

The requirement of Hermticity imposes the condition

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \bar{a_{11}} & \bar{a_{21}} \\ \bar{a_{12}} & \bar{a_{22}} \end{pmatrix}. \tag{4.13}$$

which implies that a_{11} and a_{22} are real and a_{12} and a_{21} each other's complex conjugate. Defining a_1 and a_2 to be the real and imaginary part of a_{21} and $a_0 = \frac{1}{2}(a_{11} + a_{22})$ as well as $a_3 = \frac{1}{2}(a_{11} - a_{22})$ we recover the representation (4.11) □

Often we collect the real numbers a_1, a_2, a_3 into one vector $\boldsymbol{a} = (a_1, a_2, a_3)$ in $\mathbb{R}^3$ and similarly collect the three Pauli matrices into a "vector of matrices"

$$\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3). \tag{4.14}$$

Then we use the abbreviation

$$\boldsymbol{a}\cdot\boldsymbol{\sigma} = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3. \tag{4.15}$$

As an illustration of the notation we study the following

Example 4.2.2 *Use the identity* (4.4) *to show that, for any vectors* $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^3$,

$$(\boldsymbol{p}\cdot\boldsymbol{\sigma})(\boldsymbol{q}\cdot\boldsymbol{\sigma}) = \boldsymbol{p}\cdot\boldsymbol{q}\, I + i(\boldsymbol{p} \times \boldsymbol{q})\cdot\boldsymbol{\sigma} \tag{4.16}$$

You can check the identity by writing out $\boldsymbol{p\sigma} = p_1\sigma_1 + p_2\sigma_2 + p_3\sigma_3$ and $\boldsymbol{q\sigma} = q_1\sigma_1 + q_2\sigma_2 + q_3\sigma_3$ and carrying out the multiplication term by term, using the rule (4.4). This is what you are asked to on sheet 5! □

4.3 Unitary operators in $\mathbb{C}^2$

In order to construct a parametrisation of all unitary operators in $\mathbb{C}^2$ we need the following

Lemma 4.3.1 *With the notation* (4.15) *we have, for a unit vector* $\boldsymbol{n} \in \mathbb{R}^3$*,*

$$\exp(i\phi\boldsymbol{n}\cdot\boldsymbol{\sigma}) = \cos\phi\, I + i\sin\phi\, \boldsymbol{n}\cdot\boldsymbol{\sigma}. \tag{4.17}$$

Proof: This follows by the same calculation that we carried out in example 3.4.4. The key fact is that $\boldsymbol{n}\cdot\boldsymbol{\sigma}$, like the operator H in 3.4.4 squares to I, as follows from (4.16) by setting $\boldsymbol{p} = \boldsymbol{q} = \boldsymbol{n}$. Thus

$$\begin{aligned} \exp(i\phi\boldsymbol{n}\cdot\boldsymbol{\sigma}) &= \sum_{k\,\mathrm{even}} \frac{(i\phi)^k}{k!} I + \sum_{k\,\mathrm{odd}} \frac{(i\phi)^k}{k!} \boldsymbol{n}\cdot\boldsymbol{\sigma} \\ &= \cos\phi\, I + i\sin\phi\, \boldsymbol{n}\cdot\boldsymbol{\sigma}, \end{aligned} \tag{4.18}$$

as was to be shown. □

Theorem 4.3.2 *(***Rotations***) Suppose* $\boldsymbol{n}$ *and* $\boldsymbol{m}$ *are vectors in* $\mathbb{R}^3$ *of unit length, i.e.* $\boldsymbol{n}^2 = \boldsymbol{m}^2 = 1$*. Then*

$$\exp(-\frac{i}{2}\alpha\boldsymbol{n}\cdot\boldsymbol{\sigma})\boldsymbol{m}\cdot\boldsymbol{\sigma}\exp(\frac{i}{2}\alpha\boldsymbol{n}\cdot\boldsymbol{\sigma}) = \boldsymbol{k}\cdot\boldsymbol{\sigma}, \tag{4.19}$$

where

$$\boldsymbol{k} = (\boldsymbol{n}\cdot\boldsymbol{m})\boldsymbol{n} + \cos\alpha\ (\boldsymbol{m} - (\boldsymbol{n}\cdot\boldsymbol{m})\boldsymbol{n}) + \sin\alpha\ (\boldsymbol{n}\times\boldsymbol{m}). \tag{4.20}$$

Proof: Using the formula (4.17) on the left hand side (LHS), we need to evaluate

$$LHS = (\cos\frac{\alpha}{2} - i\sin\frac{\alpha}{2}\boldsymbol{n}\cdot\boldsymbol{\sigma})\boldsymbol{m}\cdot\boldsymbol{\sigma}(\cos\frac{\alpha}{2} + i\sin\frac{\alpha}{2}\boldsymbol{n}\cdot\boldsymbol{\sigma}) \tag{4.21}$$

Now multiplying out using (4.16) we find

$$\begin{aligned} LHS &= (\cos\frac{\alpha}{2}\boldsymbol{m}\cdot\boldsymbol{\sigma} - i\sin\frac{\alpha}{2}\boldsymbol{n}\cdot\boldsymbol{m}I + \sin\frac{\alpha}{2}(\boldsymbol{n}\times\boldsymbol{m})\cdot\boldsymbol{\sigma})(\cos\frac{\alpha}{2} + i\sin\frac{\alpha}{2}\boldsymbol{n}\cdot\boldsymbol{\sigma}) \\ &= \cos^2\frac{\alpha}{2}\boldsymbol{m}\cdot\boldsymbol{\sigma} - i\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\boldsymbol{n}\cdot\boldsymbol{m}I + \sin\frac{\alpha}{2}\cos\frac{\alpha}{2}(\boldsymbol{n}\times\boldsymbol{m})\cdot\boldsymbol{\sigma} \\ &+ i\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\boldsymbol{n}\cdot\boldsymbol{m}I - \sin\frac{\alpha}{2}\cos\frac{\alpha}{2}(\boldsymbol{m}\times\boldsymbol{n})\cdot\boldsymbol{\sigma} \\ &+ \sin^2\frac{\alpha}{2}(\boldsymbol{n}\cdot\boldsymbol{m})(\boldsymbol{n}\cdot\boldsymbol{\sigma}) - \sin^2\frac{\alpha}{2}((\boldsymbol{n}\times\boldsymbol{m})\times\boldsymbol{n})\cdot\boldsymbol{\sigma}, \end{aligned} \tag{4.22}$$

where we used $(\boldsymbol{n}\times\boldsymbol{m})\cdot\boldsymbol{n} = 0$ in the last step. Now use the identiy

$$(\boldsymbol{n}\times\boldsymbol{m})\times\boldsymbol{n} = \boldsymbol{m} - (\boldsymbol{n}\cdot\boldsymbol{m})\boldsymbol{n}$$

and collect terms to find

$$LHS = (\cos^2\frac{\alpha}{2} - \sin^2\frac{\alpha}{2})\boldsymbol{m}\cdot\boldsymbol{\sigma} + 2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}(\boldsymbol{n}\times\boldsymbol{m})\cdot\boldsymbol{\sigma} + 2\sin^2\frac{\alpha}{2}(\boldsymbol{n}\cdot\boldsymbol{m})(\boldsymbol{n}\cdot\boldsymbol{\sigma})$$

Finally use the trigonometric identities

$$\begin{aligned}
\cos^2\frac{\alpha}{2} - \sin^2\frac{\alpha}{2} &= \cos\alpha, \\
2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2} &= \sin\alpha \\
2\sin^2\frac{\alpha}{2} &= 1-\cos\alpha
\end{aligned} \tag{4.23}$$

to rewrite the result as

$$LHS = \cos\alpha(\boldsymbol{m}\cdot\boldsymbol{\sigma}) + \sin\alpha(\boldsymbol{n}\times\boldsymbol{m})\cdot\boldsymbol{\sigma} + (1-\cos\alpha)(\boldsymbol{n}\cdot\boldsymbol{m})(\boldsymbol{n}\cdot\boldsymbol{\sigma}). \tag{4.24}$$

Re-arranging the terms now yields the RHS of (4.19), thus proving the claim. □

The formula (4.20) has an important **geometrical interpretation**: The vector $\boldsymbol{k}$ in (4.20) is obtained from the vector $\boldsymbol{m}$ applying a rotation by an angle α about the axis $\boldsymbol{n}$. The sense of the rotation is determined by the right-hand rule: point the thumb of your right hand in the direction of the axis; your fingers then point in the direction of the rotation. To illustrate this rule, we consider

Example 4.3.3 *Let*

$$R_{\boldsymbol{n}}[\alpha](\boldsymbol{m}) = (\boldsymbol{n}\cdot\boldsymbol{m})\boldsymbol{n} + \cos\alpha\ (\boldsymbol{m} - (\boldsymbol{n}\cdot\boldsymbol{m})\boldsymbol{n}) + \sin\alpha\ \boldsymbol{n}\times\boldsymbol{m} \tag{4.25}$$

and consider the canonical basis of $\mathbb{R}^3$

$$\boldsymbol{e}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad \boldsymbol{e}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}.$$

Compute

$$R_{\boldsymbol{e}_3}[\pi/4]\boldsymbol{e}_1, \quad R_{\boldsymbol{e}_2}[-\pi/2](\boldsymbol{e}_1), \quad \textit{and} \quad R_{\boldsymbol{e}_1}[\pi/2]\boldsymbol{e}_2.$$

Applying the formula (4.25), or thinking geometrically about the effect of rotating the vector $\boldsymbol{e}_1$ by $\pi/4$ (i.e. 45 degrees) about the axis $\boldsymbol{e}_3$ we find

$$R_{\boldsymbol{e}_3}[\pi/4]\boldsymbol{e}_1 = \frac{1}{\sqrt{2}}\boldsymbol{e}_1 + \frac{1}{\sqrt{2}}\boldsymbol{e}_2. \tag{4.26}$$

Similarly, rotating $\boldsymbol{e}_1$ by -90 degrees about $\boldsymbol{e}_2$ gives

$$R_{\boldsymbol{e}_2}[-\pi/2](\boldsymbol{e}_1) = \boldsymbol{e}_3. \tag{4.27}$$

and rotating $\boldsymbol{e}_2$ about $\boldsymbol{e}_1$ by 90 degrees we obtain

$$R_{\boldsymbol{e}_1}[\pi/2]\boldsymbol{e}_2 = \boldsymbol{e}_3. \tag{4.28}$$

□

As an immediate consquence we prove our earlier claim about the eigenvalues of the spin operators S_1 and S_2.

Example 4.3.4 *Show that the spin operators S_1, S_2 and S_3 can be conjugated into each other and therefore all have eigenvalues $\pm\frac{\hbar}{2}$*

Combining the result (4.27) from the previous example with the theorem 4.3.2 we deduce,

$$\exp(\frac{i\pi}{4}\sigma_2)\sigma_1\exp(-\frac{i\pi}{4}\sigma_2) = \sigma_3 \Rightarrow \exp(\frac{i\pi}{4}\sigma_2)S_1\exp(-\frac{i\pi}{4}\sigma_2) = S_3 \tag{4.29}$$

showing that S_3 is the diagonal form of S_1. Similarly, result (4.28) of the previous example implies

$$\exp(-\frac{i\pi}{4}\sigma_1)\sigma_2\exp(\frac{i\pi}{4}\sigma_1) = \sigma_3 \Rightarrow \exp(-\frac{i\pi}{4}\sigma_1)S_2\exp(\frac{i\pi}{4}\sigma_1) = S_3 \tag{4.30}$$

showing how to diagonalise S_2, and that the diagonal form of S_2 is S_3. Hence, S_1, S_2 and S_3 all have eigenvalues $\pm\frac{\hbar}{2}$. □

Corollary 4.3.5 *Let $\theta \in [0,\pi]$ and $\phi \in [0,2\pi)$ be angles parametrising unit vectors in $\mathbb{R}^3$ according to*

$$\boldsymbol{k}(\theta,\phi) = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} \tag{4.31}$$

Then

$$e^{-\frac{i}{2}\phi\sigma_3}e^{-\frac{i}{2}\theta\sigma_2}\sigma_3 e^{\frac{i}{2}\theta\sigma_2}e^{\frac{i}{2}\phi\sigma_3} = \boldsymbol{k}(\theta,\phi)\cdot\boldsymbol{\sigma}. \tag{4.32}$$

Proof: This follows by consecutive applications of theorem (4.3.2). First we compute

$$e^{-\frac{i}{2}\theta\sigma_2}\sigma_3 e^{\frac{i}{2}\theta\sigma_2} = \cos\theta\,\sigma_3 + \sin\theta\,\sigma_1$$

as well as

$$e^{-\frac{i}{2}\phi\sigma_3}\sigma_3 e^{\frac{i}{2}\phi\sigma_3} = \sigma_3, \quad \text{and} \quad e^{-\frac{i}{2}\phi\sigma_3}\sigma_1 e^{\frac{i}{2}\phi\sigma_3} = \cos\phi\,\sigma_1 + \sin\phi\,\sigma_2.$$

Combining, we deduce

$$e^{-\frac{i}{2}\phi\sigma_3}e^{-\frac{i}{2}\theta\sigma_2}\sigma_3 e^{\frac{i}{2}\theta\sigma_2}e^{\frac{i}{2}\phi\sigma_3} = \sin\theta\cos\phi\,\sigma_1 + \sin\theta\sin\phi\,\sigma_2 + \cos\theta\,\sigma_3, \tag{4.33}$$

which was to be shown. □

We end this subsection by giving a parametrisation of a general unitary operator in $\mathbb{C}^2$. It can be shown with the results proved in this subsection that our parametrisation captures all unitary operators. The proof is a little technical and therefore omitted (but feel free to give your own proof!)

Remark 4.3.6 *Any unitary 2×2 matrix can be written as*

$$U = e^{i\beta}\exp(i\mu\boldsymbol{k}\cdot\boldsymbol{\sigma}) \tag{4.34}$$

for angles $\beta, \mu \in [0,2\pi)$ and a unit vector $\boldsymbol{k} \in \mathbb{R}^3$.

4.4 Spin states

The spin operators S_1, S_2 and S_3 are the Hermitian operators corresponding to spin along the 1-, 2- and 3-axis. More generally we consider the operator

$$\boldsymbol{k}\cdot\boldsymbol{S} = k_1 S_1 + k_2 S_2 + k_3 S_3, \tag{4.35}$$

where $\boldsymbol{k} = (k_1, k_2, k_3)$ is a unit vector in $\mathbb{R}^3$. The operator (4.35) is the Hermitian operator corresponding to spin along the axis $\boldsymbol{k}$. According to the corollary 4.3.5 $\boldsymbol{k}\cdot\boldsymbol{\sigma}$ is conjugate to σ_3 and therefore has eigenvalues ± 1; hence $\boldsymbol{k}\cdot\boldsymbol{S}$ has eigenvalues $\pm\frac{\hbar}{2}$. In this section we find the general form of the eigenstates of $\boldsymbol{k}\cdot\boldsymbol{S}$. Furthermore, we show that, conversely, every state in $\mathbb{C}^2$ is in fact the eigenstate of $\boldsymbol{k}\cdot\boldsymbol{S}$ with eigenvalue $\frac{\hbar}{2}$ for some unit vector $\boldsymbol{k} \in \mathbb{R}^3$. This allows us to interpret an arbitrary state in $\mathbb{C}^2$ as the "spin up" state relative to some axis $\boldsymbol{k}$. In order to simplify the formula we consider the Pauli matrices σ_1, σ_2 and σ_3 instead of the corresponding spin operators here; to obtain the corresponding formulae for the spin operators you simply need to rescale by $\frac{\hbar}{2}$ at the appropriate places.

Lemma 4.4.1 *(***Spin eigenstates***) The states*

$$|(\theta,\phi)^+\rangle = e^{-\frac{i}{2}\phi\sigma_3} e^{-\frac{i}{2}\theta\sigma_2}|0\rangle \tag{4.36}$$

and

$$|(\theta,\phi)^-\rangle = e^{-\frac{i}{2}\phi\sigma_3} e^{-\frac{i}{2}\theta\sigma_2}|1\rangle \tag{4.37}$$

are eigenstates of the Hermitian operator $\boldsymbol{k}(\theta,\phi)\cdot\boldsymbol{\sigma}$ with eigenvalues respectively 1 and -1.

Proof: Using the parametrisation (4.32) of the Hermitian operator $\boldsymbol{k}(\theta,\phi)\cdot\boldsymbol{\sigma}$ we find

$$\begin{aligned} \boldsymbol{k}(\theta,\phi)\cdot\boldsymbol{\sigma}|(\theta,\phi)^+\rangle &= e^{-\frac{i}{2}\phi\sigma_3} e^{-\frac{i}{2}\theta\sigma_2}\sigma_3 e^{\frac{i}{2}\theta\sigma_2} e^{\frac{i}{2}\phi\sigma_3} e^{-\frac{i}{2}\phi\sigma_3} e^{-\frac{i}{2}\theta\sigma_3}|0\rangle \\ &= e^{-\frac{i}{2}\phi\sigma_3} e^{-\frac{i}{2}\theta\sigma_2}|0\rangle = |(\theta,\phi)^+\rangle \end{aligned} \tag{4.38}$$

where we used $\sigma_3|0\rangle = |0\rangle$. By an entirely analogous calculation, using $\sigma_3|1\rangle = -|1\rangle$, we deduce

$$\boldsymbol{k}(\theta,\phi)\cdot\boldsymbol{\sigma}|(\theta,\phi)^-\rangle = -|(\theta,\phi)^-\rangle \tag{4.39}$$

□

Example 4.4.2 *Find the components of the $\mathbb{C}^2$ vectors $|(\theta,\phi)^\pm\rangle$*

We expand

$$e^{-\frac{i}{2}\theta\sigma_2} = \cos\left(\frac{\theta}{2}\right) - i\sin\left(\frac{\theta}{2}\right)\sigma_2 = \begin{pmatrix} \cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix}$$

and

$$e^{-\frac{i}{2}\phi\sigma_3} = \begin{pmatrix} e^{-\frac{i}{2}\phi} & 0 \\ 0 & e^{\frac{i}{2}\phi} \end{pmatrix}.$$

Carrying out the matrix mutliplication we find

$$|(\theta,\phi)^+\rangle = e^{-\frac{i}{2}\phi\sigma_3}e^{-\frac{i}{2}\theta\sigma_2}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}e^{-\frac{i}{2}\phi}\cos(\frac{\theta}{2})\\ e^{\frac{i}{2}\phi}\sin(\frac{\theta}{2})\end{pmatrix} \tag{4.40}$$

and

$$|(\theta,\phi)^-\rangle = e^{-\frac{i}{2}\phi\sigma_3}e^{-\frac{i}{2}\theta\sigma_2}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-e^{-\frac{i}{2}\phi}\sin(\frac{\theta}{2})\\ e^{\frac{i}{2}\phi}\cos(\frac{\theta}{2})\end{pmatrix}. \tag{4.41}$$

□

Corollary 4.4.3 *Every vector* $|\psi\rangle \in \mathbb{C}^2$ *is eigenvector of* $\boldsymbol{k}\cdot\boldsymbol{\sigma}$ *with eigenvalue* 1 *for some unit vector* $\boldsymbol{k} \in \mathbb{R}^3$.

Given the state $|\psi\rangle = \begin{pmatrix}\alpha\\\beta\end{pmatrix} \in \mathbb{C}^2$, let us assume first that $\alpha \neq 0$. Then consider the complex number β/α. It has a unique paramerisation of the form

$$\frac{\beta}{\alpha} = \tan\left(\frac{\theta}{2}\right)e^{i\phi}, \tag{4.42}$$

where $\theta \in [0,\pi)$ and $\phi \in [0,2\pi)$. Then the state $|\psi\rangle$ must be of the form

$$|\psi\rangle = w\begin{pmatrix}e^{-\frac{i}{2}\phi}\cos(\frac{\theta}{2})\\ e^{\frac{i}{2}\phi}\sin(\frac{\theta}{2})\end{pmatrix}$$

for some complex number w and therefore proportional to (4.40). Hence it is an eigenstate of $\boldsymbol{k}(\theta,\phi)\cdot\boldsymbol{\sigma}$ with eigenvalue 1, where $\boldsymbol{k}(\theta,\phi)$ given by (4.31). If $\alpha = 0$ then

$$|\psi\rangle = \begin{pmatrix}0\\\beta\end{pmatrix}, \tag{4.43}$$

and this state is an eigenstate of $-\sigma_3$ with eigenvalue 1, i.e. an eigenstate of $\boldsymbol{k}\cdot\boldsymbol{\sigma}$ if $\boldsymbol{k} = (0,0,-1)$. □

4.5 The Stern-Gerlach experiment

In the Stern-Gerlach experiment a beam of silver atoms (which are electrically neutral and have spin 1/2) is sent through an inhomogeneous magnetic field. Each atom has a spin magnetic moment which interacts with the magnetic field. In quantum mechanics, the magnetic moment $\boldsymbol{M} = (M_1, M_2, M_3)$ is a vector of Hermitian operators, proportional to the spin vector $\boldsymbol{S}$:

$$M_a = \kappa S_a, \quad a = 1,2,3, \tag{4.44}$$

where κ is a proportionality constant which dependes on various physical quantities like the mass. Now let $\boldsymbol{k}$ be a unit vector which points from the north to the south pole of the magnet used in the Stern-Gerlach experiment. Then the inhomogeneous magnetic field causes the atom to be deflected either in the direction of $\boldsymbol{k}$ ("up") or in the opposite direction ("down"). A more detailed analysis shows that it effectively performs a measurement of the operator $\boldsymbol{k}\cdot\boldsymbol{M}$. Up to a constant of proportionality, this is the operator $\boldsymbol{k}\cdot\boldsymbol{\sigma}$ which we have studied in detail in this section. As we have seen, the eigenvalues of $\boldsymbol{k}\cdot\boldsymbol{\sigma}$ are $+1$ and -1; these eigenvalues correspond to the outcomes "deflected up" or "deflected down" in the Stern-Gerlach experiment. If we parametrise $\boldsymbol{k}$ as in (4.31), the eigenstate with eigenvalue 1 is $|(\theta,\phi)^+\rangle$ and the eigenstate with eigenvalue -1 is $|(\theta,\phi)^-\rangle$. The atoms which are deflected up are therefore in the state $|(\theta,\phi)^+\rangle$ and the atoms which are deflected down are in the state $|(\theta,\phi)^-\rangle$.

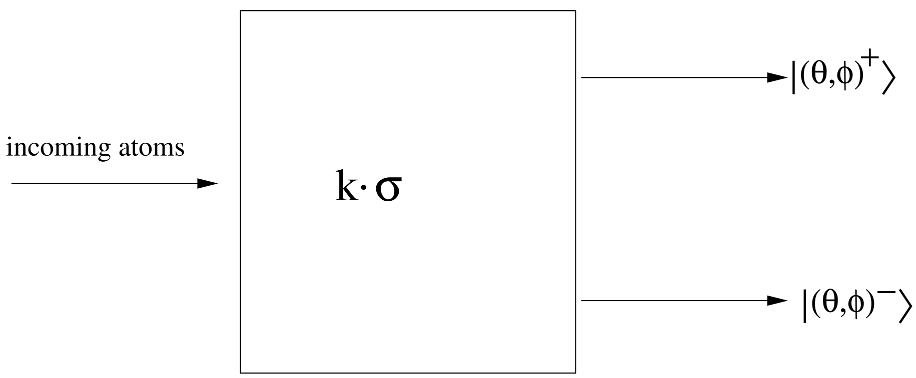

Figure 1: Schematic representation of the Stern-Gerlach experiment

The Stern-Gerlach experiment was performed in Frankfurt in 1922 by Otto Stern and Walther Gerlach with silver atoms. It played an important role in the genesis of quantum mechanics because it could not be explained with the laws of classical physics. A classical analysis of the experiment would go as follows. The electrically neutral but "spinning" atoms enter an inhomogneous magnetic field with their spin in some unknown direction. For some atoms, the spin is approximately aligned with the direction $\boldsymbol{k}$ from north to south pole, for others spin and $\boldsymbol{k}$ point in opposite directions, for most the angle between the spin and the $\boldsymbol{k}$ takes some intermediate value. The force experienced by the atoms dependes on this angle. It is such that atoms whose spin points in the direction of $\boldsymbol{k}$ ("up") should be deflected upwards and atoms whose spin points in the opposite direction of $\boldsymbol{k}$ ("down") should be deflected downwards; atoms whose spin is at right angles to $\boldsymbol{k}$ should not to be deflected at all. For intermediate angles we expect moderate deflections. However, in the Stern-Gerlach experiment, we witness that all atoms are deflected either up or down by the same amount. Quantum mechanics accounts for this, as we have seen. It allows only two outcomes of the experiment since the observable $\boldsymbol{k}\cdot\boldsymbol{M}$ being measured has precisely two eigenvalues.

Example 4.5.1 *(***Cascaded Stern-Gerlach experiments***) A beam of electrically neutral spin* $1/2$ *atoms is sent through a Stern-Gerlach apparatus with magnetic field direction* $\boldsymbol{k}_1 = \begin{pmatrix}0\\1\\0\end{pmatrix}$. *Subsequently the atoms which were deflected in the direction of* $\boldsymbol{k}_1$ *are sent through a*

Stern-Gerlach apparatus with magnetic field direction $\boldsymbol{k}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. *What is the probability of an atom being deflected "downwards" in the second apparatus, given that the initial state is* $|\psi\rangle = |0\rangle$*?*

The first Stern-Gerlach apparatus measures the operator $\boldsymbol{k}_1 \cdot \boldsymbol{\sigma} = \sigma_2$ In the parametrisation (4.31) this correponds to the angles $\theta = \frac{\pi}{2}$ and $\phi = \frac{\pi}{2}$. According to (4.40), the eigenstate with eigenvalue +1 is therefore

$$|(\frac{\pi}{2}, \frac{\pi}{2})^+\rangle = \frac{1}{2}\begin{pmatrix} 1-i \\ 1+i \end{pmatrix}. \tag{4.45}$$

Thus, according to Postulate 2, the probability of measuring the eigenvalue 1 is

$$\langle 0|(\frac{\pi}{2}, \frac{\pi}{2})^+\rangle\langle(\frac{\pi}{2}, \frac{\pi}{2})^+|0\rangle = \frac{1}{4}(1-i)(1+i) = \frac{1}{2} \tag{4.46}$$

and the state after the measurement is $|(\frac{\pi}{2}, \frac{\pi}{2})^+\rangle$. In the second Stern-Gerlach experiment, the operator $\boldsymbol{k}_2\boldsymbol{\sigma} = \sigma_3$ is measured. The outcome "downwards" corresponds to the eigenvalue -1 being measured, for which the eigenstate is $|1\rangle$. Given that the atom was in the state $|(\frac{\pi}{2}, \frac{\pi}{2})^+\rangle$ at the time of the measurement, the probability of this outcome is

$$\langle(\frac{\pi}{2}, \frac{\pi}{2})^+|1\rangle\langle 1|(\frac{\pi}{2}, \frac{\pi}{2})^+\rangle = \frac{1}{4}(1+i)(1-i) = \frac{1}{2}, \tag{4.47}$$

and the state of the atom after the measurement is $|1\rangle$. Hence the probability of measuring 1 in the first and -1 in the second Stern-Gerlach experiment is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. □

Note that in the example the state $|1\rangle$ after the second measurement is orthogonal to the initial state $|0\rangle$. If we had sent the atom only through the second Stern-Gerlach apparatus, the probability of measuring -1 would have been $\langle 0|1\rangle\langle 1|0\rangle = 0$.

5 The density operator

5.1 Ensembles of states

In this section we are going to generalise the notions of "state" and "expectation value", and formulate more general versions of the postulates of quantum mechanics. The drawback of the description of the measurement process in 3.3 is that it requires a precises knowledge of the state $|\psi\rangle$ of the system before the measurement. However, since it is eigenvalues of Hermitian operators and not the states which are the outcomes of measurements, we can only prepare the system in a given state $|\psi\rangle$ if that state is uniquely characterised by being the eigenstate of one or several Hermitian operators. This is the case when $|\psi\rangle$ is the unique (up to phase) eigenstate corresponding to the eigenvalue λ of a Hermitian operator A, or when $|\psi\rangle$ is the unique (up to phase) eigenstate corresponding to eigenvalues $\lambda, \mu, \ldots$ of several commuting Hermitian operators $A, B \ldots$. If, on the other hand, a Hermitian operator A has an eigenvalue λ with a two- (or higher) dimensional eigenspace, the measurement outcome λ by itself does not tell us the state of the system. We encountered this situation in discussing the example 3.3.1, where the observable A had a two-dimensional eigenspace for the eigenvalue $\lambda_2 = 2$ spanned by $|b_{2,1}\rangle, |b_{2,2}\rangle$. If we had measured the eigenvalue $\lambda_2 = 2$ of the observable A without knowledge of the state of the system before the measurement we would only know that the state of the system after the measurement is $|b_{2,1}\rangle$ or $|b_{2,2}\rangle$ or indeed any superposition of these two states. If we were to perform a further measurement of a different observable, we would not be able to use Postulate 2 to calculate probabilities and the state after the measurement since we do not know which initial state $|\psi\rangle$ to use.

The usual way of parametrising ignorance in science is to ascribe probabilities to the various possibilities. Consider a generalisation of the example, where we have a collection of orthonormal states $|\psi_k\rangle$, $k = 1, \ldots, K$. Suppose we know that the system is in of the states $|\psi_k\rangle$, but we do not know which. Instead we have probabilities p_k, $k = 1, \ldots, K$, for each of the states $|\psi_k\rangle$. The set

$$\mathcal{E} = \{(p_k, |\psi_k\rangle)\}_{k=1,\ldots,K} \tag{5.1}$$

is called a **ensemble of states**. Given an ensemble of states we reformulate Postulate 2 about the measurement of an observable A as follows.

Suppose the observable has the spectral decomposition

$$A = \sum_{i=1}^{m} \lambda_m P_m. \tag{5.2}$$

in terms of orthogonal projection operators P_i and eigenvalues λ_i, $i = 1, \ldots, m$. The possible outcomes in a measurement of A are the eigenvalues $\lambda_1, \ldots, \lambda_m$. If the state of the system is described by the ensemble (5.1) then we know that

$$\text{Probability of system being in state } |\psi_k\rangle \; = p_k \tag{5.3}$$

and

$$\text{Probability of measuring } \lambda_i \text{ given that system is in state } |\psi_k\rangle = p_{\psi_k}(\lambda_i). \tag{5.4}$$

Hence, using the standard "and" and "or" rules of classical probability, the probability of measuring the eigenvalue λ_i is

$$p_{\mathcal{E}}(\lambda_i) = \sum_{k=1}^{K} p_k p_{\psi_k}(\lambda_i), \tag{5.5}$$

Using the formula (3.1) for $p_{\psi_k}(\lambda_i)$ we have the equivalent expression

$$p_{\mathcal{E}}(\lambda_i) = \sum_{n=1}^{N} p_k \langle\, \psi_k | P_i | \psi_k \,\rangle. \tag{5.6}$$

In computing expectation values of the observable A we average the expectation values for each of the states in the ensemble:

$$E_{\mathcal{E}}(A) = \sum_{k=1}^{K} p_k \langle\, \psi_k | A | \psi_k \,\rangle. \tag{5.7}$$

What is the ensemble after the measurement? Applying the projection rule (3.2) to each of the states $|\psi_n\,\rangle$ of the ensemble, the ensemble after the measurement contains the states $P_i|\psi_k\,\rangle$, $k = 1, \ldots, K$. Again using standard probability theory for conditional probabilities

Probability of system being in state $|\psi_k\,\rangle$ given that λ_i has been measured

$$\begin{aligned} &= \frac{\text{Probability of system being in state } |\psi_k\,\rangle \textbf{ and } \text{measuring } \lambda_i}{\text{Probability of measuring } \lambda_i} \\ &= \frac{p_k p_{\psi_k}(\lambda_i)}{p_{\mathcal{E}}(\lambda_i)} \end{aligned} \tag{5.8}$$

Hence the ensemble after the measurement is

$$\tilde{\mathcal{E}} = \{ \left(\frac{p_k p_{\psi_k}(\lambda_i)}{p_{\mathcal{E}}(\lambda_i)}, \frac{1}{\sqrt{p_{\psi_k}(\lambda_i)}} P_i |\psi_k\,\rangle \right) \}_{k=1,\ldots,K} \tag{5.9}$$

Extending the measurement postulate by using the notion of an ensemble addresses our original concern. If we only know that the state of a system is in some K-dimensional subspace W of the full Hilbert space V, we might pick an orthonormal basis $|\psi_k\,\rangle$ of W and, based on our total ignorance, assign equal probabilities $p_k = \frac{1}{K}$ to each of the basis states $|\psi_k\,\rangle$. Using the rules (5.5), (5.9) and (5.7) we can then analyse measurements and compute expectation values

Example 5.1.1 *Consider the Hilbert space $\mathbb{C}^2$ and the observable*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{5.10}$$

In order to see the difference between a superposition and an ensemble, consider the state

$$|\psi\,\rangle = \alpha|0\,\rangle + \beta|1\,\rangle,$$

where α and β are complex numbers satisfying $|\alpha|^2 + |\beta|^2 = 1$, and the ensemble

$$\mathcal{E} = \{(|\alpha|^2, |0\rangle), (|\beta|^2, |1\rangle)\}$$

For both $|\psi\rangle$ and $\mathcal{E}$, compute the probability of measuring the eigenvalue 2 of the observable A, and give the state, respectively the ensemble, after the measurement. Also compute the expectation value of A for both the state $|\psi\rangle$ and the ensemble $\mathcal{E}$.

The eigenvector for the eigenvalue 2 of A is $|v\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. The probability of measuring this eigenvalue, given that the system is in the state $|\psi\rangle$, is

$$p_\psi(2) = \langle\psi|v\rangle\langle v|\psi\rangle = \frac{1}{2}(|\alpha|^2 + \alpha\bar{\beta} + \bar{\alpha}\beta + |\beta|^2) = \frac{1}{2}|\alpha+\beta|^2,$$

and the state after the measurement is

$$\frac{1}{\sqrt{p_\psi(2)}}|v\rangle\langle v|\psi\rangle = \frac{\alpha+\beta}{|\alpha+\beta|}|v\rangle,$$

i.e. up to the phase $e^{i\delta} := (\alpha+\beta)/|\alpha+\beta|$ the state after the measurement is the eigenstate $|v\rangle$ for the eigenvalue 2. For the expectation value we find

$$E_\psi(A) = |\alpha|^2 + \alpha\bar{\beta} + \bar{\alpha}\beta + |\beta|^2 = |\alpha+\beta|^2.$$

In order to analyse the measurement from the point of view of the ensemble $\mathcal{E}$ we need the probability of measuring the eigenvalue 2 given that the system was in the state $|0\rangle$

$$p_0(2) = \langle 0|v\rangle\langle v|0\rangle = \frac{1}{2}$$

and the probability of measuring the eigenvalue 2 given that the system was in the state $|1\rangle$

$$p_1(2) = \langle 1|v\rangle\langle v|1\rangle = \frac{1}{2}.$$

Hence the probability of measuring 2 if the system is described by the ensemble $\mathcal{E}$ is

$$p_{\mathcal{E}}(2) = |\alpha|^2 p_0(2) + |\beta|^2 p_1(2) = \frac{1}{2}(|\alpha|^2 + |\beta|^2) = \frac{1}{2}.$$

To find the ensemble after the measurement we note that

$$\frac{1}{\sqrt{p_0(2)}}|v\rangle\langle v||0\rangle = |v\rangle, \qquad \frac{1}{\sqrt{p_1(2)}}|v\rangle\langle v||1\rangle = |v\rangle$$

and therefore the ensemble after the measurement is

$$\mathcal{E}' = \{(|\alpha|^2, |v\rangle), (|\beta|^2, |v\rangle)\}; \tag{5.11}$$

Since the state $|v\rangle$ appears twice, with a total probability $|\alpha|^2 + |\beta|^2 = 1$, the state of the system after the measurement is $|v\rangle$. Finally, the expectation value is

$$E_{\mathcal{E}}(A) = |\alpha|^2 + |\beta|^2 = 1. \qquad \square$$

The example shows that calculations with ensembles can be cumbersome, in particular the determination of the ensemble after the measurement. The example also highlights a subtlety in the notion of a state which we discussed after stating Postulate 1 in Sect. 3. The vectors $|v\,\rangle$ and $e^{i\delta}|v\,\rangle$, where δ is an arbitrary real number, are eigenvectors of A with the same eigenvalue 2 and they are both normalised to unit length. In quantum mechanics we can identify $|v\,\rangle$ and $e^{i\delta}|v\,\rangle$, i.e. we can consider them to be the same state. We have not done that in our formulation of Postulate 1 mainly for pedagogical reasons. However, we shall see that the new formulation of the postulates in this section takes care of this problem. Our new notion of states will not distinguish between $|v\,\rangle$ and $e^{i\delta}|v\,\rangle$.

The key idea for the new formulation of the postulates is to associate to each normalised vector ψ the projection operator

$$P_\psi = |\psi\,\rangle\langle\,\psi|. \tag{5.12}$$

Clearly, the projection operator is the same for $|\psi\,\rangle$ and $e^{i\delta}|\psi\,\rangle$, since the phase drops out in (5.12). Furthermore we note

Lemma 5.1.2 *For any Hermitian operator A acting in the Hilbert space V and any state $|\psi\,\rangle \in V$*

$$\langle\,\psi|A|\psi\,\rangle = \mathrm{tr}(P_\psi A). \tag{5.13}$$

Proof: Complete the $|\psi\,\rangle$ to an orthonormal basis $\{|\psi\,\rangle, |b_2\,\rangle, \dots |b_n\,\rangle\}$ of V. Then

$$\begin{aligned} \mathrm{tr}(P_\psi A) &= \langle\,\psi|P_\psi A|\psi\,\rangle + \sum_{j=2}^{n}\langle\,b_j|P_\psi A|b_j\,\rangle \\ &= \langle\psi|\psi\,\rangle\langle\,\psi|A|\psi\,\rangle + \sum_{j=2}^{n}\langle b_j|\psi\,\rangle\langle\,\psi|A|b_j\,\rangle \\ &= \langle\,\psi|A|\psi\,\rangle, \end{aligned} \tag{5.14}$$

where we used the orthonormality of the basis $\{|\psi\,\rangle, |b_2\,\rangle, \dots |b_n\,\rangle\}$ of V. □

Using this lemma we write the probability $p_\psi(\lambda_i)$ as

$$p_\psi(\lambda_i) = \mathrm{tr}(P_\psi P_i) \tag{5.15}$$

and the expectation value $E_\psi(A)$ as

$$E_\psi(A) = \mathrm{tr}(P_\psi A) \tag{5.16}$$

Thus, if we associate the operator

$$\rho_{\mathcal{E}} = \sum_{k=1}^{K} p_k|\psi_k\,\rangle\langle\,\psi_k| \tag{5.17}$$

to the ensemble $\mathcal{E}$ in (5.1), we can write the probability (5.5) as

$$p_{\mathcal{E}}(\lambda_i) = \mathrm{tr}(\rho_{\mathcal{E}} P_i) \tag{5.18}$$

and the expectation value (5.7) as

$$E_{\mathcal{E}}(A) = \mathrm{tr}(\rho_{\mathcal{E}} A). \tag{5.19}$$

Operators like (5.17) are called density operators. We give a careful definition of such operators below, and will rephrase our quantum mechanical postulates in terms of them. In order to formulate all of the quantum mechanical postulates in terms of density operators we need the following

Lemma 5.1.3 *If $\langle\,\psi|$ is the bra corresponding to the ket $|\psi\,\rangle$ in a Hilbert space V and A is an operator $V \to V$ then the bra corresponding to the ket $A|\psi\,\rangle$ is $\langle\,\psi|A^\dagger$.*

Proof: If you are happy with the extension of the definition of † to bra's and ket's in(2.63) and (2.64) you will like the following one-line calculation of the bra corrsponding to $A|\psi\,\rangle$:

$$(A|\psi\,\rangle)^\dagger = |\psi\,\rangle^\dagger A^\dagger = \langle\,\psi|A^\dagger. \tag{5.20}$$

A proof starting from first principles goes as follows. Recall that, by definition, the bra $\langle\,\psi|$ is the map

$$\langle\,\psi| : V \to V, \quad |\varphi\,\rangle \mapsto \langle\psi|\varphi\,\rangle = (|\psi\,\rangle, |\varphi\,\rangle) \tag{5.21}$$

Thus the bra associated to $A|\psi\,\rangle$ is the map

$$|\varphi\,\rangle \mapsto (A|\psi\,\rangle, \varphi) = (|\psi\,\rangle, A^\dagger\varphi) \tag{5.22}$$

which is the compositon of the maps

$$|\varphi\,\rangle \mapsto A^\dagger|\varphi\,\rangle \mapsto (|\psi\,\rangle, A^\dagger\varphi), \tag{5.23}$$

and this is precisely the definition of $\langle\,\psi|A^\dagger$. □

It follows in particular that if P is an orthogonal (i.e. Hermitian) projection operator then the bra corresponding to $P|\psi\,\rangle$ is $\langle\,\psi|P$. Hence the density operator constructed from the ensemble (5.9) after the measurement is

$$\rho_{\tilde{\mathcal{E}}} = \sum_{k=1}^{K} \frac{p_k}{p_{\mathcal{E}}(\lambda_i)} P_i|\psi_k\,\rangle\langle\,\psi_k|P_i \tag{5.24}$$

Note that the dependence on $p_{\psi_k}(\lambda_i)$ drops out. Recalling the formula (5.18) we can write the density operator after the measurement very elegantly in terms of the density operator before the measurement and the projection operator P_i:

$$\rho_{\tilde{\mathcal{E}}} = \frac{P_i \rho_{\mathcal{E}} P_i}{\mathrm{tr}(\rho_{\mathcal{E}} P_i)}. \tag{5.25}$$

Finally we note that the time evoluation postulate can also be formulated very simply in terms of the density operator. If the time evolution of the states $|\psi_k\,\rangle$ in the ensemble $\mathcal{E}$

from time t to time t' is given by the unitary operator U, so that the states at t' are given by

$$|\psi_k'\rangle = U|\psi_k\rangle \tag{5.26}$$

then the corresponding density operator evolves to

$$\rho_{\mathcal{E}'} = \sum_{k=1}^{K} U p_k|\psi_k\rangle\langle\psi_k|U^\dagger = U\rho_{\mathcal{E}}U^{-1}, \tag{5.27}$$

where we used the unitarity of U.

Before we re-write the postulates of quantum mechanics in terms of density operators, we give a general definition. The definition is motivated by two properties of the density operators we have considered so far.

Definition 5.1.4 *(***Density operator***) A density operator in a Hilbert space V is any Hermitian operator $\rho : V \to V$ satisfying the conditions*

1. *(***Trace condition***)* $\mathrm{tr}(\rho) = 1$
2. *(***Positivity***) ρ is a positive operator, i.e. for any state $|\psi\rangle \in V$, $\langle\psi|\rho|\psi\rangle \geq 0$.*

It is not difficult to check that the density operator $\rho_{\mathcal{E}}$ (5.17) associated to the ensemble $\mathcal{E}$ (5.1) satisfies the conditions. Complement the orthonormal set $\{|\psi_1\rangle, \ldots, |\psi_K\rangle\}$ to an orthonormal basis $\{|\psi_1\rangle, \ldots, |\psi_K\rangle, |b_{K+1}\rangle, \ldots, |b_n\rangle\}$ of the n-dimensional Hilbert space V. Then, using the orthogonality of the basis,

$$\begin{aligned}\mathrm{tr}(\rho_{\mathcal{E}}) &= \sum_{j=1}^{K}\sum_{k=1}^{K} p_k\langle\psi_j|\psi_k\rangle\langle\psi_k|\psi_j\rangle + \sum_{j=K+1}^{n}\sum_{k=1}^{K} p_k\langle b_j|\psi_k\rangle\langle\psi_k|b_j\rangle \\ &= \sum_{j=1}^{K}\sum_{k=1}^{K} p_k\delta_{jk} = \sum_{k=1}^{K} p_k = 1 \end{aligned} \tag{5.28}$$

by the requirement that probabilities add up to 1. Furthermore, for any state $|\psi\rangle$

$$\langle\psi|\rho|\psi\rangle = \sum_{k=1}^{K} p_k\langle\psi|\psi_k\rangle\langle\psi|\psi_k\rangle = \sum_{k=1}^{K} p_k|\langle\psi|\psi_k\rangle|^2 \geq 0 \tag{5.29}$$

since each term in the sum is non-negative.

Perhaps more surprisingly, the reverse is also true:

Theorem 5.1.5 *Let ρ be a density operator, i.e. an operator acting in a Hilbert space V and satisfying the conditions in the definition 5.1.4. Then there exists an ensemble*

$$\mathcal{E} = \{(p_k, |\psi_k\rangle)\}_{k=1,\ldots,K}$$

with $K \leq n = \mathrm{dim}V$ so that

$$\rho = \rho_{\mathcal{E}} = \sum_{k=1}^{K} p_k|\psi_k\rangle\langle\psi_k| \tag{5.30}$$

Proof: By assumption, ρ is Hermitian and therefore has a spectral decomposition

$$\rho = \sum_{i=1}^{n} \lambda_i |b_i\,\rangle\langle\, b_i| \tag{5.31}$$

in terms of an orthonormal basis $|b_1\,\rangle, \ldots, |b_n\,\rangle$ of V. By the positivity of ρ

$$\langle\, b_i|\rho|b_i\,\rangle = \lambda_i \geq 0 \tag{5.32}$$

for all $i = 1, \ldots, n$. Computing the trace we also find

$$\mathrm{tr}(\rho) = \sum_{i=1}^{n} \lambda_i = 1. \tag{5.33}$$

However, if a sum of positive numbers is 1, each of the positive numbers must lie between 0 and 1. We can therefore interpret them as probabilities. After dropping the basis elements $|b_i\,\rangle$ for which $\lambda_i = 0$ and renaming the remaining eigenvalues $\lambda_k \to p_k$ and the the remaining states $|b_k\,\rangle \to |\psi_k\,\rangle$ we obtain the required ensemble. □

Motivated by our calculations with the density operator $\rho_{\mathcal{E}}$ we now reformulate the postulates of quantum mechanics.

5.2 The postulates of quantum mechanics in terms of density operators

Postulate 1': State space

Associated to every isolated physical system is a complex vector space V with inner product (Hilbert space) called the state space of the system. At any give time the physical state of the system is completely described by a density operator, which is Hermitian operator $V \to V$ satisfying the conditions in the definition 5.1.4.

The density operators made from a single ket $|\psi\,\rangle$ - our old notion of "state" - still play a special role and are called pure states. They can be characterised as follows.

Definition 5.2.1 *We say that a density operator ρ defines a* **pure state** *if it has precisely one non-zero eigenvalue (which must then be equal to 1). Otherwise, the density operator is said to characterise a* **mixed state**

Lemma 5.2.2 *(***Criterion for pure states***) Every density operator ρ satisfies*

$$\mathrm{tr}(\rho^2) \leq 1 \tag{5.34}$$

The equality $\mathrm{tr}(\rho^2) = 1$ holds if and only if ρ describes a pure state.

Proof: Using the spectral decomposition

$$\rho = \sum_{k=1}^{K} p_k |\psi_k\,\rangle\langle\, \psi_k| \tag{5.35}$$

and (3.20) we compute

$$\rho^2 = \sum_{k=1}^{K} p_k^2 |\psi_k\rangle\langle\psi_k|. \tag{5.36}$$

Since $0 \le p_k \le 1$ we have $p_k^2 \le p_k$. Hence

$$\mathrm{tr}(\rho^2) = \sum_{k=1}^{K} p_k^2 \le \sum_{k=1}^{K} p_k = 1. \tag{5.37}$$

The equality $p_k^2 = p_k$ holds iff p_k is either 1 or 0. However, since $\sum_{k=1}^{K} p_k = 1$ this can only happen if precisely one of the p_k is 1 and the others are 0 i.e. if ρ describes a pure state. Hence the equality $\mathrm{tr}(\rho^2) = 1$ holds iff ρ describes a pure state. □

Example 5.2.3 *For each of the following density operators decide if they describe pure or mixed states. If they describe a pure state, find a ket $|\psi\rangle$ so that $\rho = |\psi\rangle\langle\psi|$.*

$$(i)\, \rho = \frac{1}{4}\begin{pmatrix} 1 & -1 \\ 3 & 3 \end{pmatrix} \quad (ii)\, \rho = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{5.38}$$

(i) ρ^2 has diagonal entries $\frac{-2}{16}$ and $\frac{6}{16}$ (don't bother working out all entries!) so $\mathrm{tr}(\rho^2) = \frac{1}{4} < 1$ and ρ is a mixed state.

(ii) $\rho^2 = \rho$ in this case, so $\mathrm{tr}(\rho^2) = 1$ and the state is pure. The ket $|b_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ is eigenvector with eigenvalue 0 and the ket $|b_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ is eigenvector with eigenvalue 1. Hence

$$\rho = |b_2\rangle\langle b_2|$$

is the required representation of ρ. □

Example 5.2.4 *Show that the most general density operator in $\mathbb{C}^2$ is of the form*

$$\rho = \frac{1}{2}(I + \boldsymbol{r}\cdot\boldsymbol{\sigma}), \tag{5.39}$$

where $\boldsymbol{r}$ is a vector in $\mathbb{R}^3$ of length at most 1.

Density operators are Hermitian, and we saw in Sect. 4 that any Hermitian operator can be written as

$$\rho = a_0 I + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 \tag{5.40}$$

in terms of real numbers a_0, a_1, a_2, a_3, see equation (4.9). The condition $\mathrm{tr}(\rho) = 1$ for density operators implies

$$\mathrm{tr}(\rho) = 2a_0 = 1 \Rightarrow a_0 = \frac{1}{2}.$$

With the notation $\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ the requirement of positivity means that for any $|\psi\rangle \in \mathbb{C}^2$ with $\langle\psi, |\psi\rangle = 1$

$$\langle\psi|\rho|\psi\rangle = \frac{1}{2} + \langle\psi|\boldsymbol{a}\cdot\boldsymbol{\sigma}|\psi\rangle \geq 0 \Rightarrow \langle\psi|\boldsymbol{a}\cdot\boldsymbol{\sigma}|\psi\rangle \geq -\frac{1}{2}. \tag{5.41}$$

Now write $\boldsymbol{a} = a\boldsymbol{k}$, where $\boldsymbol{k}$ is a unit vector and $a = |\boldsymbol{a}|$. Then we know from Sect. 4 that the operator $\boldsymbol{k}\cdot\boldsymbol{\sigma}$ has eigenvalues ± 1. Let us denote the eigenvectors by $|+\rangle$ and $|-\rangle$ for brevity. Expanding

$$|\psi\rangle = \alpha|+\rangle + \beta|-\rangle$$

with $|\alpha|^2 + |\beta|^2 = 1$ we deduce

$$\langle\psi|\boldsymbol{a}\cdot\boldsymbol{\sigma}|\psi\rangle = a(|\alpha|^2 - |\beta|^2) \geq -a.$$

Hence, comparing with (5.41) we obtain a positive operator if we pick $a \leq \frac{1}{2}$. Defining $\boldsymbol{r} = 2\boldsymbol{a}$, the most general density operator is therefore of the form (5.39) with $|\boldsymbol{r}| \leq 1$. $\square$.

Postulate 2': Observables and measurements

The physically observable quantities of a physical system, also called the observables, are mathematically described by Hermitian operators acting on the state space V of the system. The possible outcomes of measurements of an observable A are given by the eigenvalues $\lambda_1, \ldots \lambda_m$ of A. If the system is in a state with density operator ρ at the time of the measurement, the probability of obtaining the outcome λ_i is

$$p_\rho(\lambda_i) = \mathrm{tr}(\rho P_i), \tag{5.42}$$

where P_i is the orthogonal projection operator onto the eigenspace of λ_i. Given that this outcome occurred, the state of the system immediately after the measurement has the density operator

$$\tilde{\rho} = \frac{P_i \rho P_i}{\mathrm{tr}(\rho P_i)}. \tag{5.43}$$

We compute expectation values of an observable A in a state with density operator ρ according to the rule

$$E_\rho(A) = \mathrm{tr}(\rho A) \tag{5.44}$$

and standard deviations according to

$$\Delta_\rho^2(A) = \mathrm{tr}(\rho A^2) - (\mathrm{tr}(\rho A))^2. \tag{5.45}$$

Example 5.2.5 *In a system with Hilbert space $V = \mathbb{C}^3$ the observable A with matrix*

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

relative to the canonical basis is measured when the system is in the state with density operator ρ. The matrix representing ρ relative to the canonical basis is

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

What is the probability of measuring the eigenvalue 2 in a measurement of A? If the eigenvalue 2 is measured, what is the density operator of the system after the measurement? Find the expectation value and standard deviation of A in the state described by ρ.

The observable A was studied in detail in example 3.3.1. There we saw that it has eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 2$, and gave the projectors onto both eigenspaces. Since the observable A and the density operator ρ are given in terms of its matrix relative to the canonical basis, it is easiest to do the entire calculation with matrices. The matrix representation for P_2 is

$$P_2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.46}$$

Then

$$\rho P_2 = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{8} & \frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix}. \tag{5.47}$$

Hence the probability of measuring $\lambda_2 = 2$ is

$$p_\rho(\lambda_2) = \mathrm{tr}(\rho P_2) = \frac{1}{4} + \frac{1}{8} + \frac{1}{4} = \frac{5}{8} \tag{5.48}$$

and the state after the measurement has the density matrix

$$\tilde{\rho} = \frac{8}{5} P_2 \rho P_2 = \frac{8}{5} \begin{pmatrix} \frac{3}{16} & \frac{3}{16} & 0 \\ \frac{3}{16} & \frac{3}{16} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{10} & \frac{3}{10} & 0 \\ \frac{3}{10} & \frac{3}{10} & 0 \\ 0 & 0 & \frac{2}{5} \end{pmatrix}. \tag{5.49}$$

Finally the expectation value of A is

$$E_\rho(A) = \mathrm{tr}(\rho A) = \frac{5}{4}, \tag{5.50}$$

where we used the fact that $A = 2\rho$ and the result (5.48). Since $A^2 = 2A$ we have

$$\Delta_\rho^2 = E_\rho(A^2) - (E_\rho(A))^2 = \frac{5}{2} - \frac{25}{16} = \frac{15}{16}. \tag{5.51}$$

□

Postulate 3": Time evolution is unitary

The time evolution of a closed system is described by a unitary transformation. If the state of the system is given by the density operator ρ at time t and by the density operator ρ' at time t' then there is a unitary operator U so that

$$\rho' = U \rho U^\dagger. \tag{5.52}$$

Example 5.2.6 *The system with Hilbert space $\mathbb{C}^2$ is in the state with density operator*

$$\rho = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{3}{4} \end{pmatrix}$$

at time $t = 0$ seconds. The time evolution operator from time $t = 0$ seconds to $t = 1$ second is

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

Find the density operator of the system at time $t = 1$ second. If the observable

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

is measured at time $t = 1$ second, what is the probability of obtaining the eigenvalue -1?

The density operator at time $t = 1$ second is

$$\rho' = U\rho U^\dagger = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{3}{4} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{3}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}. \tag{5.53}$$

The eigenstate with eigenvalue -1 of the observable A is $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ so that the projector onto this eigenspace has the matrix representation

$$P = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5.54}$$

Therefore the probability of measuring -1 is at time $t = 1$ second is

$$p_{\rho'} = \mathrm{tr}(\rho' P) = \frac{1}{4}. \tag{5.55}$$

□

6 Composite systems

6.1 Tensor products

6.1.1 Basic definitions, notation

Given two vector spaces V and W one can construct a new vector space out of them in two ways. One is called the direct sum and the other the tensor product. In quantum mechanics, the composition of vector spaces via the tensor product plays an important role.

Definition 6.1.1 *(**Tensor product**) Consider two complex vector spaces V and W. The tensor product of V and W is a complex vector space consisting on all linear combinations of elements of the form $|v\rangle \otimes |w\rangle$, where $|v\rangle \in V$ and $|w\rangle \in W$. It satisfies the following properties*

1. $\alpha(|v\rangle \otimes |w\rangle) = (\alpha|v\rangle) \otimes |w\rangle = |v\rangle \otimes (\alpha|w\rangle)$ *for all* $\alpha \in \mathbb{C}, |v\rangle \in V, |w\rangle \in W$.
2. $(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle$ *for all* $|v_1\rangle, |v_2\rangle \in V, |w\rangle \in W$.
3. $|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle$ *for all* $|v\rangle \in V, |w_1\rangle, |w_2\rangle \in W$.

Note that rules 1-3 in the definition are natural rules for a product: Rule 1 is similar to the rule $\alpha AB = (\alpha A)B = A(\alpha B)$ for matrices A, B and a complex number α. Rules 2 and 3 are "distributive laws" which also hold for addition and multiplication of ordinary numbers. Note however, that the product $\otimes$, unlike the product of ordinary numbers, is not commutative

$$|v\rangle \otimes |w\rangle \neq |w\rangle \otimes |v\rangle. \tag{6.1}$$

The following lemma, which we will not prove, summarises important properties of the tensor product.

Lemma 6.1.2 *(**Bases of tensor products**) Let V and W be vector spaces dimensions m and n with bases $D = \{|d_1\rangle, \ldots, |d_m\rangle\}$ and $E = \{|e_1\rangle, \ldots, |e_m\rangle\}$. Then the tensor product $V \otimes W$ has dimension $m \times n$ and the set*

$$P = \{|d_i\rangle \otimes |e_j\rangle\}_{i=1,\ldots,m, \quad j=1,\ldots,n} \tag{6.2}$$

is a basis of $V \otimes W$

When working with tensor products we often adopt a simplified notation, writing $|v\rangle|w\rangle$ or even $|vw\rangle$ for $|v\rangle \otimes |w\rangle$. The latter notation is particularly convenient for tensor products of the basic qubit vector space $\mathbb{C}^2$.

Example 6.1.3 *Write out the basis of tensor product $\mathbb{C}^2 \otimes \mathbb{C}^2$ constructed from the canonical basis of $\mathbb{C}^2$. Also give the dimension and a basis for the triple tensor product $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$.*

Writing out all possible products of $|0\rangle$ and $|1\rangle$ we obtain

$$P = \{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}$$

or, in simplified notation,

$$P = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}. \tag{6.3}$$

Note that there is a natural ordering of the basis elements by interpreting the labels $00, 01, 10, 11$ as binary numbers.

The triple tensor product $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ can be thought of as the tensor product of $\mathbb{C}^2 \otimes \mathbb{C}^2$ with $\mathbb{C}^2$. Thus the product basis is

$$P = \{|00\rangle \otimes |0\rangle, |01\rangle \otimes |0\rangle, |10\rangle \otimes |0\rangle, |11\rangle \otimes |0\rangle, |00\rangle \otimes |1\rangle, |01\rangle \otimes |1\rangle, |10\rangle \otimes |1\rangle, |11\rangle \otimes |1\rangle\}$$

or, in simplified notation,

$$P = \{|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle\}, \tag{6.4}$$

again ordered by interpreting the three digits as a binary representation of the numbers $0, \ldots, 7$. □

It is very important for the applications of tensor products in quantum computing that there are elements in a tensor product space $V \otimes W$ which cannot be written as the tensor product $|v\rangle \otimes |w\rangle$ for $|v\rangle \in V$ and $|w\rangle \in W$.

Definition 6.1.4 (**Factorisable and entangled states**) *Let V, W be vector spaces. A state in $|\psi\rangle \in V \otimes W$ is called a factorisable or product state if it can be written as*

$$|\psi\rangle = |v\rangle \otimes |w\rangle \tag{6.5}$$

for $|v\rangle \in V$ and $|w\rangle \in W$. States which are not product states are called entangled states

Example 6.1.5 *Consider $V = W = \mathbb{C}^2$. Show that the state $|\varphi\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ is entangled and that the state $|\psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$ is factorisable.*

Suppose we could find $|v\rangle = v_0|0\rangle + v_1|1\rangle$ and $|w\rangle = w_0|0\rangle + w_1|1\rangle$ so that $|\varphi\rangle = |v\rangle \otimes |w\rangle$. Then we would have the equality

$$\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) = v_0 w_0|00\rangle + v_0 w_1|01\rangle + v_1 w_0|10\rangle + v_1 w_1|11\rangle.$$

Comparing coefficients we deduce

$$v_0 w_0 = 0, \quad v_0 w_1 = \frac{1}{\sqrt{2}}, \quad v_1 w_0 = \frac{1}{\sqrt{2}}, \quad v_1 w_1 = 0.$$

These equations cannot be satisfied simultaneously. Suppose there was a solution. Then take the product of the second and the third to deduce $v_0 w_1 v_1 w_0 = \frac{1}{2}$; on the other hand taking the product of the first and the fourth we deduce $v_0w_1v_1w_0 = 0$, which is a contradiction. In order to write $|\psi\rangle$ as a product state we need to find v_0, v_1, w_0, w_1 so that

$$|\psi\rangle = v_0w_0|00\rangle + v_0w_1|01\rangle + v_1w_0|10\rangle + v_1w_1|11\rangle.$$

Comparing with the expression for $|\psi\rangle$ we find

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

□

6.1.2 *Inner products*

If both the space V and W are equipped with an inner product, the tensor product space $V \otimes W$ inherits an inner product as follows. If $|v_1\rangle, |v_2\rangle \in V$ and $|w_1\rangle, |w_2\rangle \in W$ then define

$$(|v\rangle \otimes |w\rangle, |v'\rangle \otimes |w'\rangle) = (|v\rangle, |v'\rangle)(|w\rangle, |w'\rangle) = \langle v, |v'\rangle\langle w, |w'\rangle \tag{6.6}$$

In order to compute the inner product of linear combinations

$$|\varphi\rangle = \sum_i \alpha_i|v_i\rangle|w_i\rangle, \qquad |\psi\rangle = \sum_j \beta_j|v'_j\rangle|w'_j\rangle,$$

we extend the above definition linearly in the second argument and conjugate-linearly in the first:

$$\langle\varphi|\psi\rangle = \sum_{i,j} \bar{\alpha}_i\beta_j\langle v_i, |v'_j\rangle\langle w_i, |w'_j\rangle. \tag{6.7}$$

Example 6.1.6 *Consider the two kets $|\varphi\rangle = |00\rangle + |11\rangle$ and $|\psi\rangle = \frac{i}{3}(|00\rangle + |01\rangle + |10\rangle)$. Compute their norms and their inner product.*

$\langle\varphi|\varphi\rangle = 1 + 1 = 2$, so $||\varphi\rangle| = \sqrt{2}$. $\langle\psi|\psi\rangle = \frac{3}{9}$, so $||\varphi\rangle| = \frac{1}{\sqrt{3}}$. Finally $\langle\varphi|\psi\rangle = \frac{i}{3}$. □

6.1.3 *Linear operators*

Suppose $A : V_1 \to V_2$ is a linear operator from the vector space V_1 to the vector space V_2 and $B : W_1 \to W_2$ is a linear operator from the vector space W_1 to the vector space W_2. Then we define a linear operator

$$A \otimes B : V_1 \otimes W_1 \to V_2 \otimes W_2 \tag{6.8}$$

by the rule

$$A \otimes B(|v\rangle \otimes |w\rangle) = A(|v\rangle) \otimes B(|w\rangle) \tag{6.9}$$

and the requirement of linearity i.e.

$$A \otimes B(\alpha|v\rangle \otimes |w\rangle + \beta|v'\rangle \otimes |w'\rangle) = \alpha A(|v\rangle) \otimes B(|w\rangle) + \beta A(|v'\rangle) \otimes B(|w'\rangle) \tag{6.10}$$

Example 6.1.7 *Linear operators $A, B : \mathbb{C}^2 \to \mathbb{C}^2$ are defined via*

$$A|0\,\rangle = \frac{1}{\sqrt{2}}(|0\,\rangle + |1\,\rangle), \quad A|1\,\rangle = \frac{1}{\sqrt{2}}(|1\,\rangle - |0\,\rangle)$$

and

$$B|0\,\rangle = |1\,\rangle, \quad B|1\,\rangle = |0\,\rangle.$$

Find the images of $A \otimes B$ when acting on the ket $|\varphi\,\rangle = |00\,\rangle + |11\,\rangle$.

$$A \otimes B|\varphi\,\rangle = \frac{1}{\sqrt{2}}(|0\,\rangle + |1\,\rangle) \otimes |1\,\rangle + \frac{1}{\sqrt{2}}(|1\,\rangle - |0\,\rangle) \otimes |0\,\rangle = \frac{1}{\sqrt{2}}(|01\,\rangle + |11\,\rangle + |10\,\rangle - |00\,\rangle)$$

□

The matrix representation of an operator of the form $A \otimes B$ is defined as for any linear operator. It takes a particularly simple form in the tensor product basis (6.2). Let us for simplicity consider the situation where $V_1 = V_2$ and $W_1 = W_2$ i.e. $A : V \to V$ and $B : W \to W$. Recall that the matrix representations of A and B relative to the bases D and E are defined via

$$A(|d_i\,\rangle) = \sum_{k=1}^{m} A_{ki}|d_k\,\rangle, \qquad B(|e_j\,\rangle) = \sum_{l=1}^{n} B_{lj}|e_l\,\rangle \tag{6.11}$$

Then, acting on the elements of the basis (6.2) of $V \otimes W$ we find

$$\begin{aligned} A \otimes B|d_i\,\rangle \otimes |e_j\,\rangle &= A(|d_i\,\rangle) \otimes B(|e_j\,\rangle) \\ &= (\sum_{k=1}^{m} A_{ki}|d_k\,\rangle) \otimes (\sum_{l=1}^{n} B_{lj}|e_l\,\rangle) \\ &= \sum_{k=1}^{m}\sum_{l=1}^{n} A_{ki}B_{lj}|d_k\,\rangle \otimes |e_l\,\rangle). \end{aligned} \tag{6.12}$$

This defines the matrix representation of $A \otimes B$ relative to the product basis B (6.2). Although it looks complicated, it has a simple interpretation in terms of the matrix representations of A and B relative to the bases D and E (given before (6.2)). To see this consider the special case $V = W = \mathbb{C}^2$ and recall the basis given in (6.3) and the ordering described there. Suppose the linear maps A and B have the following actions on the basis elements $|0\,\rangle$ and $|1\,\rangle$ of $\mathbb{C}^2$:

$$\begin{aligned} A|0\,\rangle &= A_{00}|0\,\rangle + A_{10}|1\,\rangle \\ A|1\,\rangle &= A_{01}|0\,\rangle + A_{11}|1\,\rangle \\ B|0\,\rangle &= B_{00}|0\,\rangle + B_{10}|1\,\rangle \\ B|1\,\rangle &= B_{01}|0\,\rangle + B_{11}|1\,\rangle \end{aligned} \tag{6.13}$$

so that the matrix representations relative to the canonical basis are

$$A = \begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix} \qquad B = \begin{pmatrix} B_{00} & B_{01} \\ B_{10} & B_{11} \end{pmatrix}. \tag{6.14}$$

The 4×4-matrix representing $A \otimes B$ relative to the canonical basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ is then

$$A \otimes B = \begin{pmatrix} A_{00}B_{00} & A_{00}B_{01} & A_{01}B_{00} & A_{01}B_{01} \\ A_{00}B_{10} & A_{00}B_{11} & A_{01}B_{10} & A_{01}B_{11} \\ A_{10}B_{00} & A_{10}B_{01} & A_{11}B_{00} & A_{11}B_{01} \\ A_{10}B_{10} & A_{10}B_{11} & A_{11}B_{10} & A_{11}B_{11} \end{pmatrix}. \tag{6.15}$$

We obtain this matrix by writing down the matrix A and multiplying every matrix element of A with a copy of the matrix B:

$$A \otimes B = \begin{pmatrix} A_{00}B & A_{01}B \\ A_{10}B & A_{11}B \end{pmatrix}. \tag{6.16}$$

Example 6.1.8 *If*

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

and

$$B = \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}$$

find $A \otimes B$ *and* $B \otimes A$

Following the above rule, we find

$$A \otimes B = \begin{pmatrix} 3 & 4 & 6 & 8 \\ 5 & 6 & 10 & 12 \\ -3 & -4 & 3i & 4i \\ -5 & -6 & 5i & 6i \end{pmatrix}$$

and

$$B \otimes A = \begin{pmatrix} 3 & 6 & 4 & 8 \\ -3 & 3i & -4 & 4i \\ 5 & 10 & 6 & 12 \\ -5 & 5i & -6 & 6i \end{pmatrix}$$

In particular $A \otimes B \neq B \otimes A$. □

Consider now a general linear map

$$C : V \otimes W \to V \otimes W.$$

Its matrix representation relative to the product basis $\{|d_i\rangle \otimes |e_j\rangle\}_{i=1,\dots,m, j=1\dots,n}$ is defined via

$$C|d_k\rangle \otimes |e_l\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{ikjl} |d_i\rangle \otimes |e_j\rangle. \tag{6.17}$$

In the case of $V = W$ being two dimensional we obtain the matrix

$$C = \begin{pmatrix} C_{1111} & C_{1112} & C_{1211} & C_{1212} \\ C_{1121} & C_{1122} & C_{1221} & C_{1222} \\ C_{2111} & C_{2112} & C_{2211} & C_{2212} \\ C_{2121} & C_{2122} & C_{2221} & C_{2222} \end{pmatrix}. \tag{6.18}$$

Such matrices need not be of the product form $A \otimes B$ - there are “entangled” matrices which cannot be factorised, just like there are entangled states in $V \otimes W$. On sheet 4 you are asked to show that a given matrix representation of a linear map cannot be a tensor product of two matrices.

As for any pair of linear maps, we can compose two linear maps $C, D : V \otimes W \to V \otimes W$. The matrix of the product CD is

$$(CD)_{ikjl} = \sum_{p=1}^{m} \sum_{q=1}^{n} C_{ipjq} D_{pkql}. \tag{6.19}$$

Finally we define the trace as for any matrix.

$$\mathrm{tr}(C) = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{iijj} \tag{6.20}$$

If C is of the form $A \otimes B$ we have the useful formula

$$\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\mathrm{tr}(B). \tag{6.21}$$

This follows directly from the definition

$$\mathrm{tr}(A \otimes B) = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ii} B_{jj} = \sum_{i=1}^{m} A_{ii} \sum_{j=1}^{n} B_{jj} = \mathrm{tr}(A)\mathrm{tr}(B). \tag{6.22}$$

In addition, we can use the structure of the tensor product $V \otimes W$ to define partial traces.

Definition 6.1.9 *Let $C : V \otimes W \to V \otimes W$ be a linear map with matrix representation C_{ikjl} relative to the tensor product basis $\{|d_i\rangle \otimes |e_j\rangle\}_{i=1,\ldots,m,j=1\ldots,n}$. Then the partial trace of C over V is the linear map $C^W = \mathrm{tr}_V(C) : W \to W$ with matrix elements*

$$C^W_{jl} = \sum_{i=1}^{m} C_{iijl} \tag{6.23}$$

relative to the basis $\{|d_j\rangle\}_{j=1,\ldots,n}$ of W. Similarly the partial trace of C over W is the linear map $C^V = \mathrm{tr}_W(C) : V \to V$ with matrix elements

$$C^V_{ik} = \sum_{j=1}^{n} C_{ikjj} \tag{6.24}$$

relative to the basis $\{|e_i\rangle\}_{i=1,\ldots,m}$ of V.

The following lemma is very useful for computing partial traces of tensor products.

Lemma 6.1.10 *For any linear map of the product form $A \otimes B : V \otimes W \to V \otimes W$*

$$(A \otimes B)^V = \mathrm{tr}(B)\, A \qquad (A \otimes B)^W = \mathrm{tr}(A)\, B. \tag{6.25}$$

Proof: Using the bases D and E of V and W defined in (6.2) to define the matrix representations of A and B we have

$$(A \otimes B)^V_{ik} = \sum_{j=1}^{n} A_{ik} B_{jj} = \mathrm{tr}(B) A_{ik}.$$

Since $(A \otimes B)^V$ and $\mathrm{tr}(B)A$ have the same matrix representation with respect the basis D of V, they are equal as linear maps. Similarly

$$(A \otimes B)^W_{jl} = \sum_{i=1}^{m} A_{ii} B_{jl} = \mathrm{tr}(A) B_{jl},$$

showing that $(A \otimes B)^W$ and $\mathrm{tr}(A)B$ are the same linear map. $\Box$.

Example 6.1.11 *(i) For the matrices A and B from example 6.1.8 compute* $\mathrm{tr}(A), \mathrm{tr}(B), \mathrm{tr}(A \otimes B)$ *and* $\mathrm{tr}(B \otimes A)$*, and check the formula* (6.21)*.*
(ii) Consider the operator $C : V \otimes W \to V \otimes W$, where $V = W = \mathbb{C}^2$, with matrix representation

$$C = \begin{pmatrix} i & 1 & 2 & -1 \\ 1 & -i & 1 & 0 \\ -i & 0 & -i & i \\ i & -1 & i & 1 \end{pmatrix}$$

relative to the canonical basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ of $\mathbb{C}^2 \otimes \mathbb{C}^2$. Compute its partial trace C^W with respect to the first component V of the tensor product and its partial trace C^V with respect to the second component W of the tensor product. Also compute its full trace. Check that $\mathrm{tr}_V(C^V) = \mathrm{tr}_W(C^W) = \mathrm{tr}_{V\otimes W}(C)$

(i) We find $\mathrm{tr}(A) = 1+i$, $\mathrm{tr}(B) = 9$. Also $\mathrm{tr}(A \otimes B) = 3+6+3i+6i = 9+9i = \mathrm{tr}(B \otimes A) = \mathrm{tr}(A)\mathrm{tr}(B)$.
(ii) We find

$$C^W = \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix} + \begin{pmatrix} -i & i \\ i & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1+i \\ 1+i & 1-i \end{pmatrix}$$

and

$$C^V = \begin{pmatrix} i-i & 2-0 \\ -i-1 & -i+1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -1-i & 1-i \end{pmatrix}$$

so that $\mathrm{tr}_W(C^W) = 1 - i = \mathrm{tr}_V(C^V) = \mathrm{tr}_{V\otimes W}(C)$ $\Box$

Lemma 6.1.12 *Consider two Hilbert spaces V and W and the tensor product $V \otimes W$ equipped with its canonical inner product. Given two linear operators $A : V \to V$ and $B : W \to W$ with adjoints $A^\dagger$ and $B^\dagger$, the adjoint of the tensor product of A and B is*

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger. \tag{6.26}$$

Proof: Let $|v_1\rangle, |v_2\rangle \in V$ and $|w_1\rangle, |w_2\rangle \in W$. Then

$$\begin{aligned}(|v_1\rangle \otimes |w_1\rangle, A \otimes B|v_2\rangle \otimes |w_2\rangle) &= (|v_1\rangle, A|w_1\rangle)(|v_2\rangle, B|w_2\rangle) \\ &= (A^\dagger|v_1\rangle, |w_1\rangle)(B^\dagger|v_2\rangle, |w_2\rangle) \\ &= (A^\dagger \otimes B^\dagger|v_1\rangle \otimes |w_1\rangle, |v_2\rangle \otimes |w_2\rangle) \end{aligned} \tag{6.27}$$

Since this holds for all product states $|v_1\rangle \otimes |w_1\rangle, |v_2\rangle \otimes |w_2\rangle \in V \otimes W$, and since the product states span $V \otimes W$, we deduce that the adjoint of $A \otimes B$ is $A^\dagger \otimes B^\dagger$. □.

Example 6.1.13 *Suppose $A : V \to V$ and $B : W \to W$ are Hermitian operators. Since A and B are Hermitian there exists a basis of eigenvectors$\{|d_i\rangle\}_{i=1,\ldots m}$ of A for V and a basis of eigenvectors $\{|e_j\rangle\}_{j=1,\ldots n}$ of B for W. Denote the corresponding eigenvalues by λ_i and μ_j i.e.*

$$A|d_i\rangle = \lambda_i|d_i\rangle \quad B|e_j\rangle = \mu_j|e_j\rangle. \tag{6.28}$$

Show that $\{|d_i\rangle \otimes |e_j\rangle\}_{i=1,\ldots m, j=1,\ldots n}$ is a basis of eigenvectors of $A \otimes B$ for $V \otimes W$

Since

$$A \otimes B\,|d_i\rangle \otimes |e_j\rangle = A|d_i\rangle \otimes B|e_j\rangle = \lambda_i\mu_j|d_i\rangle \otimes |e_j\rangle$$

the vectors $\{|d_i\rangle \otimes |e_j\rangle\}$ are eigenvectors with eigenvalues $\lambda_i\mu_j$. They form a basis of $V \otimes W$ by Lemma 6.1.2. □

6.2 Quantum mechanics of composite systems

Postulate 4: Composite systems *The state space of a composite physical system is the tensor product of the component physical systems. If we have N systems with label $i = 1, \ldots, N$, then if the i-the system is prepared in the pure state $|\psi_i\rangle$, the state of the total system is the pure state $|\psi_1\rangle \otimes \ldots |\psi_N\rangle$.*

In order to investigate the properties of composite systems we apply postulate 2 (observables and measurement) and postulate 3 (time evolution) to tensor product spaces.

Example 6.2.1 *(***Measurement in composite systems***)*

Consider the system made by composing two systems with Hilbert space $\mathbb{C}^2$. Suppose the system is in the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{6.29}$$

when the observable $A = \sigma_1 \otimes \sigma_3$ is measured. Show that A has eigenvalues ± 1 and find the probability of measuring the eigenvalue 1. Find the state after the measurement and the expectation value of A in the state $|\psi\rangle$.

We know from Sect. 4 that σ_3 has eigenvalues 1 and -1 with eigenvectors $|0\rangle$ and $|1\rangle$, and that σ_1 has eigenvalues 1 and -1 with eigenvectors $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Hence

$$\begin{aligned} |v\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle), \\ |w\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \otimes |1\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |11\rangle) \end{aligned} \tag{6.30}$$

are eigenstates of $\sigma_1 \otimes \sigma_3$ for the eigenvalue 1. The projector onto the eigenspace spanned by $|v\rangle$ and $|w\rangle$ is

$$P_1 = |v\rangle\langle v| + |w\rangle\langle w|$$

and therefore the probability of measuring the outcome 1, given that the system is in the state $|\psi\rangle$, is

$$p_\psi(1) = |\langle\psi|v\rangle|^2 + |\langle\psi|w\rangle|^2 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

The state after the measurement is

$$\begin{aligned} \frac{1}{\sqrt{p_\psi(1)}} P_1|\psi\rangle &= \sqrt{2}\left(\frac{1}{2\sqrt{2}}(|00\rangle + |10\rangle) - \frac{1}{2\sqrt{2}}(|01\rangle - |11\rangle)\right) \\ &= \frac{1}{2}(|00\rangle + |10\rangle - |01\rangle + |11\rangle). \end{aligned} \tag{6.31}$$

The expectation value of A is

$$\langle\psi|A|\psi\rangle = \frac{1}{\sqrt{2}}(\langle 00| + \langle 11|)(|10\rangle - |01\rangle) = 0 \tag{6.32}$$

$\square$

Example 6.2.2 *(***Partial measurement***) Consider again the system made by composing two systems with Hilbert space $\mathbb{C}^2$, and suppose the system is in the state*

$$|\varphi\rangle = \frac{1}{\sqrt{3}}(|00\rangle + |01\rangle + |10\rangle). \tag{6.33}$$

Give a precise mathematical formulation and then an answer for the question "what is the probability that the first qubit is in the state $|0\rangle$?".

In order to answer this question in the formalism of quantum mechanics we need to give an operator such that one of its eigenspaces contains all states of the form $|0\rangle \otimes |\psi\rangle$, where $|\psi\rangle$ is an arbitrary state of the second qubit. The operator

$$P = |0\rangle\langle 0| \otimes I \tag{6.34}$$

is a projection operator since $P^2 = |0\rangle\langle 0|0\rangle\langle 0| \otimes I = P$. According to example 6.1.13 its eigenspace for eigenvalue 1 consists of all states of the form $|0\rangle \otimes |v\rangle$, were $|v\rangle$ is an arbitrary state in $\mathbb{C}^2$ and its eigenspace for eigenvalue 0 consists of all states of the form $|1\rangle \otimes |v\rangle$, were $|v\rangle$ is again an arbitrary state in $\mathbb{C}^2$. Hence we formulate the question "what is the probability that the first qubit is in the state $|0\rangle$?" as "what is the probability of measuring the eigenvalue 1 of the operator P?". The answer is

$$p_\varphi(P=1) = \langle\varphi|P|\varphi\rangle = \frac{1}{3}(\langle 00| + \langle 01| + \langle 10|)(|00\rangle + |01\rangle) = \frac{2}{3}$$

□

In studying composite systems we often need to adress questions which only concern one of the subsystems that make the system, as illustrated by the previous example. There is a systematic way of answering such questions which works as follows. Consider a system with Hilbert space $V \otimes W$ and the observable $A : V \to V$ of the subsystem with Hilbert space V. We would like to compute the possible outcomes, probabilities and expectation values in measurements of A, but we only have density operator of the total system $\rho : V \otimes W \to V \otimes W$. In order to compute expectation values of A we "embed" the observable A into the total system and compute with $\tilde{A} = A \otimes I$. This is what we did in the example above. However, using the bases D and E of V and W as before, and working with the matrix representations of A and ρ we have

$$\begin{aligned}
\mathrm{tr}_{V\otimes W}(\rho\tilde{A}) &= \sum_{i,p=1}^{m}\sum_{j,q=1}^{n} \rho_{ipjq} A_{pi} \delta_{qj} \\
&= \sum_{i,p=1}^{m}\sum_{j}^{n} \rho_{ipjj} A_{pi} \\
&= \mathrm{tr}_W(\rho^V A).
\end{aligned} \tag{6.35}$$

In other words, the quantum mechanical predicitions for measurements of observables of the subsystem V determined by the partial trace ρ^V of ρ.

Definition 6.2.3 (**Reduced density operator**) *Let ρ be a density operator for the composite system with Hilbert space $V \otimes W$. Then the reduced density operators for the subsystems V and W are given by the partial traces ρ^V and ρ^W of ρ as defined in 6.1.9.*

Example 6.2.4 (**Expectation values**) *The density operator of the two-qubit system with Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2$ is given by*

$$\rho = \frac{1}{4}(I + \sigma_1) \otimes (I + \sigma_2).$$

Find the expectation values of the observables

$$C = \sigma_3 \otimes \sigma_3 + \sigma_1 \otimes I$$

and

$$D = \sigma_2 \otimes I.$$

Also find the reduced density operator for the first qubit and compute the expectation value of σ_2 in the first qubit.

Since

$$\rho C = \frac{1}{4}(\sigma_3 - i\sigma_2) \otimes (\sigma_3 + i\sigma_1) + \frac{1}{4}(\sigma_1 + I) \otimes (I + \sigma_2)$$

we have

$$\mathrm{tr}(\rho C) = \frac{1}{4}\mathrm{tr}(I)\mathrm{tr}(I) = 1,$$

where we used that all Pauli matrices are traceless. Similarly

$$\mathrm{tr}(\rho D) = \frac{1}{4}\mathrm{tr}\left((\sigma_2 + i\sigma_3) \otimes (I + \sigma_2)\right) = 0$$

The partial trace of ρ over the second qubit gives

$$\rho^V = \frac{1}{2}(I + \sigma_1)$$

and hence

$$\mathrm{tr}(\rho^V \sigma_2) = \frac{1}{2}\mathrm{tr}(\sigma_2 + i\sigma_3) = 0,$$

which agrees with $\mathrm{tr}(\rho D)$, as it should. □

Example 6.2.5 *(***Time evolution in composite systems***) The time evolution of a state $|\psi(t)\rangle$ in $\mathbb{C}^2 \otimes \mathbb{C}^2$ is given by the Schrödinger equation*

$$i\hbar\frac{d}{dt}|\psi\rangle = H|\psi\rangle \tag{6.36}$$

where the Hamiltonian H is given by

$$H = \sigma_1 \otimes \sigma_3. \tag{6.37}$$

Find the time evolution operator. If the state of the system at time $t = 0$ is $|\psi_0\rangle = |11\rangle$ find the state of the system at time t.

Since H satisfies $H^2 = I \otimes I$ we can compute the time evolution operator as in(3.40):

$$U(t) = \exp(-i\frac{t}{\hbar}H) = \cos(\frac{t}{\hbar})I \otimes I - i\sin(\frac{t}{\hbar})\sigma_1 \otimes \sigma_3.$$

Hence the the state at time t is

$$|\psi(t)\rangle = \cos(\frac{t}{\hbar})I \otimes I|11\rangle - i\sin(\frac{t}{\hbar})\sigma_1 \otimes \sigma_3|11\rangle = \cos(\frac{t}{\hbar})|11\rangle + i\sin(\frac{t}{\hbar})|01\rangle.$$

□

6.3 Schmidt decomposition and purification

We have seen that states in tensor product spaces $V \otimes W$ are either product states or entangled. In this section we give an algorithm for determining if a state is a product state or entangled, and introduce a measure for the degree of "entangledness" of entangled states. We begin with a technical lemma. It generalises the representation of a Hermitian matrix $A = UDU^{-1}$ in terms of a real, diagonal matrix D and a unitary matrix U.

Lemma 6.3.1 (**Singular value decomposition**) *Let S be a complex $n\times n$ matrix. Then there exist unitary $n\times n$ matrices U and $\tilde{U}$ and a diagonal matrix D with real, non-negative diagonal entries such that*

$$S = UD\tilde{U}. \tag{6.38}$$

The eigenvalues of D (but not their ordering) are uniquely determined by S.

We omit the proof, which is a little technical but not difficult - see e.g. Nielsen and Chuang, Quantum Computation and quantum Information, page 78 ff.

Theorem 6.3.2 (**Schmidt decomposition**) *Suppose V and W are Hilbert spaces of dimension n and $|\psi\rangle \in V\otimes W$ has norm 1. Then there exist orthonormal bases $\{|v_1\rangle,\dots,|v_n\rangle\}$ and $\{|w_1\rangle,\dots,|w_n\rangle\}$ of V and, respectively, W such that*

$$|\psi\rangle = \sum_{k=1}^{n} \lambda_k |v_k\rangle \otimes |w_k\rangle, \tag{6.39}$$

where the coefficients λ_i are non-negative real numbers satisfying

$$\sum_{k=1}^{n} \lambda_k^2 = 1. \tag{6.40}$$

Proof: Let $D=\{|d_1\rangle,\dots,|d_n\rangle\}$ and $E=\{|e_1\rangle,\dots,|e_n\rangle\}$ be bases of V and W. Then a given state $|\psi\rangle$ can be expanded

$$|\psi\rangle = \sum_{i,j=1}^{n} S_{ij}|d_i\rangle \otimes |e_j\rangle, \tag{6.41}$$

with complex numbers S_{ij}, $i,j=1,\dots,n$. Now decompose the complex $n\times n$ matrix S according to the singular value decomposition (6.38) so that

$$S_{ij} = \sum_{k,l=1}^{n} U_{ik} D_{kl} \tilde{U}_{lj}$$

for unitary matrices U and $\tilde{U}$ and a positive, diagonal matrix D. Writing the matrix elements of D as

$$D_{kl} = \delta_{kl}\lambda_k$$

for non-negative numbers λ_k the expansion (6.41) becomes

$$|\psi\rangle = \sum_{i,j,k=1}^{n} \lambda_k U_{ik}|d_i\rangle \otimes \tilde{U}_{kj}|e_j\rangle. \tag{6.42}$$

Now define the Schmidt basis

$$|v_k\rangle = \sum_{i=1}^{n} U_{ik}|d_i\rangle, \qquad |w_k\rangle = \sum_{j=1}^{n} \tilde{U}_{kj}|e_j\rangle. \tag{6.43}$$

It follows from the unitarity of U and $\tilde{U}$ that

$$\langle v_k | v_l \rangle = \sum_{i=1}^{n} \bar{U}_{ik} U_{il} = \delta_{kl}$$

and

$$\langle w_k | w_l \rangle = \sum_{j=1}^{n} \bar{\tilde{U}}_{kj} \tilde{U}_{lj} = \delta_{kl},$$

so that (6.42) gives the promised expansion (6.39) in terms of orthonormal states and non-negative numbers λ_k. The condition (6.40) follows from the normalisation of $|\psi\rangle$:

$$1 = \langle \psi | \psi \rangle = \sum_{k,l=1}^{n} \lambda_k \lambda_l \langle v_l | v_k \rangle \langle w_l | w_k \rangle = \sum_{k=1}^{n} \lambda_k^2. \tag{6.44}$$

□

Definition 6.3.3 *(***Schmidt coefficients and Schmidt number***) The real numbers in the decomposition* (6.39) *are called the Schmidt coefficients of the state* $|\psi\rangle$*. The number of non-zero Schmidt coefficients is called the Schmidt number of the state* $|\psi\rangle$*.*

Lemma 6.3.4 *The Schmidt number and the Schmidt coefficients of a state* $|\psi\rangle$ *in the tensor product* $V \otimes W$ *are well-defined. Moreover, a pure state* $|\psi\rangle$ *of a composite system is a product state if and only if its Schmidt number is* 1

Proof: If we had expanded the state $|\psi\rangle$ in different bases D' and E' of V and W we would have obtained a matrix S' which is related to the matrix S in (6.41) via

$$S' = TSR$$

where T and R are unitary matrices. We thus obtain a singular value decomposition of

$$S' = U' D \tilde{U}'$$

where $U' = TU$ and $\tilde{U}' = \tilde{U}R$, but D is unchanged. According to lemma 6.3.1 the eigenvalues of D in any singular value decomposition are the same. In particular, the number of non-zero eigenvalues in any Schmidt decomposition of a given state $|\psi\rangle$ is therefore the same.

If $|\psi\rangle$ is a product state then the formula

$$|\psi\rangle = |v\rangle \otimes |w\rangle \tag{6.45}$$

is a Schmidt decomposition of $|\psi\rangle$ with one Schmidt coefficient equal to 1 and the others 0. The Schmidt number of the state is therefore 1. Conversely, if we know that the Schmidt number of a given state is 1 we deduce from (6.40) that the only non-zero Schmidt coefficient is 1, and that the Schmidt decomposition takes the form (6.45).

When $V = W$ and the matrix S with matrix elements S_{ij} defined via the equation (6.41) is Hermitian, we can find the Schmidt decomposition by diagonalising S. Suppose we have

$$S = UDU^{-1}$$

where U is unitary and D is diagonal. Provided that the eigenvalues of D are non-negative, we obtain a Schmidt basis (6.43) via

$$|v_k\rangle = \sum_{i=1}^{n} U_{ik}|d_i\rangle, \qquad |w_k\rangle = \sum_{j=1}^{n} \bar{U}_{jk}|e_j\rangle. \tag{6.46}$$

where we used that $U^{-1} = \bar{U}^t$ for unitary matrices. If some of the eigenvalues λ_k of D are negative, we multiply the corrsponding basis vectors w_k by -1.

Example 6.3.5 *Compute the Schmidt number of the state*

$$|\psi\rangle = \frac{1}{4}(|00\rangle - \sqrt{3}|01\rangle - \sqrt{3}|10\rangle + 3|11\rangle). \tag{6.47}$$

Reading off the matrix S in the expansion

$$|\psi\rangle = S_{00}|00\rangle + S_{01}|01\rangle + S_{10}|10\rangle + S_{11}|11\rangle \tag{6.48}$$

we find

$$S = \frac{1}{4}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{pmatrix}.$$

Since S is Hermitian, we find its singular value decomposition by diagonalising it. The eigenvalues are 1 and 0, and therefore the Schmidt number is 1. The eigenvector for the eigenvalue 1 is $\frac{1}{2}(|0\rangle - \sqrt{3}|1\rangle)$ and therefore the state factorises

$$|\psi\rangle = \frac{1}{2}(|0\rangle - \sqrt{3}|1\rangle) \otimes \frac{1}{2}(|0\rangle - \sqrt{3}|1\rangle).$$

□.

The main application of the Schmidt decomposition is the following theorem, which links the notion of factorisable states in a composite system with that of pure states in the constituent systems.

Theorem 6.3.6 *The state $|\psi\rangle$ in the composite system with Hilbert space $V \otimes W$ is a factorisable state if and only if the reduced density matrices ρ^V and ρ^W obtained form the density operator $\rho = |\psi\rangle\langle\psi|$ are pure states.*

Note that, because of the formulation "if and only if" the theorem also says that the state $|\psi\rangle$ in the composite system is entangled if and only if the reduced density matrices ρ^V and ρ^W obtained form $\rho = |\psi\rangle\langle\psi|$ are mixed states.

Proof: Starting from the Schmidt decomposition of the state $|\psi\rangle$

$$|\psi\rangle = \sum_{k=1}^{n} \lambda_k |v_k\rangle \otimes |w_k\rangle,$$

the density operator is

$$\rho = |\psi\rangle\langle\psi| = \sum_{l,k=1}^{n} \lambda_k \lambda_l (|v_k\rangle\langle v_l|) \otimes (|w_k\rangle\langle w_l|).$$

Hence

$$\rho^V = \mathrm{tr}_W(\rho) = \sum_{i=1}^{n} \sum_{l,k=1}^{n} \lambda_k \lambda_l (|v_k\rangle\langle v_l|) \otimes (\langle w_i|w_k\rangle\langle w_l|w_i\rangle) = \sum_{i=1}^{n} \lambda_i^2 |v_i\rangle\langle v_i|$$

and

$$\rho^W = \mathrm{tr}_V(\rho) = \sum_{i=1}^{n} \sum_{l,k=1}^{n} \lambda_k \lambda_l (\langle v_i|v_k\rangle\langle v_l|v_i\rangle) \otimes (|w_k\rangle\langle w_l|) = \sum_{i=1}^{n} \lambda_i^2 |w_i\rangle\langle w_i|,$$

where we used the orthonormality conditions

$$\langle v_i|v_k\rangle = \delta_{ik}, \quad \langle w_l|w_i\rangle = \delta_{li}.$$

Thus the reduced density operators describe pure states if and only one of the λ_i is 1 and all the others 0. This holds if and only if the state $|\psi\rangle$ is a product state. □

Example 6.3.7 *Consider the so-called Bell states in* $\mathbb{C}^2 \otimes \mathbb{C}^2$*:*

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \qquad |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

Give their Schmidt decomposition and compute their reduced density operators for the first and second qubit.

The state $|\Phi^-\rangle$ is almost in the form required for Schmidt decomposition, except for the minus sign in front of $|11\rangle$. Factoring $-1 = i \times i$ we write

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}|0\rangle \otimes |0\rangle + \frac{1}{\sqrt{2}}(i|1\rangle) \otimes (i|1\rangle).$$

Thus with $|v_1\rangle = |w_1\rangle = |0\rangle$ and $|v_2\rangle = |w_2\rangle = i|1\rangle$ we have a decomposition of the type (6.39) with Schmidt coefficients $\frac{1}{\sqrt{2}}$ and $\frac{1}{\sqrt{2}}$. The reduced density operator for both qubits is

$$\rho^V_{\Phi^-} = \rho^W_{\Phi^-} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$$

(note that the i drops out in the density operator!) so that the matrix representation with respect to the canonical basis is

$$\rho^V_{\Phi^-} = \rho^W_{\Phi^-} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

To find the Schmidt decomposition of the state $|\Psi^+\rangle$ we write it as

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}|0\rangle \otimes |1\rangle + \frac{1}{\sqrt{2}}(i|1\rangle) \otimes (i|0\rangle) \tag{6.49}$$

Thus with $|v_1\rangle = |0\rangle$ and $|v_2\rangle = i|1\rangle$ but $|w_1\rangle = |1\rangle$ and $|w_2\rangle = i|0\rangle$ we have a decomposition of the type (6.39) and the Schmidt coefficients are again $\frac{1}{\sqrt{2}}$ and $\frac{1}{\sqrt{2}}$. The reduced density operators are

$$\rho^V_{\Psi^-} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}I$$

and

$$\rho^W_{\Psi^-} = \frac{1}{2}|1\rangle\langle 1| + \frac{1}{2}|0\rangle\langle 0| = \frac{1}{2}I,$$

where factors of i have again cancelled out. □

Both Bell states studied in the example have 2=dim($\mathbb{C}^2$) Schmidt coefficients so that their Schmidt number takes its largest possible value. Moreover, the Schmidt coefficients are all the same. Hence both states are "maximally different" from a product state, whose Schmidt coefficients would be 1 and 0, The Bell states are therefore also called "maximally entangled". The example shows that the reduced density operators constructed from pure but maximally entangled states are "maximally mixed": since the density operator is proportional to the identity, all states are assigned the same probability in the ensemble interpretation of the density operator. We will return to this point in the Sect. 6.4.

We have seen that reduced density operators obtained from entangled states in tensor product are mixed states in each of the constituent spaces. It is possible to reverse this process i.e. to start with a density operator in a system with Hilbert space V and to give a pure (but entangled) state in a tensor product $V \otimes W$ such that the reduced density operator gives the original state. This process is called purification, and the Hilbert space W used for the defining the composite system is called the auxiliary space.

Definition 6.3.8 (**Purification**) *Suppose ρ is a density operator in the system with Hilbert space V. A pure state $|\psi\rangle$ in the tensor product $V \otimes W$, where W is called the auxiliary Hilbert space, is called the purification of ρ if*

$$\rho = \rho^V_\psi, \quad \textit{where} \quad \rho = |\psi\rangle\langle\psi|. \tag{6.50}$$

Example 6.3.9 *Show that the state in $\mathbb{C}^2$ with density matrix*

$$\rho = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

relative to the canonical basis of $\mathbb{C}^2$ is a mixed state and find the purification of it in the Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2$.

Since

$$\mathrm{tr}(\rho^2) = \frac{10}{16} < 1$$

the state is mixed. In order to find a purification, we diagonlise ρ. It has eigenvalues $\lambda_1 = \frac{3}{4}$ and $\lambda_2 = \frac{1}{4}$ with normalised eigenstates $|v_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|v_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Hence

$$\rho = \frac{3}{4}|v_1\rangle\langle v_1| + \frac{1}{4}|v_2\rangle\langle v_2|$$

and the purification is given by ρ_ψ where

$$|\psi\rangle = \sqrt{\lambda_1}|v_1\rangle \otimes |v_1\rangle + \sqrt{\lambda_2}|v_2\rangle \otimes |v_2\rangle = \frac{\sqrt{3}}{2}|v_1\rangle \otimes |v_1\rangle + \frac{1}{2}|v_2\rangle \otimes |v_2\rangle.$$

$\square$

6.4 The EPR (thought) experiment

The EPR thought experiment demonstrates that the result of a measurement performed on one part of a quantum system can have an instantaneous effect on the result of a measurement performed on another part, regardless of the distance separating the two parts. This appears to violate Einstein's theory of special relativity, which states that information cannot be transmitted faster than the speed of light. "EPR" stands for Albert Einstein, Boris Podolsky, and Nathan Rosen, who introduced the thought experiment in a 1935 paper to argue that quantum mechanics is not a complete physical theory. The version of the thought experiment we will discuss here is due to David Bohm.

The EPR thought experiment is often referred to as a paradox (not by the authors!). It is a paradox in the following sense: if one takes quantum mechanics and adds some seemingly reasonable conditions (referred to as "locality", "realism", and "completeness"), then one obtains a contradiction. However, quantum mechanics by itself does not appear to be internally inconsistent, nor does it contradict relativity. As a result of further theoretical and experimental developments since the original EPR paper, most physicists today regard the EPR paradox as an illustration of how quantum mechanics violates classical intuition, and not as an indication that quantum mechanics is fundamentally flawed.

In Bohm's version of the EPR thought experiments, two spin 1/2 particles with Hilbert space are prepared in the state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right). \tag{6.51}$$

An important property of this state is that measurements of the particles' spins along any axis are correlated. If one particle is found to be in the "spin up" state along any axis $\boldsymbol{k}$, the other must be in the "spin down" state. Before we enter the discussion of the experiment, we make this statement mathematically precise, and prove it.

Lemma 6.4.1 *Let $\boldsymbol{k}$ be a unit vector in $\mathbb{R}^3$, and recall that $\frac{\hbar}{2}\boldsymbol{k}\cdot\boldsymbol{\sigma}$ is the spin operator along the axis $\boldsymbol{k}$. Then the state $|\Psi^-\rangle$ satisfies*

$$(\boldsymbol{k}\cdot\boldsymbol{\sigma} \otimes I + I \otimes \boldsymbol{k}\cdot\boldsymbol{\sigma})|\Psi^-\rangle = 0. \tag{6.52}$$

Proof: We first note that

$$\sigma_3 \otimes I(|01\rangle - |10\rangle) = |01\rangle + |10\rangle \tag{6.53}$$

and

$$I \otimes \sigma_3(|01\rangle - |10\rangle) = -|01\rangle - |10\rangle \tag{6.54}$$

so that

$$(\sigma_3 \otimes I + I \otimes \sigma_3)|\Psi^-\rangle = 0. \tag{6.55}$$

Hence the claim holds for $\boldsymbol{k} = (0,0,1)^t$. Now recall from Eq. (4.32) that the operator for the spin along an arbitrary axis

$$\boldsymbol{k}(\theta,\phi) = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}, \tag{6.56}$$

with $\theta \in [0,\theta)$ and $\phi \in [0,2\pi)$ is proportional to

$$\boldsymbol{k}\cdot\boldsymbol{\sigma} = U(\theta,\phi)\sigma_3 U^{-1}(\theta,\phi), \tag{6.57}$$

where the unitary operator $U(\theta,\phi)$ is

$$U(\theta,\phi) = e^{-\frac{i}{2}\phi\sigma_3}e^{-\frac{i}{2}\theta\sigma_2}. \tag{6.58}$$

However, for any unitary operator

$$U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{6.59}$$

we have

$$U \otimes U|01\rangle = \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} \otimes \begin{pmatrix} \beta \\ \delta \end{pmatrix} = \alpha\beta|00\rangle + \alpha\delta|01\rangle + \gamma\beta|10\rangle + \gamma\delta|11\rangle \tag{6.60}$$

and

$$U \otimes U|10\rangle = \begin{pmatrix} \beta \\ \delta \end{pmatrix} \otimes \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = \alpha\beta|00\rangle + \beta\gamma|01\rangle + \delta\alpha|10\rangle + \gamma\delta|11\rangle \tag{6.61}$$

so that

$$U \otimes U(|01\rangle - |10\rangle) = (\alpha\delta - \gamma\beta)(|01\rangle - |10\rangle) = \det(U)(|01\rangle - |10\rangle). \tag{6.62}$$

we therefore have

$$\begin{aligned} & (\sigma_3 \otimes I + I \otimes \sigma_3)|\Psi^-\rangle = 0 \\ \Rightarrow\ & U \otimes U(\sigma_3 \otimes I + I \otimes \sigma_3)\left(U^{-1} \otimes U^{-1}\right)(U \otimes U)|\Psi^-\rangle = 0 \\ \Rightarrow\ & (\boldsymbol{k}\cdot\boldsymbol{\sigma} \otimes I + I \otimes \boldsymbol{k}\cdot\boldsymbol{\sigma})\det(U)|\Psi^-\rangle = 0 \\ \Rightarrow\ & (\boldsymbol{k}\cdot\boldsymbol{\sigma} \otimes I + I \otimes \boldsymbol{k}\cdot\boldsymbol{\sigma})|\Psi^-\rangle = 0, \end{aligned} \tag{6.63}$$

where we used that $\det(U)$ cannot be zero since unitary operators are invertible. $\square$

Remark: For any unitary matrix

$$1 = \det(UU^{-1}) = \det(UU^\dagger) = \det(U)\det(U^\dagger) = \det(U)\overline{\det(U)} \qquad (6.64)$$

so that $\det(U)$ has unit modulus and can be written as $\det(U) = e^{i\epsilon}$ for some real ϵ. However, as explained in the paragraph preceding Eq. (5.12), we do not consider kets differing by a phase $e^{i\epsilon}$ as different states in quantum mechanics. Hence the state $|\Psi^-\rangle$ is actually invariant under transformations of the form $U \otimes U$

$$U \otimes U|\Psi^-\rangle = |\Psi^-\rangle \quad \text{up to a phase.} \qquad (6.65)$$

$\square$.

In the EPR thought experiment, two spin 1/2 particles are prepared in the state $|\Psi^-\rangle$ and then separated. The particles's spins are subsequently measured by two observers, traditionally called Alice and Bob, see Fig. 2. Both Alice and Bob only perform measurements

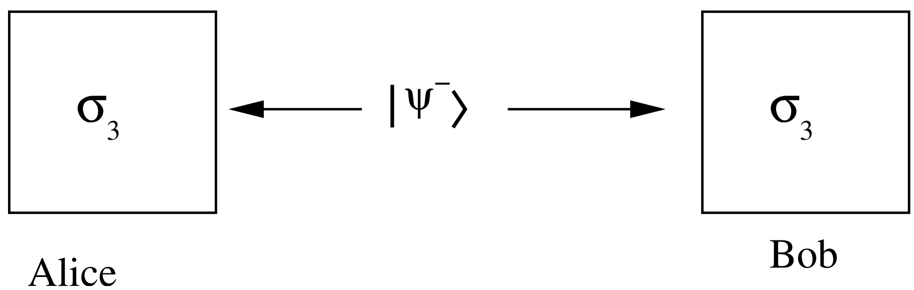

Figure 2: Schematic representation of Bohm's version of the EPR thought experiment

on "their" particle, i.e. Alice measures observables of the form $A \otimes I$ and Bob measures observables of the kind $I \otimes B$. We saw in the previous section that we can calculate all quantum mechanical predictions for measurements for such observables from the reduced density matrices. Thus starting with the density operator

$$\rho = |\Psi^-\rangle\langle\Psi^-| \qquad (6.66)$$

we compute the reduced density operator for the first qubit (Alice's) by tracing over the second (Bob's). Using our earlier calculation in Example 6.3.7 we have the following matrix representation with respect to the canonical basis:

$$\rho^A = \rho^B = \frac{1}{2}I \qquad (6.67)$$

(A for Alice and B for Bob). Now suppose that Alice measures the spin along the 3-axis of her particle, using, for example, a Stern-Gerlach apparatus. In the language of quantum mechanics, she performs a measurement of the operator $\frac{\hbar}{2}\sigma_3 \otimes I$. According to Sect. 4 the only possible outcomes of the measurement are $\pm\frac{\hbar}{2}$, corresponding to the spin pointing up or down. Suppose Alice finds the outcome "spin up" i.e. $\frac{\hbar}{2}$. The projection operator onto this eigenspace is

$$P = |0\rangle\langle 0| \otimes I \qquad (6.68)$$

and therefore the probability of the outcome “spin up” is

$$p_{\Psi^-}\left(\frac{\hbar}{2}\right) = \langle\, \Psi^-|P|\Psi^-\,\rangle = \frac{1}{\sqrt{2}}\langle\, \Psi^-|01\rangle = \frac{1}{2}$$

and the state after the measurement is

$$|\,\Psi\,\rangle = \sqrt{2}P|\,\Psi^-\,\rangle = |\,01\,\rangle. \tag{6.69}$$

This is a product state, and the reduced density operators for both Alice and Bob after Alice's measurement are pure states:

$$\tilde{\rho}^A = |\,0\,\rangle\langle\,0|, \qquad \tilde{\rho}^B = |\,1\,\rangle\langle\,1|. \tag{6.70}$$

Alice's state has changed

$$\rho^A = \frac{1}{2}I \mapsto \tilde{\rho}^B = |\,0\,\rangle\langle\,0| \tag{6.71}$$

as a result of her measurement in accordance with Postulate 2. However, Bob's state has also changed

$$\rho^B = \frac{1}{2}I \mapsto \tilde{\rho}^A = |\,1\,\rangle\langle\,1| \tag{6.72}$$

as a result of Alice's measurement. If we arrange for Alice and Bob to be well-separated at the time of Alice's measurement this result seems to imply that an event (Alice's measurement) can have an instantaneous effect in an arbitrarily far removed location (Bob). However, according to the special theory of relativity, there is a maximal speed with which information can be transmitted between two observers, namely the speed of light. Special relativity rules out “action at a distance” and therefore appears to be inconsistent with the quantum mechanical account of Bob's measurement and its effect on Alice's particle. Einstein, Podolsky and Rosen concluded that the quantum mechanical description of the situation in terms of the state $|\,\Psi^-\,\rangle$ is therefore **incomplete**. Recall that the knowledge of $|\,\Psi^-\,\rangle$ allows us to deduce that the spins of the two particles must be equal and opposite, but does not tell us anything about the direction of the spins. After Alice's measurement, the spin of Alice's particle is known to be in the 3-direction and, due to the correlation imposed by the state $|\,\Psi^-\,\rangle$, the spin of Bob's particle has to point in the opposite direction. The result of this argument is that at least one of three statements must be true:

1. The particles must be exchanging information instantaneously i.e. faster than light;
2. There are hidden variables, so the results of both Alice's and Bob's experiments are pre-ordained;
3. Quantum theory is not exactly true in these rather special experiments.

The first possibility may be described as the renunciation of the principle of locality, whereby signals cannot be passed from one particle to another faster than the speed of light. This suggestion was anathema to Einstein. EPR therefore concluded that if quantum theory was correct, i.e. if one ruled out possibility (3), then (2) must be true. In Einstein's terms, quantum theory was not complete but needed to be supplemented by hidden variables.

6.5 Bell's inequality

The particle physicist John Bell (1928-1990) derived a testable prediction from the assumption of local hidden variables, which has become known as Bell's inequaltity. There are several Bell inequalities, and we will consider one which is closely linked to our version of the EPR thought experiment. We will show that the quantum mechanical analysis of the EPR experiment shows that Bell's inequality should be violated, whereas the inequality should hold in any theory with local hidden variables. The different predictions can be and have been put to experimental tests. All such tests confirm the violation of Bell's inequalities, precisely as predicted by quantum mechanics.

The key idea is to allow Alice and Bob to conduct several measurements of spin along different axes, and to study the correlations between their findings. Consider spin operators

$$Q = \boldsymbol{q}\cdot\boldsymbol{\sigma} \otimes I, \quad R = \boldsymbol{r}\cdot\boldsymbol{\sigma} \otimes I, \quad S = I \otimes \boldsymbol{s}\cdot\boldsymbol{\sigma}, \quad \text{and} \quad T = I \otimes \boldsymbol{t}\cdot\boldsymbol{\sigma}. \tag{6.73}$$

associated to unit vectors $\boldsymbol{q}, \boldsymbol{r}, \boldsymbol{s}$ and $\boldsymbol{t}$ in $\mathbb{R}^3$. Two pairs of spin 1/2 particles are prepared in the state $|\Psi^-\rangle$ and separated. When Alice receives her particles, she picks two directions $\boldsymbol{q}$ and $\boldsymbol{r}$ at random and performs measurements of Q and R. When Bob receives his particles he picks two directions $\boldsymbol{s}$ and $\boldsymbol{t}$ at random and measures S and T, see Fig. 3. The experiment

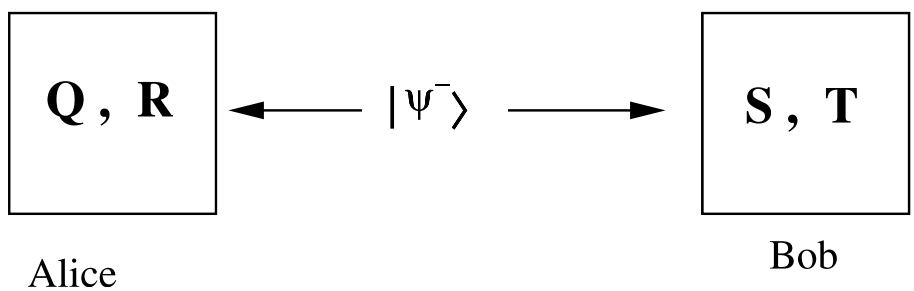

Figure 3: Schematic representation of the experimental set-up for the Bell inequality

is so arranged that Alice and Bob perform their measurements at the same time so that no measurement which Alice performs can disturb Bob' measurement and vice-versa. The possible outcomes of each of the measurements are 1 or -1.

The assumption of local hidden variables is tantamount to the following two mathematical assumptions

1. There is a probability space Λ and the observed outcomes by both Alice and Bob result by random sampling of the ("hidden") variable $\lambda \in \Lambda$.

2. The values observed by Alice or Bob are functions of the local detector settings and the hidden variable only. Thus the value of the spin observed by Alice with detector set to measure spin along the axis $\boldsymbol{q}$ is $A(\boldsymbol{q}, \lambda)$. Similarly, the value observed by Bob with detector set to measure spin along the axis $\boldsymbol{s}$ is $B(\boldsymbol{s}, \lambda)$.

It is implicit in assumption 1. above that the hidden variable space Λ has a probability measure $\rho(\lambda)d\lambda$ (satisfying $\int_\Lambda \rho(\lambda)d\lambda = 1$). The expectation of a random variable X on Λ

with respect to λ is written as

$$E(X) = \int_\Lambda X(\lambda)\rho(\lambda)d\lambda.$$

With the abbreviations

$$\begin{aligned} q(\lambda) = A(\boldsymbol{q}, \lambda), \quad r(\lambda) = A(\boldsymbol{r}, \lambda) \\ s(\lambda) = B(\boldsymbol{s}, \lambda), \quad t(\lambda) = B(\boldsymbol{t}, \lambda) \end{aligned} \tag{6.74}$$

we can compute the expectation values of Alice and Bob's measurements according to

$$E(q) = \int_\Lambda q(\lambda)\rho(\lambda)d\lambda, \tag{6.75}$$

etc. and we can compute expectation values of products of Alice's and Bob's measurement results according to

$$E(qs) = \int_\Lambda q(\lambda)s(\lambda)\rho(\lambda)d\lambda = \int_\Lambda A(\boldsymbol{q}, \lambda)B(\boldsymbol{s}, \lambda)\rho(\lambda)d\lambda, \tag{6.76}$$

with corresponding formulae of $E(rs), E(rt)$ and $E(qt)$.

Now consider a fixed $\lambda \in \Lambda$. Since $q(\lambda), r(\lambda) = \pm 1$ we must have either $(r(\lambda)+q(\lambda))s(\lambda) = 0$ (in which case $r(\lambda) - q(\lambda) = \pm 2$) or $(r(\lambda) - q(\lambda))t(\lambda) = 0$ (in which case $r(\lambda) + q(\lambda) = \pm 2$). In either case

$$qs + rs + rt - qt = (q + r)s + (r - q)t = \pm 2, \tag{6.77}$$

where all functions are evaluated at λ. Hence

$$E(qs + rs + rt - qt) \leq E(2) = 2. \tag{6.78}$$

Thus that we arrive a the Bell inequality

$$E(qs) + E(rs) + E(rt) - E(qt) \leq 2. \tag{6.79}$$

This particular version of the Bell inequality is called CHSH inequality, after its discoverers Clauser, Horne, Shimony and Holt.

We now show that Bell's inequality can be violated by quantum mechanical expectation values. Choose

$$\begin{aligned} Q = \sigma_3 \otimes I \quad , \qquad & R = \sigma_1 \otimes I, \\ S = -I \otimes \frac{1}{\sqrt{2}}(\sigma_1 + \sigma_3) \quad , \qquad & T = I \otimes \frac{1}{\sqrt{2}}(\sigma_3 - \sigma_1). \end{aligned} \tag{6.80}$$

Then

$$\begin{aligned} QS &= -\frac{1}{\sqrt{2}}\sigma_3 \otimes (\sigma_1 + \sigma_3) \\ RS &= -\frac{1}{\sqrt{2}}\sigma_1 \otimes (\sigma_1 + \sigma_3) \\ RT &= \frac{1}{\sqrt{2}}\sigma_1 \otimes (\sigma_3 - \sigma_1) \\ QT &= \frac{1}{\sqrt{2}}\sigma_3 \otimes (\sigma_3 - \sigma_1). \end{aligned} \tag{6.81}$$

It is straightforward to check that

$$\langle \Psi^- | \sigma_3 \otimes \sigma_3 | \Psi^- \rangle = \langle \Psi^- | \sigma_1 \otimes \sigma_1 | \Psi^- \rangle = -1 \tag{6.82}$$

and

$$\langle \Psi^- | \sigma_1 \otimes \sigma_3 | \Psi^- \rangle = \langle \Psi^- | \sigma_3 \otimes \sigma_1 | \Psi^- \rangle = 0. \tag{6.83}$$

It follows that

$$E_{\Psi^-}(QS) = \frac{1}{\sqrt{2}}, \quad E_{\Psi^-}(RS) = \frac{1}{\sqrt{2}}, \quad E_{\Psi^-}(RT) = \frac{1}{\sqrt{2}}, \quad E_{\Psi^-}(QT) = -\frac{1}{\sqrt{2}} \tag{6.84}$$

so that

$$E_{\Psi^-}(QS) + E_{\Psi^-}(RS) + E_{\Psi^-}(RT) - E_{\Psi^-}(QT) = 2\sqrt{2} > 2. \tag{6.85}$$

The calculation shows:

Theorem 6.5.1 (**Bell's theorem**) *No theory which uses local hidden variables can reproduce the predictions of quantum mechanics for all experiments.*

Furthermore, as mentioned in the introductory remarks, experiments show that Bell's inequality is indeed violated in precisely the way which quantum mechanics predicts, thus ruling out local hidden variable theories and corroborating quantum mechanics.

7 Quantum circuits and quantum algorithms

7.1 Classical versus quantum circuits

A classical circuit is made up of wires, which carry information, and gates, which perform simple computational tasks. It takes k input bits (i.e. a binary number with k digits) and produces l output bits. Mathematically, a classical circuit is therefore a function

$$f : \{0,1\}^k \to \{0,1\}^l. \tag{7.1}$$

Each gate is itself a map of this type. The wires indicate how the maps for each gate are to be composed to give the map for the entire circuit. The simplest non-trivial example of

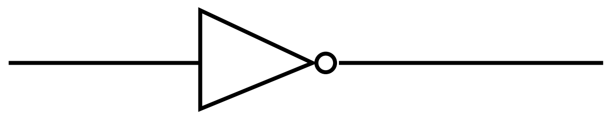

Figure 4: The NOT gate

a gate (and hence of a circuit) is the NOT gate, graphically represented as shown in Fig. 4 and corresponding to the map

$$n : \{0,1\} \to \{0,1\}, \quad n(x) = x \oplus 1, \tag{7.2}$$

where $\oplus$ denotes addition modulo 2, i.e.

$$0 \oplus 1 = 1, \quad 1 \oplus 1 = 0, \tag{7.3}$$

so that $n(0) = 1$ and $n(1) = 0$. If we interpret 1 as "true" and 0 as "false", the NOT gate negates, turning "true" into "false" and vice-versa.

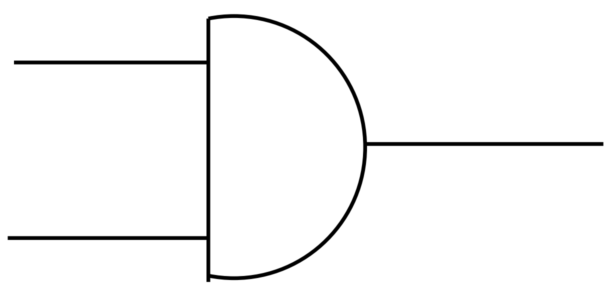

Figure 5: The AND gate

Other elementary examples are the AND gate, shown in Fig. 5. It corresponds to the map

$$a : \{0,1\}^2 \to \{0,1\}, \quad a(x,y) = xy. \tag{7.4}$$

Again interpreting 0 as "false" and 1 as "true", the AND gate takes "true" and "true" into "true", but gives the output "false" when either of the inputs is false. We can compose the AND and NOT gates to construct the circuit for NAND, shown in Fig. 6 We obtain

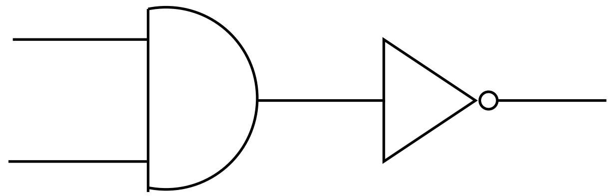

Figure 6: The NAND circuit

the mathematical function describing this circuit by composing the functions for AND and NOT, obtaining

$$na : \{0,1\}^2 \to \{0,1\}, \quad n(x,y) = xy \oplus 1. \tag{7.5}$$

Quantum circuits take qubits as input and produce qubits as output. We introduce the notation

$$\bigotimes^k \mathbb{C}^2 = \underbrace{\mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2}_{k\,\text{times}} \tag{7.6}$$

for the k-fold tensor product of the single qubit Hilbert space $\mathbb{C}^2$ with itself. Then quantum circuits are mathematically presented by maps

$$F : \bigotimes^k \mathbb{C}^2 \to \bigotimes^k \mathbb{C}^2. \tag{7.7}$$

Note that, unlike in classical circuits, the number of input qubits is equal to the number of output qubits. The basic reason for this lies in the nature of the two fundamental quantum mechanical processes, time evolution and measurement, on which quantum computing rests. As we have seen, for pure states both are mathematically represented by maps of the type (7.7), preserving the number of qubits. Time evolution of a k-qubit system is given by a unitary map

$$U : \bigotimes^k \mathbb{C}^2 \to \bigotimes^k \mathbb{C}^2, \tag{7.8}$$

and measurement is implemented by projection and rescaling

$$P : \bigotimes^k \mathbb{C}^2 \to \bigotimes^k \mathbb{C}^2, \quad |\psi\rangle \mapsto \frac{1}{\sqrt{\langle\psi|P|\psi\rangle}} P|\psi\rangle. \tag{7.9}$$

(Note that, because of the division by $\sqrt{\langle\psi|P|\psi\rangle}$ this map is not linear). The gates used in quantum computing makes use of these two types of operations and are correspondingly called **unitary gates** and **measurement gates**.

7.2 Unitary quantum gates

The simplest unitary quantum gates perform an operation on one qubit only i.e. they are unitary operators

$$U : \mathbb{C}^2 \to \mathbb{C}^2. \tag{7.10}$$

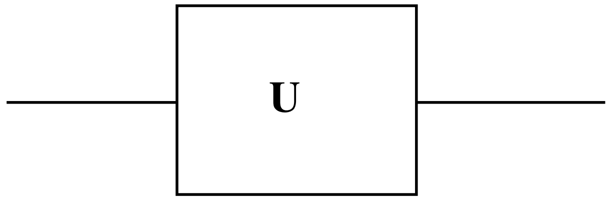

Figure 7: Circuit diagram for the unitary operator U

We studied all such operators in detail in Sect. 4, so we only introduce some additional nomenclature here. The circuit diagram for the unitary operator U is shown in Fig. 7. In particular there are gates correspoding to the three Pauli matrices. There is a special

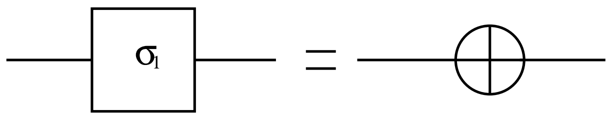

Figure 8: Two representations of the quantum gate σ_1

diagram for the Pauli gate σ_1 since it is a quantum analogue of the classical NOT gate, see Fig 8. We can write its action on the canonical basis states as

$$\sigma_1 : |x\rangle \mapsto |x \oplus 1\rangle \tag{7.11}$$

where $x \in \{0, 1\}$ and $\oplus$ is the addition modulo 2 as before. The gate represented by the operator with matrix representation

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{7.12}$$

relative to the canonical basis is called the **Hadamard gate**. Other simple gates which play a role in quantum computing are the phase gate S with matrix representation

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \tag{7.13}$$

and the so-called $\pi/8$ gate T with matrix representation

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} . \tag{7.14}$$

More interesting and useful gates involving two qubits are gates for **controlled operations**. The first qubit plays the role of the controller, the other that of the target. If the control qubit is in the state $|0\rangle$, the target qubit is left unchanged. If the control qubit is in the state $|1\rangle$, a prescribed unitary transformation U is performed on the target qubit. We depict the gate as shown in Fig. 9. An important example is $U = \sigma_1$. The resulting gate is called the CNOT gate, depicted in Fig. 10.It has the following action on the canonical basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$:

$$\begin{aligned} |00\rangle &\mapsto |00\rangle, \quad |01\rangle \mapsto |01\rangle \\ |10\rangle &\mapsto |11\rangle, \quad |11\rangle \mapsto |10\rangle \end{aligned} \tag{7.15}$$

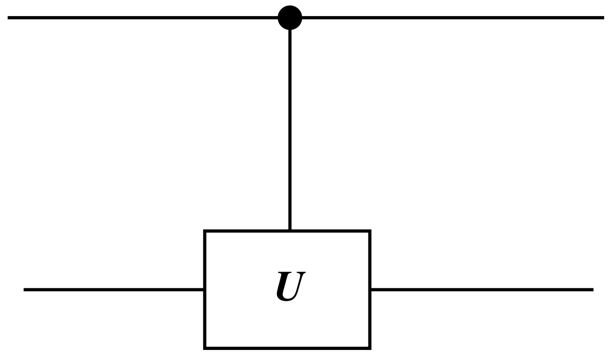

Figure 9: The gate for the controlled U operation

so that it is represented by the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{7.16}$$

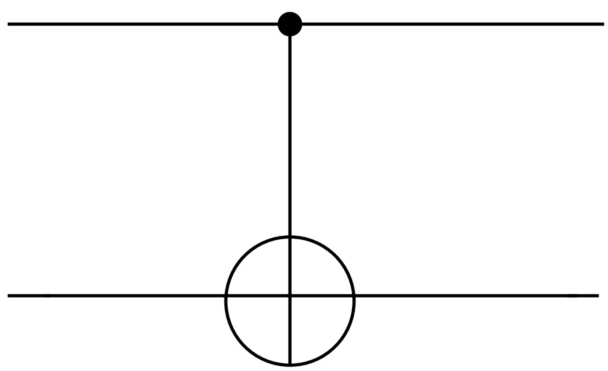

Figure 10: The gate for the controlled NOT operation

Composing the Hadamard gate on the control qubit with the CNOT gate we obtain our first interesting quantum circuit, shown in Fig. 11. We work out its effect on the canonical basis states by composing the operations

$$\begin{aligned} |00\rangle &\mapsto \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \mapsto \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ |01\rangle &\mapsto \frac{1}{\sqrt{2}}(|01\rangle + |11\rangle) \mapsto \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\ |10\rangle &\mapsto \frac{1}{\sqrt{2}}(|00\rangle - |10\rangle) \mapsto \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \\ |11\rangle &\mapsto \frac{1}{\sqrt{2}}(|01\rangle - |11\rangle) \mapsto \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \end{aligned} \tag{7.17}$$

so that the images of $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ are the entangled Bell states, conventionally denoted $|\Phi^+\rangle, |\Psi^+\rangle, |\Phi^-\rangle$ and, respectively, $|\Psi^-\rangle$.

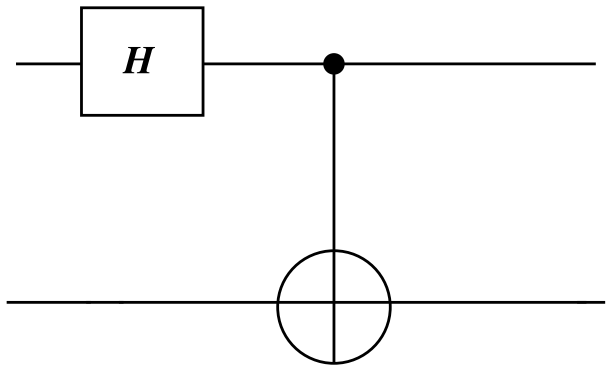

Figure 11: Quantum circuit to create Bell states

7.3 Measurement: the circuit for quantum teleportation

Measurement gates are depicted as shown inf Fig 12, with the outcome of the measurement (a real number) denoted m. We now combine the unitary gates of the previous section with

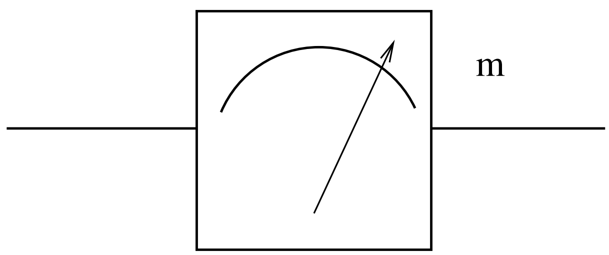

Figure 12: Graphical representation of a measurement gate

measurement gates to understand something non-trivial and surprising: quantum teleportation. The task is to send a quantum state in $\mathbb{C}^2$ to a recipient by only transmitting classical information *without knowing the state*. This can be achieved by using one of the Bell states constructed in the previous section, and three qubits. Of these, the first two belong to the sender (Alice) and the third to the recipient (Bob). Suppose that Alice and Bob generated the Bell state $|\Phi^+\rangle$ sometime in the past and then each took one qubit when they separated, Alice the first and Bob the second. The state to be teletransported is

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{7.18}$$

where α and β are unknown complex numbers satisfying $|\alpha|^2 + |\beta|^2 = 1$. The state

$$|\psi\rangle \otimes |\Phi^+\rangle = \frac{1}{\sqrt{2}}[\alpha(|000\rangle + |011\rangle) + \beta(|100\rangle + |111\rangle)] \tag{7.19}$$

is then input into the quantum circuit shown in Fig. 13. Alice sends her two qubits through a CNOT gate, resulting in the state

$$\frac{1}{\sqrt{2}}[\alpha(|000\rangle + |011\rangle) + \beta(|110\rangle + |101\rangle)]. \tag{7.20}$$

She then sends the first qubit through a Hadamard gate, leading to

$$\frac{1}{2}[\alpha(|000\rangle + |100\rangle + |011\rangle + |111\rangle) + \beta(|010\rangle - |110\rangle + |001\rangle - |101\rangle)]. \tag{7.21}$$

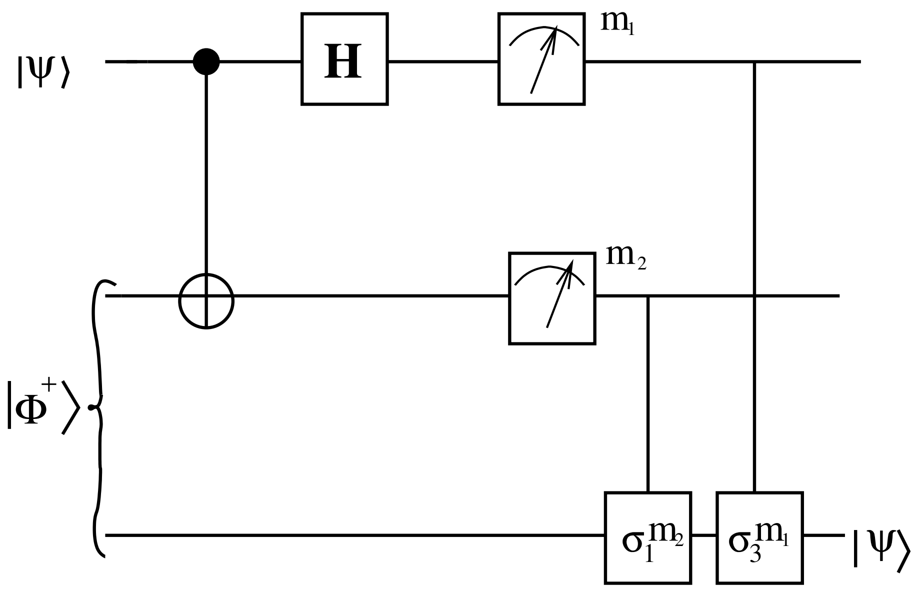

Figure 13: Quantum circuit for teleporting a qubit

which can be written as

$$\begin{aligned} &\frac{1}{2}[|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle) \\ + \quad &|10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)]. \end{aligned} \tag{7.22}$$

Now Alice performs measurements on her two qubits. She measures the observable $|1\rangle\langle 1|$ on the first and then on the second on her qubit, i.e she measures the commuting observables

$$|1\rangle\langle 1| \otimes I \otimes I \quad \text{and} \quad I \otimes |1\rangle\langle 1| \otimes I. \tag{7.23}$$

The possible outcomes of the measurements are $(m_1, m_2) = (0,0), (0,1), (1,0), (1,1)$ and correspondingly the state of her two qubits after the measurements are $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. If her qubits are in the state $|00\rangle$ she can tell Bob (by classical means - e.g. a phone call) that his state is now $|\psi\rangle$ i.e. she has successfully teleported her state. If her qubits are in the state $|01\rangle$, i.e. $m_2 = 1$, then Bob can recover the state $|\psi\rangle$ by passing his state through a σ_1-gate, which maps

$$(\alpha|1\rangle + \beta|0\rangle) \mapsto (\alpha|0\rangle + \beta|1\rangle) = |\psi\rangle \tag{7.24}$$

If Alice found the state $|10\rangle$, i.e. $m_1 = 1$, then Bob can recover the state $|\psi\rangle$ by passing his state through a σ_3-gate:

$$(\alpha|0\rangle - \beta|1\rangle) \mapsto (\alpha|0\rangle + \beta|1\rangle) = |\psi\rangle. \tag{7.25}$$

Finally, if Alice found the state $|11\rangle$ she can tell Bob to recover the state $|\psi\rangle$ by passing his state through a σ_1-gate and then a σ_3 gate:

$$(\alpha|1\rangle - \beta|0\rangle) \mapsto (\alpha|0\rangle - \beta|1\rangle) \mapsto (\alpha|0\rangle + \beta|1\rangle) = |\psi\rangle. \tag{7.26}$$

Thus, in general Bob can recover the state $|\psi\rangle$ by applying the transformations $\sigma_3^{m_1}\sigma_1^{m_2}$.

7.4 The Deutsch algorithm

In this final section we consider a quantum algorithm which "integrates" a function

$$f : \{0,1\} \to \{0,1\} \tag{7.27}$$

in a single evaluation of f. Classical computers would need to evaluate the function at both arguments 0 and 1 and then add the results. The algorithm we are about to discuss, called Deutsch algorithm after its inventor David Deutsch, therefore illustrates how quantum algorithms can outperform classical algorithms.

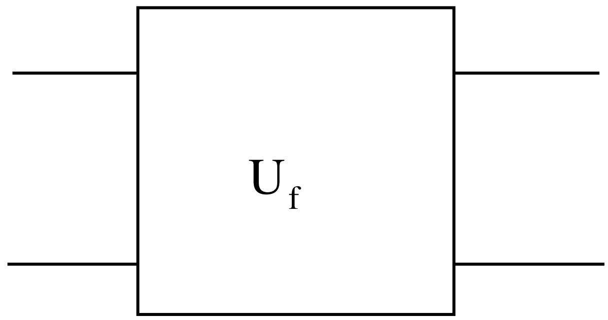

Figure 14: Quantum circuit implementing the unitrary transformation U_f

To construct the circuit, we first note that the linear operator $U_f : \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$ defined by its action on the canonical basis

$$U_f : |x, y\rangle \mapsto |x, y \oplus f(x)\rangle \tag{7.28}$$

is unitary. This can be checked explicitly by going through the four possibilities

$$f(0) = 0, f(1) = 0, \quad f(0) = 0, f(1) = 1, \quad f(0) = 1, f(1) = 0, \quad f(0) = 1, f(1) = 1. \tag{7.29}$$

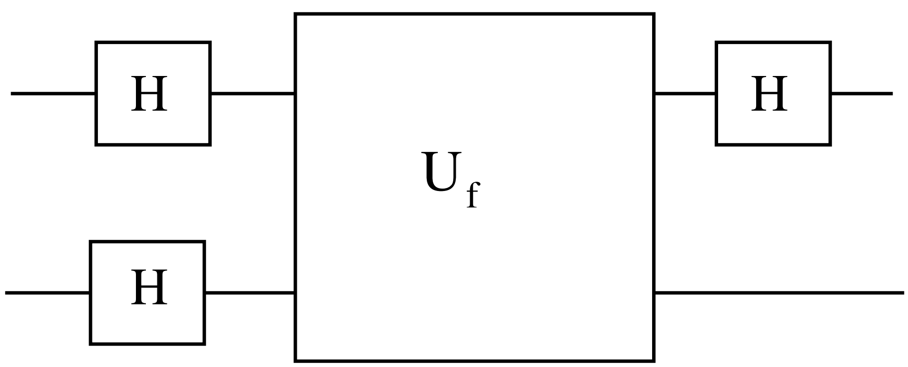

Figure 15: Quantum circuit implementing Deutsch's algorithm

Let us assume that we have a two-qubit gate that implements this transformation, diagrammatically shown in Fig. 14. The circuit diagram for the Deutsch algorithm is obtained by composing this gate with Hadamard gates, see Fig. 15. Suppose we input the state $|01\rangle$ into this circuit. The state after passing through the two Hadamard gates is

$$|\psi_1\rangle = \frac{1}{2}(|0\rangle + |1\rangle) \otimes (|0\rangle - |1\rangle). \tag{7.30}$$

Now note that applying U_f to the state $|x\rangle \otimes (|0\rangle - |1\rangle)$ gives $|x\rangle \otimes (|0\rangle - |1\rangle)$ if $f(x) = 0$ and $|x\rangle \otimes (|1\rangle - |0\rangle)$ if $f(x) = 1$. We can write this as

$$U_f(|x\rangle \otimes (|0\rangle - |1\rangle)) = (-1)^{f(x)}|x\rangle \otimes (|0\rangle - |1\rangle). \tag{7.31}$$

Hence, after passing through U_f the state $|\psi_1\rangle$ is

$$|\psi_2\rangle = \begin{cases} \frac{1}{2}(-1)^{f(0)}(|0\rangle + |1\rangle) \otimes (|0\rangle - |1\rangle) & \text{if} \quad f(0) = f(1) \\ \frac{1}{2}(-1)^{f(0)}(|0\rangle - |1\rangle) \otimes (|0\rangle - |1\rangle) & \text{if} \quad f(0) \neq f(1) \end{cases} \tag{7.32}$$

Applying the final Hadamard gate to the first qubit gives

$$|\psi_3\rangle = \begin{cases} (-1)^{f(0)}|0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) & \text{if} \quad f(0) = f(1) \\ (-1)^{f(0)}|1\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) & \text{if} \quad f(0) \neq f(1) \end{cases} \tag{7.33}$$

Now note that $f(0) \oplus f(1) = 0$ if $f(0) = f(1)$ and $f(0) \oplus f(1) = 1$ if $f(0) \neq f(1)$ to write the final state as

$$|\psi_3\rangle = (-1)^{f(0)}|f(0) + f(1)\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{7.34}$$

Hence by measuring the first qubit we obtain $|f(0) + f(1)|$, the promised "integral" of a function $f : \{0, 1\} \to \{0, 1\}$. Remarkably, this algorithm calculates a global property of the function f by a single evaluation of the function in (single) use of the gate U_f.