Introduction to quantum information theory

Suzie Brown

August 22, 2018

1 Quantum information

This section is based on the course by Noah Linden (University of Bristol) and discussion with Thomas Hebdige and David Jennings (Imperial College London). For a comprehensive introduction see for example Nielsen and Chuang (2002) or Wilde (2013).

The basis of quantum mechanics is simply linear algebra in a Hilbert space. The definitions in Sections 1.2 to 1.4 are exactly what you would find in a linear algebra course, except for the notation. The famously strange quantum behaviour arises from the "rules of quantum mechanics" (Section 1.5) pertaining to measurements. The majority of phenomena besides are described perfectly naturally by linear algebra, and are only counter-intuitive when one departs from the safety of mathematical abstraction and attempts to interpret quantum behaviour within the framework of the natural world, as perceived on a human scale. It is not surprising that such attempts should lead to confusion, since quantum mechanics is based on complex numbers and so diverges from human reality even at its very fundamentals.

1.1 Dirac notation

Dirac notation is a convenient way of denoting vectors such that it is easy to visually identify inner and outer products, and thus quickly recognise scalars, vectors and matrices:

- $|v\rangle$ denotes a column vector
- $\langle v|$ denotes a row vector
- $\langle u|v\rangle$ denotes an inner product (resulting in a scalar)
- $|u\rangle\langle v|$ denotes an outer product (resulting in a matrix)

Additionally, $\overline{\alpha}$ denotes the complex conjugate of a scalar α, and $U^\dagger$ denotes the adjoint (conjugate transpose) of an operator U.

1.2 Hilbert space

A *Hilbert space* is a vector space with an inner product $\langle\cdot|\cdot\rangle$ satisfying the following:

- $\langle u|\left(\alpha|v\rangle + \beta|w\rangle\right) = \alpha\langle u|v\rangle + \beta\langle u|w\rangle$
- $\langle u|v\rangle = \overline{\langle v|u\rangle}$
- $\langle v|v\rangle \geq 0$ with equality if and only if $|v\rangle$ is the zero vector.

1.3 Orthonormal bases

An *orthonormal basis* of a Hilbert space $\mathcal{H}$ is a set of vectors $\{v_1, \ldots, v_n\}$ in $\mathcal{H}$ such that:

- $\text{span}\{v_1, \ldots, v_n\} = \mathcal{H}$
- $\langle v_i|v_j\rangle = \delta_{ij}$

We now restrict to the Hilbert space $\mathbb{C}^2$, which is the state space of a single quantum bit (qubit), and is sufficient to describe the basics of quantum information. The *computational basis* $\{|0\rangle, |1\rangle\} = \{(1,0)^T, (0,1)^T\}$ is taken as the canonical basis for $\mathbb{C}^2$ and is henceforth used wherever not specified otherwise. Since it is an orthonormal basis, every vector $|v\rangle$ in $\mathbb{C}^2$ has a unique representation

$$|v\rangle = \alpha|0\rangle + \beta|1\rangle \equiv (\alpha, \beta)^T$$

for some $\alpha, \beta \in \mathbb{C}$. For reasons which will probably not become apparent in this treatment, we restrict ourselves to *normalised* vectors, requiring also $|\alpha|^2 + |\beta|^2 = 1$. We will also consider two vectors equivalent if they differ only by an overall phase, i.e. $|u\rangle \equiv |v\rangle$ if $|u\rangle = e^{i\theta}|v\rangle$ for some θ, since it is impossible to distinguish between two such vectors with any measurement.

To ensure coherency with the properties of the inner product, we have that

$$\langle v| = \overline{\alpha}\langle 0| + \overline{\beta}\langle 1|.$$

The inner product of $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ with $|u\rangle = \gamma|0\rangle + \delta|1\rangle$ is therefore computed as

$$\begin{aligned}\langle v|u\rangle &= (\overline{\alpha}\langle 0| + \overline{\beta}\langle 1|)(\gamma|0\rangle + \delta|1\rangle) \\ &= \overline{\alpha}\gamma\,\langle 0|0\rangle + \overline{\alpha}\delta\,\langle 0|1\rangle + \overline{\beta}\gamma\,\langle 1|0\rangle + \overline{\beta}\delta\,\langle 1|1\rangle \\ &= \overline{\alpha}\gamma + \overline{\beta}\delta.\end{aligned}$$

One alternative choice of orthonormal basis which is worth mentioning is given by $\{|+\rangle, |-\rangle\}$, consisting of the states

$$\begin{aligned}|+\rangle &:= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ |-\rangle &:= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).\end{aligned}$$

1.4 Linear operators

A linear operator is an operator with the property

$$A\left(\alpha|u\rangle + \beta|v\rangle\right) = \alpha A|u\rangle + \beta A|v\rangle.$$

It is therefore fully defined by its action on an orthonormal basis. For instance, the quantum NOT operator (usually denoted X) is defined by

$$\begin{aligned}X|0\rangle &= |1\rangle \\ X|1\rangle &= |0\rangle\end{aligned}$$

Equivalently, X can be expressed as a matrix with respect to the computational basis:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

In Dirac notation, operators are written in terms of outer products of basis states:

$$X = |0\rangle\langle 1| + |1\rangle\langle 0|$$

This is equivalent to the matrix form:

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

Then the action of X on a general state $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ can be calculated:

$$\begin{aligned}X|v\rangle &= (|0\rangle\langle 1| + |1\rangle\langle 0|)(\alpha|0\rangle + \beta|1\rangle) \\ &= \alpha|0\rangle\langle 1|0\rangle + \alpha|1\rangle\langle 0|0\rangle + \beta|0\rangle\langle 1|1\rangle + \beta|1\rangle\langle 0|1\rangle \\ &= \alpha|1\rangle + \beta|0\rangle.\end{aligned}$$

Definition 1. Let U be a linear operator.

- U is said to be *self-adjoint* (or *hermitian*) if $U^\dagger = U$.
- U is said to be *unitary* if $UU^\dagger = U^\dagger U = I$.

For example, the operator X is both self-adjoint and unitary. Crucially, unitary operators preserve normalisation, so they map states to states. It is also obvious that unitary transformations are always reversible (i.e. the inverse operator exists).

1.5 Rules of Quantum Mechanics

1. *States* of a quantum mechanical system correspond to normalised vectors in Hilbert space, up to an overall phase.
2. *Evolutions* of the system correspond to unitary operators.
3. *Measurements* on quantum states correspond to self-adjoint operators — see below.

1.5.1 Spectral theorem and measurement

The outcome of a measurement depends on the current state of the system and the type of measurement performed (i.e. which self-adjoint oeprator is applied). The spectral theorem states that every self-adjoint operator A can be represented by its spectral decomposition

$$A = \sum_i \lambda_i P_i \tag{1}$$

where $\{\lambda_1, \ldots, \lambda_k\}$ is the set of *distinct* eigenvalues of A, and P_i is the *projection* operator onto the eigenspace corresponding to eigenvalue λ_i.

When we measure a state $|x\rangle$ using operator A, the measurement outcome we observe is one of the eigenvalues of A. In particular, we observe λ_i with probability $\langle x|P_i|x\rangle$. Making a measurement causes the system to collapse onto the eigenspace corresponding to the observed eigenvalue; that is, the state after measurement is proportional to $P_i|x\rangle$.

Example 1. To give a concrete example, let us consider again the operator X. This operator has eigenvalues ± 1 corresponding to eigenvectors (or *eigenstates* in quantum mechanics) $|+\rangle$ and $|-\rangle$ respectively. Therefore X admits the diagonal representation

$$X = |+\rangle\langle +| - |-\rangle\langle -|$$

which is of the form (1). Now suppose we make a measurement on the state $|v\rangle = |0\rangle$, using X. The outcome of the measurement will be an eigenvalue of X: either $+1$ or -1. We observe the outcome $+1$ with probability

$$\langle v|P_{+1}|v\rangle = \langle 0|+\rangle \langle +|0\rangle = 1/2$$

in which case the state after measurement is

$$P_{+1}|v\rangle = |+\rangle\langle +|0\rangle \propto |+\rangle.$$

Similarly, with probability $1/2$ we observe the outcome -1 and the state after measurement is $|-\rangle$.

In this example, the two outcomes are equally likely because the state $|0\rangle$ is 'equidistant' from the two eigenstates of X. In general, outcomes that leave the state after measurement closer to the original state are more likely. This notion is formalised in Section 1.8.

A particularly important type of measurement is measurement in the computational basis. In this case the self-adjoint operator used is $0 \cdot |0\rangle\langle 0| + 1 \cdot |1\rangle\langle 1| = |1\rangle\langle 1|$. Measuring a state $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ in the computational basis returns 0 with probability $|\alpha|^2$ or 1 with probability $|\beta|^2$. This is useful in quantum computing algorithms because it deterministically converts the basis states $|0\rangle$ and $|1\rangle$ to their classical equivalents 0 and 1, ready for classical output.

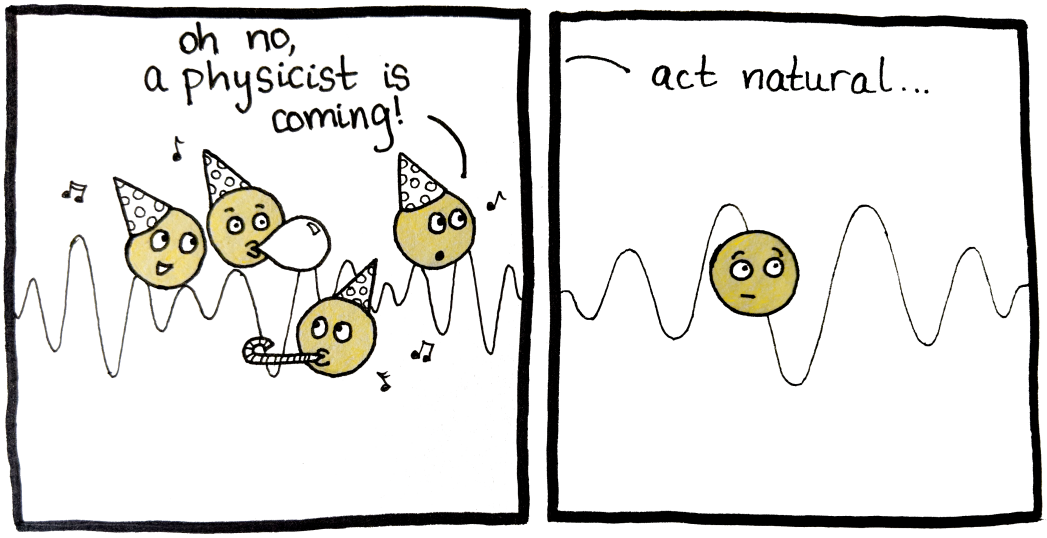


Figure 1: When a quantum system is in a superposition of states, making a measurement causes the system to collapse into the eigenstate corresponding to the observed eigenvalue.

1.6 Quantum randomness

It is worth remarking at this point on the difference between quantum and classical randomness. As statisticians we are used to dealing with randomness, but we accept that randomness is simply part of a model, and is not purported to exist in nature. We artificially introduce random variables into our models to account for a lack of information, either about the state of the system or about the (presumably deterministic) processes that govern certain phenomena.

On the contrary, our uncertainty about the outcome of a measurement on a quantum system is of a different kind. This uncertainty is not a symptom of our lack of knowledge: even when we know exactly the state of the system, as in Example 1, we are still unable to predict with certainty the outcome of a measurement on the system. The randomness here is intrinsic; the natural processes governing quantum measurement are truly stochastic.

1.7 Pure and mixed states

So far we have only encountered *pure states*, like $|0\rangle$ or $|+\rangle$. But we can also consider probabilistic mixtures of pure states, known as *mixed states*, like

$$|v\rangle = \sum_i p_i |v_i\rangle$$

where $\{p_i\}$ is a probability distribution and the $|v_i\rangle$ are pure states. You can think of this as a 'source' which emits a state $|v_i\rangle$ with probability p_i. A classical analogue is a bit which takes the value 1 with probability p or 0 with probability $1-p$ (a p-coin). In this case, we can say that the expected value of the bit is p.

There is no full ordering on the set of quantum states, so it is not possible to define the expected value of a mixed state. However, we can calculate the expected value of a measurement on a mixed state, using the state's *density operator*.

1.7.1 Density operators

The density operator of a mixed state $|v\rangle = \sum_i p_i |v_i\rangle$ is given by

$$\rho = \sum_i p_i |v_i\rangle\langle v_i|.$$

Suppose we measure this mixed state using the self-adjoint operator $A = \sum_k \lambda_k P_k$, where $\{\lambda_k\}$ are the distinct eigenvalues of A and P_k the projections onto the corresponding eigenspaces. Recall (from Section 1.5.1) that the outcome of the measurement will be an eigenvalue of A. Applying the law of total probability, the probability of

observing a particular outcome λ_j is

$$\mathbb{P}(\text{observe } \lambda_j) = \sum_i p_i \langle v_i | P_j | v_i \rangle = tr(\rho P_j)$$

where $tr(\cdot)$ denotes the trace of an operator. Note that $tr(\rho) = 1$ because the states are normalised and $\{p_i\}$ is a probability distribution. The state $\rho = I/2$ is called the *maximally mixed* state; it expresses complete ignorance, like a classical (1/2)-coin.

Different mixtures of states may have the same density operator, but in this case is is impossible to distinguish between them by any measurement.

Example 2. For example, a source $|v\rangle$ which emits $|0\rangle$ or $|1\rangle$ each with probability 1/2 has the same density operator $\rho = I/2$ as a source $|u\rangle$ which emits $|+\rangle$ or $|-\rangle$ each with probability 1/2. Suppose we measured both of these mixed states in the computational basis. The probability of observing the outcome 0 from $|u\rangle$ is

$$\langle u | P_0 | u \rangle = \frac{1}{2} \langle 0|0\rangle \langle 0|0\rangle + \frac{1}{2} \langle 1|0\rangle \langle 0|1\rangle = 1/2 \tag{2}$$

while the probability of observing 0 from $|v\rangle$ is

$$\langle v | P_0 | v \rangle = \frac{1}{2} \langle +|0\rangle \langle 0|+\rangle + \frac{1}{2} \langle -|0\rangle \langle 0|-\rangle = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = 1/2. \tag{3}$$

Therefore measurement in the computational basis can provide no information to distinguish between $|u\rangle$ and $|v\rangle$. And indeed nor can any other measurement — proved easily by exchanging the projection operator $|0\rangle\langle 0|$ for general P in the above — so the two sources are indistinguishable.

1.8 The Bloch sphere

Density operators have a neat geometrical representation as points in a unit sphere (Figure 2b). This is called the *Bloch sphere*, and can be thought of as the state space of a qubit.

An arbitrary pure state can be written in the form

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

Since the overall phase $e^{i\psi}$ can be ignored, the state is parametrised by two angles θ and ϕ. Taking these as spherical coordinates with unit radius, we can think of each pure state as a point on the unit sphere (or equivalently, a unit vector). Orthogonal states lie at opposite points on the Bloch sphere.

Mixed states can also be represented in the Bloch sphere. The Bloch vector for a mixed state is found by taking the weighted average of the Bloch vectors of its component pure states. Due to the triangle inequality, mixed states therefore lie strictly inside the unit sphere. This illustrates that several different mixtures of states can have the same density operator. Quantum states are distinguishable if and only if they correspond to different points in the Bloch sphere.

There is an analogy here with classical information. The state space of a classical bit is the unit interval: its value could be 0 or 1, or we could take a probabilistic mixture of the two. The 'pure' states then are 0 and 1, which lie on the 'surface' of the interval, while 'mixed' states lie strictly inside the interval (Figure 2a).

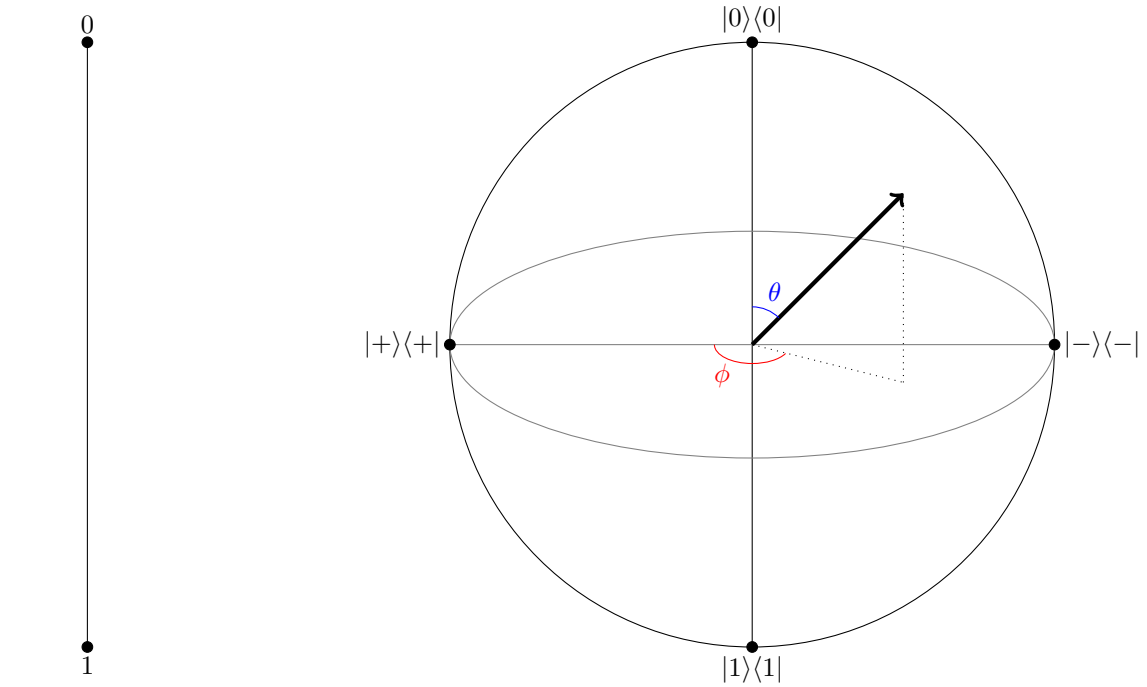
Finally, unitary operators correspond to rotations of the Bloch sphere. This illustrates that unitary transformations are reversible.

1.8.1 Geometric view of measurement

Suppose we are going to measure a pure state $|v\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle$ in the computational basis. There is a simple relationship between the probability of each outcome and the position of $|v\rangle$ in the Bloch sphere.

Figure 3 shows the semicircular slice of the Bloch sphere on which $|v\rangle$ lies. The value of ϕ determines which slice this is, and conditional on that the state only depends on θ — the following calculations are independent of ϕ and apply to an arbitrary pure state. The state $|v\rangle$ is projected onto the line segment between $|0\rangle$ and $|1\rangle$ as shown. Since the sphere has unit radius, the distance labelled x on Figure 3 is

$$x = 1 - \cos\theta.$$



(a) Classical bit

(b) Qubit

Figure 2: State space of classical bits versus qubits. The state space of a classical bit is the line segment $[0, 1]$. The 'pure' states 0 and 1 are at the endpoints, and probabilistic mixtures of the two lie inbetween. The state space of a qubit is the Bloch sphere. Pure states lie on the surface of the sphere, and mixed states lie in the interior. The centre of the sphere is the maximally mixed state $\rho = I/2$.

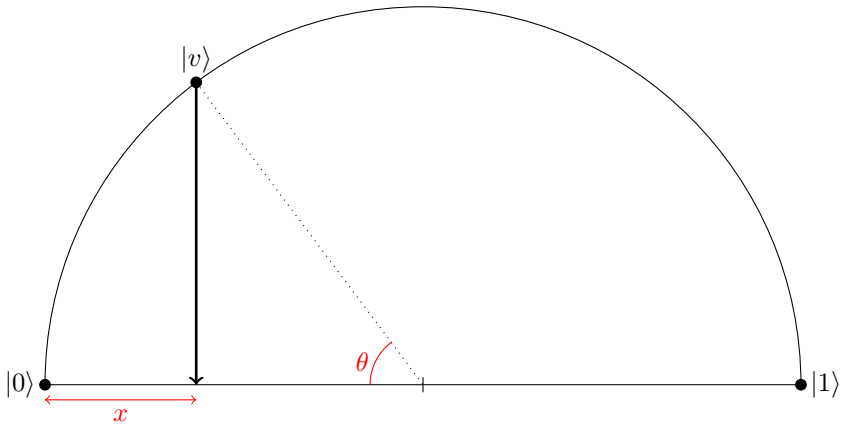


Figure 3: When measuring a state in the computational basis, the probability of observing a certain eigenvalue is related to the projected distance x from the corresponding eigenstate.

Using the expansion of $|v\rangle$ in the computational basis, we can calculate the probability of observing measurement outcome 0:

$$\mathbb{P}(\text{observe } 0) = \langle v|0\rangle \langle 0|v\rangle = \cos^2\left(\frac{\theta}{2}\right) = \frac{\cos\theta + 1}{2} = 1 - \frac{x}{2}.$$

This relationship between the projected distance and measurement probabilities also applies to mixed states. For instance, if we take any mixture of states all with the same value of θ (but different values of ϕ), the result will hold, because x and $\mathbb{P}$(observe 0) will be the same for every component state, and hence for the mixed state too. But the result holds for general mixed states as well. Any point in the sphere whose projection onto the line segment between $|0\rangle$ and $|1\rangle$ lies a distance x from $|0\rangle$ will measure 0 with probability $1 - x/2$.

Furthermore, this illustrates a more general concept, which applies not only to measurement in the computational basis but to any quantum measurement. Any valid measurement operator on one qubit (that is non-trivial in the sense of having more than one possible outcome) has two eigenstates that lie opposite each other on the Bloch sphere. Projecting a state $|v\rangle$ onto the line segment between the two eigenstates gives the probabilities of observing the associated eigenvalues.

1.9 Entanglement

Suppose we now have two qubits and would like to describe their joint state. The joint state space will be $\mathbb{C}^2 \otimes \mathbb{C}^2$, where $\otimes$ denotes the tensor product. This state space has canonical basis

$$\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}.$$

We also use the shorthand notations $|0\rangle \otimes |0\rangle = |0\rangle|0\rangle = |00\rangle$. The inner product on $\mathbb{C}^2 \otimes \mathbb{C}^2$ of $|u_1\rangle|u_2\rangle$ with $|v_1\rangle|v_2\rangle$ is defined as $\langle u_1|v_1\rangle \langle u_2|v_2\rangle$.

The joint state of two qubits $|u\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|v\rangle = \gamma|0\rangle + \delta|1\rangle$ can be expanded in the computational basis:

$$|u\rangle \otimes |v\rangle = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle.$$

States of this form are called *product states* as they can be written as the tensor product of two states in $\mathbb{C}^2$. However, the set of all product states is only a subset of $\mathbb{C}^2 \otimes \mathbb{C}^2$. In general, a state $a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$ may not be expressible as a product of two states in $\mathbb{C}^2$. In this case, it is called an *entangled* state.

For example, consider the state $(|00\rangle + |11\rangle)/\sqrt{2}$. This state is entangled: if we try to write it in the form $|u\rangle \otimes |v\rangle$ as above, we find that either α or δ must be zero, in which case the coefficient of either $|00\rangle$ or $|11\rangle$ must be zero, producing a contradiction.

The set of four entangled states

$$\left\{|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\right\} := \left\{\frac{|00\rangle + |11\rangle}{\sqrt{2}}, \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \frac{|01\rangle - |10\rangle}{\sqrt{2}}\right\}$$

is another important basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$, called the *Bell basis*.

References

Nielsen, M. A. and Chuang, I. (2002), *Quantum computation and quantum information*, AAPT.

Wilde, M. M. (2013), *Quantum information theory*, Cambridge University Press.