Proceedings of Symposia in Applied Mathematics
Volume **68**, 2010

A Rosetta Stone for Quantum Mechanics with an Introduction to Quantum Computation

Samuel J. Lomonaco, Jr.

Abstract. The purpose of these lecture notes is to provide readers, who have some mathematical background but little or no exposure to quantum mechanics and quantum computation, with enough material to begin reading the research literature in quantum computation, quantum cryptography, and quantum information theory. This paper is a written version of the first of eight one hour lectures given in the American Mathematical Society (AMS) Short Course on Quantum Computation held in conjunction with the Annual Meeting of the AMS in Washington, DC, USA in January 2000.

Part 1 of the paper is a preamble introducing the reader to the concept of the qubit

Part 2 gives an introduction to quantum mechanics covering such topics as Dirac notation, quantum measurement, Heisenberg uncertainty, Schrödinger's equation, density operators, partial trace, multipartite quantum systems, the Heisenberg versus the Schrödinger picture, quantum entanglement, EPR paradox, quantum entropy.

Part 3 gives a brief introduction to quantum computation, covering such topics as elementary quantum computing devices, wiring diagrams, the no-cloning theorem, quantum teleportation.

Many examples are given to illustrate underlying principles. A table of contents as well as an index are provided for readers who wish to "pick and choose." Since this paper is intended for a diverse audience, it is written in an informal style at varying levels of difficulty and sophistication, from the very elementary to the more advanced.

2000 *Mathematics Subject Classification.* Primary 81P68, 81-02, 81-01, 81P15, 22E70; Secondary 15A90, 94A15.

Key words and phrases. Quantum mechanics, quantum computation, quantum algorithms, entanglement, quantum information.

This work was partially supported by Army Research Office (ARO) Grant #P-38804-PH-QC, by the National Institute of Standards and Technology (NIST), by the Defense Advanced Research Projects Agency (DARPA) and Air Force Materiel Command USAF under agreement number F30602-01-0522, and by the L-O-O-P Fund. The author gratefully acknowledges the hospitality of the University of Cambridge Isaac Newton Institute for Mathematical Sciences, Cambridge, England, where some of this work was completed. I would also like to thank the other AMS Short Course lecturers, Howard Brandt, Dan Gottesman, Lou Kauffman, Alexei Kitaev, Peter Shor, Umesh Vazirani and the many Short Course participants for their support. (Copyright 2000.)

Contents

Part 1. Preamble

1. Introduction

These lecture notes were written for the American Mathematical Society (AMS) Short Course on Quantum Computation held 17-18 January 2000 in conjunction with the Annual Meeting of the AMS in Washington, DC in January 2000. They are intended for readers with some mathematical background but with little or no exposure to quantum mechanics. The purpose of these notes is to provide such readers with enough material in quantum mechanics and quantum computation to begin reading the vast literature on quantum computation, quantum cryptography, and quantum information theory.

The paper was written in an informal style. Whenever possible, each new topic was begun with the introduction of the underlying motivating intuitions, and then followed by an explanation of the accompanying mathematical finery. Hopefully, once having grasped the basic intuitions, the reader will find that the remaining material easily follows.

Since this paper is intended for a diverse audience, it was written at varying levels of difficulty and sophistication, from the very elementary to the more advanced. A large number of examples have been included. An index and table of contents are provided for those readers who prefer to "pick and choose." Hopefully, this paper will provide something of interest for everyone.

Because of space limitations, these notes are, of necessity, far from a complete overview of quantum mechanics. For example, only finite dimensional Hilbert spaces are considered, thereby avoiding the many pathologies that always arise when dealing with infinite dimensional objects. Many important experiments that are traditionally part of the standard fare in quantum mechanics texts (such as for example, the Stern-Gerlach experiment, Young's two slit experiment, the Aspect experiment) have not been mentioned in this paper. We leave it to the reader to decide if these notes have achieved their objective.

2. The classical world

2.1. Introducing the Shannon bit.

Since one of the objectives of this paper is to discuss quantum information, we begin with a brief discussion of classical information.

The Shannon bit is so well known in our age of information that it needs little, if any, introduction. As we all know, the Shannon bit is like a very decisive individual. It is either 0 or 1, but by no means both at the same time. The Shannon bit has become so much a part of our every day lives that we take many of its properties for granted. For example, we take for granted that Shannon bits can be copied.

2.2. Polarized light: Part I. The classical perspective.

Throughout this paper the quantum polarization states of light will be used to provide concrete illustrations of underlying quantum mechanical principles. So we also begin with a brief discussion of polarized light from the classical perspective.

Light waves in the vacuum are transverse electromagnetic (EM) waves with both electric and magnetic field vectors perpendicular to the direction of propagation and also to each other. (See figure 1.)

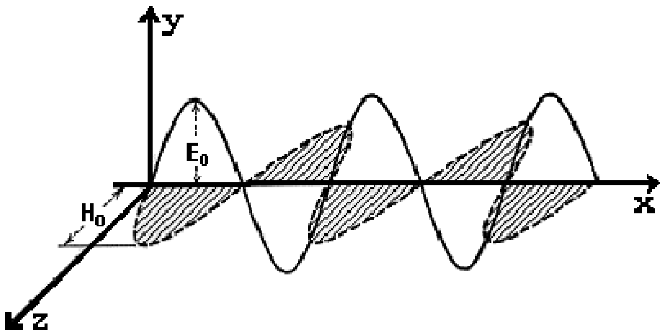


Figure 1. A linearly polarized electromagnetic wave.

If the electric field vector is always parallel to a fixed line, then the EM wave is said to be **linearly polarized**. If the electric field vector rotates about the direction of propagation forming a right-(left-)handed screw, it is said to be **right** (**left**) **elliptically polarized**. If the rotating electric field vector inscribes a circle, the EM wave is said to be right- or left-**circularly polarized**.

3. The quantum world

3.1. Introducing the qubit – But what is a qubit?

Many of us may not be as familiar with the quantum bit of information, called a **qubit**. Unlike its sibling rival, the Shannon bit, the qubit can be both 0 and 1 at the same time. Moreover, unlike the Shannon bit, the qubit can not be duplicated[1]. As we shall see, qubits are like very slippery, irascible individuals, exceedingly difficult to deal with.

One example of a qubit is a spin $\frac{1}{2}$ particle, which can be in a spin-up state $|1\rangle$ which we label as "1", in a spin-down state $|0\rangle$ which we label as "0", or in a **superposition** of these states, which we interpret as being both 0 and 1 at the same time. (The term "superposition" will be explained shortly.)

Another example of a qubit is the polarization state of a photon. A photon can be in a vertically polarized state $|\updownarrow\rangle$. We assign a label of "1" to this state. It can be in a horizontally polarized state $|\leftrightarrow\rangle$. We assign a label of "0" to this state. Or, it can be in a superposition of these states. In this case, we interpret its state as representing both 0 and 1 at the same time.

Anyone who has worn polarized sunglasses is familiar with the polarization states of light. Polarized sunglasses eliminate glare by letting through only vertically polarized light, while filtering out the horizontally polarized light. For that reason, they are often used to eliminate road glare, i.e., horizontally polarized light reflected from the road.

[1] This is a result of the no-cloning theorem of Dieks[**24**], Wootters and Zurek[**93**]. A proof of the no-cloning theorem is given in Section 11 of this paper.

3.2. Where do qubits live? – But what is a qubit?

But where do qubits live? They live in a Hilbert space $\mathcal{H}$. By a Hilbert space, we mean:

A **Hilbert space** $\mathcal{H}$ is a vector space over the complex numbers $\mathbb{C}$ with a complex valued inner product

$$(-,-) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$$

which is complete with respect to the norm

$$\|u\| = \sqrt{(u,u)}$$

induced by the inner product.

Remark 1. *By a complex valued inner product, we mean a map*

$$(-,-) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$$

from $\mathcal{H} \times \mathcal{H}$ into the complex numbers $\mathbb{C}$ such that:

1) $(u,u) = 0$ *if and only if* $u = 0$
2) $(u,v) = (v,u)^*$
3) $(u, v+w) = (u,v) + (u,w)$
4) $(u, \lambda v) = \lambda(u,v)$

where '' denotes the complex conjugate.*

Remark 2. *Please note that* $(\lambda u, v) = \lambda^*(u,v)$.

3.3. A qubit is ... [2]

A **qubit** is a quantum system $\mathcal{Q}$ whose state lies in a two dimensional Hilbert space $\mathcal{H}$.

Part 2. An Introduction to Quantum Mechanics

[2] Barenco et al in [**3**] define a qubit as a quantum system with a two dimensional Hilbert space, capable of existing in a superposition of Boolean states, and also capable of being entangled with the states of other qubits. Their more functional definition will take on more meaning as the reader progresses through this paper.

4. The beginnings of quantum mechanics

4.1. A Rosetta stone for Dirac notation: Part I. Bras, kets, and bra-(c)-kets.

The elements of a Hilbert space $\mathcal{H}$ will be called **ket vectors**, **state kets**, or simply **kets**. They will be denoted as:

$$|\, label\, \rangle$$

where '*label*' denotes some label, i.e., some chosen string of symbols that designate a state[3].

Let $\mathcal{H}^*$ denote the Hilbert space of all Hilbert space morphisms of $\mathcal{H}$ into the Hilbert space of all complex numbers $\mathbb{C}$, i.e.,

$$\mathcal{H}^* = Hom_{\mathbb{C}}\left(\mathcal{H}, \mathbb{C}\right).$$

The elements of $\mathcal{H}^*$ will be called **bra vectors**, **state bras**, or simply **bras**. They will be denoted as:

$$\langle\, label\, |$$

where once again '*label*' denotes some label.

Also please note that the complex number

$$\langle\, label_1\, | \,(|\, label_2\, \rangle)$$

will simply be denoted by

$$\langle\, label_1 \;|\; label_2\, \rangle$$

and will be called the **bra-(c)-ket** product of the bra $\langle\, label_1\, |$ and the ket $|\, label_2\, \rangle$.

There is a monomorphism (which is an isomorphism if the underlying Hilbert space is finite dimensional)

$$\mathcal{H} \xrightarrow{\dagger} \mathcal{H}^*$$

defined by

$$|\, label\, \rangle \longmapsto (\;|\, label\, \rangle\, , -)$$

[3] It should be mentioned that the bra and ket vectors of physics form a more general space than Hilbert space. P.A.M. Dirac [**28**, page 40] writes:

> The space of bra and ket vectors when the vectors are restriced to be of finite length and to have finite scalar products is called by mathematicians a *Hilbert space*. The bra and ket vectors that we now use form a more general space than Hilbert space.

Readers interested in pursuing this further should refer to the the theory of **Gel'fand triplets**, also known as **rigged Hilbert spaces**.

The bra $(\mid label \rangle, -)$ is denoted by $\langle label \mid$.

Hence,

$$\langle label_1 \mid label_2 \rangle = (\mid label_1 \rangle, \mid label_2 \rangle)$$

REMARK 3. *Please note that* $(\lambda \mid label \rangle)^\dagger = \lambda^* \langle label|$ *and* $\| \; |label\rangle \; \| = \sqrt{\langle \; label \mid label \; \rangle}$.

The **tensor product**[4] $\mathcal{H} \otimes \mathcal{K}$ of two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ is simply the "simplest" Hilbert space such that

1) $(h_1 + h_2) \otimes k = h_1 \otimes k + h_2 \otimes k$, for all h_1, $h_2 \in \mathcal{H}$ and for all $k \in \mathcal{K}$, and
2) $h \otimes (k_1 + k_2) = h \otimes k_1 + h \otimes k_2$ for all $h \in \mathcal{H}$ and for all k_1, $k_2 \in \mathcal{K}$.
3) $\lambda (h \otimes k) \equiv (\lambda h) \otimes k = h \otimes (\lambda k)$ for all $\lambda \in \mathbb{C}$, $h \in \mathcal{H}$, $k \in \mathcal{K}$.

It follows that, if $\{ \; e_1, e_2, \ldots, e_m \; \}$ and $\{ \; f_1, f_2, \ldots, f_n \; \}$ are respectively bases of the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, then $\{ \; e_i \otimes f_j \mid 1 \le i \le m, \, 1 \le j \le n \; \}$ is a basis of $\mathcal{H} \otimes \mathcal{K}$. Hence, the dimension of the Hilbert space $\mathcal{H} \otimes \mathcal{K}$ is the product of the dimensions of the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, i.e.,

$$Dim\,(\mathcal{H} \otimes \mathcal{K}) = Dim\,(\mathcal{H}) \cdot Dim\,(\mathcal{K}) \quad .$$

Finally, if $\mid label_1 \rangle$ and $\mid label_2 \rangle$ are kets respectively in Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, then their tensor product will be written in any one of the following three ways:

$$\mid label_1 \rangle \otimes \mid label_2 \rangle$$

$$\mid label_1 \rangle \mid label_2 \rangle$$

$$\mid label_1 \, , \, label_2 \rangle$$

4.2. Quantum mechanics: Part I. The state of a quantum system.

The states of a quantum system $\mathcal{Q}$ are represented by state kets in a Hilbert space $\mathcal{H}$. Two kets $|\alpha\rangle$ and $|\beta\rangle$ represent the same state of a quantum system $\mathcal{Q}$ if they differ by a non-zero multiplicative constant. In other words, $|\alpha\rangle$ and $|\beta\rangle$ represent the same quantum state $\mathcal{Q}$ if there exists a non-zero $\lambda \in \mathbb{C}$ such that

$$|\alpha\rangle = \lambda \, |\beta\rangle$$

[4] Readers well versed in homological algebra will recognize this informal definition as a slightly disguised version of the more rigorous universal definition of the tensor product. For more details, please refer to [**16**], or any other standard reference on homological algebra.

Hence, quantum states are simply elements of the manifold

$$\mathcal{H}/\tilde{\ } = \mathbb{C}P^{n-1}$$

where n denotes the dimension of $\mathcal{H}$, and $\mathbb{C}P^{n-1}$ denotes **complex projective** $(n-1)$**-space** .

> **Convention:** Since a quantum mechanical state is represented by a state ket up to a multiplicative constant, we will, unless stated otherwise, choose those kets $|\alpha\rangle$ which are of unit length, i.e., such that
>
> $$\langle \alpha \mid \alpha \rangle = 1 \Longleftrightarrow \| \, |\alpha\rangle \| = 1$$

4.3. Polarized light: Part II. The quantum mechanical perspective.

As an illustration of the above concepts, we consider the polarization states of a photon.

The polarization states of a photon are represented as state kets in a two dimensional Hilbert space $\mathcal{H}$. One orthonormal basis of $\mathcal{H}$ consists of the kets

$$|\circlearrowleft\rangle \text{ and } |\circlearrowright\rangle$$

which represent respectively the quantum mechanical states of left- and right-circularly polarized photons[5]. Another orthonormal basis consists of the kets

$$|\updownarrow\rangle \text{ and } |\leftrightarrow\rangle$$

representing respectively vertically and horizontally linearly polarized photons. And yet another orthonormal basis consists of the kets

$$|\nearrow\rangle \text{ and } |\searrow\rangle$$

for linearly polarized photons at the angles $\theta = \pi/4$ and $\theta = -\pi/4$ off the vertical, respectively.

These orthonormal bases are related as follows:

$$\left\{ \begin{array}{rcl} |\nearrow\rangle & = & \frac{1}{\sqrt{2}} \left(|\updownarrow\rangle + |\leftrightarrow\rangle \right) \\ |\searrow\rangle & = & \frac{1}{\sqrt{2}} \left(|\updownarrow\rangle - |\leftrightarrow\rangle \right) \end{array} \right. \qquad \left\{ \begin{array}{rcl} |\nearrow\rangle & = & \frac{1+i}{2} |\circlearrowright\rangle + \frac{1-i}{2} |\circlearrowleft\rangle \\ |\searrow\rangle & = & \frac{1-i}{2} |\circlearrowright\rangle + \frac{1+i}{2} |\circlearrowleft\rangle \end{array} \right.$$

$$\left\{ \begin{array}{rcl} |\updownarrow\rangle & = & \frac{1}{\sqrt{2}} \left(|\nearrow\rangle + |\searrow\rangle \right) \\ |\leftrightarrow\rangle & = & \frac{1}{\sqrt{2}} \left(|\nearrow\rangle - |\searrow\rangle \right) \end{array} \right. \qquad \left\{ \begin{array}{rcl} |\updownarrow\rangle & = & \frac{1}{\sqrt{2}} \left(|\circlearrowright\rangle + |\circlearrowleft\rangle \right) \\ |\leftrightarrow\rangle & = & \frac{i}{\sqrt{2}} \left(|\circlearrowright\rangle - |\circlearrowleft\rangle \right) \end{array} \right.$$

$$\left\{ \begin{array}{rcl} |\circlearrowright\rangle & = & \frac{1}{\sqrt{2}} \left(|\updownarrow\rangle - i \, |\leftrightarrow\rangle \right) \\ |\circlearrowleft\rangle & = & \frac{1}{\sqrt{2}} \left(|\updownarrow\rangle + i \, |\leftrightarrow\rangle \right) \end{array} \right. \qquad \left\{ \begin{array}{rcl} |\circlearrowright\rangle & = & \frac{1-i}{2} |\nearrow\rangle + \frac{1+i}{2} |\searrow\rangle \\ |\circlearrowleft\rangle & = & \frac{1+i}{2} |\nearrow\rangle + \frac{1-i}{2} |\searrow\rangle \end{array} \right.$$

[5] Please refer to [**84**, pages 9-10] for a discussion in regard to the circular polarization states of light.

The bracket products of the various polarization kets are given in the table below:

	$\lvert\updownarrow\rangle$	$\lvert\leftrightarrow\rangle$	$\lvert\nearrow\rangle$	$\lvert\searrow\rangle$	$\lvert\circlearrowright\rangle$	$\lvert\circlearrowleft\rangle$
$\langle\updownarrow\rvert$	1	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
$\langle\leftrightarrow\rvert$	0	1	$\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	$-\frac{i}{\sqrt{2}}$	$\frac{i}{\sqrt{2}}$
$\langle\nearrow\rvert$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	1	0	$\frac{1-i}{2}$	$\frac{1+i}{2}$
$\langle\searrow\rvert$	$\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	0	1	$\frac{1+i}{2}$	$\frac{1-i}{2}$
$\langle\circlearrowright\rvert$	$\frac{1}{\sqrt{2}}$	$\frac{i}{\sqrt{2}}$	$\frac{1+i}{2}$	$\frac{1-i}{2}$	1	0
$\langle\circlearrowleft\rvert$	$\frac{1}{\sqrt{2}}$	$-\frac{i}{\sqrt{2}}$	$\frac{1-i}{2}$	$\frac{1+i}{2}$	0	1

In terms of the basis $\{\lvert\updownarrow\rangle, \lvert\leftrightarrow\rangle\}$ and the dual basis $\{\langle\updownarrow\rvert, \langle\leftrightarrow\rvert\}$, these kets and bras can be written as matrices as indicated below:

$$\begin{cases} \langle\updownarrow\rvert = \begin{pmatrix} 1 & 0 \end{pmatrix}, & \lvert\updownarrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \langle\leftrightarrow\rvert = \begin{pmatrix} 0 & 1 \end{pmatrix}, & \lvert\leftrightarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{cases}$$

$$\begin{cases} \langle\nearrow\rvert = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \end{pmatrix}, & \lvert\nearrow\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \langle\searrow\rvert = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \end{pmatrix}, & \lvert\searrow\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{cases}$$

$$\begin{cases} \langle\circlearrowright\rvert = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \end{pmatrix}, & \lvert\circlearrowright\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} \\ \langle\circlearrowleft\rvert = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \end{pmatrix}, & \lvert\circlearrowleft\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} \end{cases}$$

In this basis, for example, the tensor product $\lvert\nearrow\circlearrowright\rangle$ is

$$\lvert\nearrow\circlearrowright\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \otimes \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -i \\ 1 \\ -i \end{pmatrix}$$

and the projection operator $\lvert\circlearrowleft\rangle\langle\circlearrowleft\rvert$ is:

$$\lvert\circlearrowleft\rangle\langle\circlearrowleft\rvert = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} \otimes \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$$

Remark 4. *Please note that, if bras and kets are represented as matrices, then the adjoint is nothing more than the conjugate transpose.*

4.4. A Rosetta stone for Dirac notation: Part II. Operators.

An **(linear) operator** or **transformation** $\mathcal{O}$ on a ket space $\mathcal{H}$ is a Hilbert space morphism of $\mathcal{H}$ into $\mathcal{H}$, i.e., is an element of

$$Hom_{\mathbb{C}}(\mathcal{H}, \mathcal{H})$$

The **adjoint** $\mathcal{O}^{\dagger}$ of an operator $\mathcal{O}$ is that operator such that

$$\left(\mathcal{O}^{\dagger} \,|\, label_1 \,\rangle \,, |\, label_2 \,\rangle\right) = \left(|\, label_1 \,\rangle \,, \mathcal{O} \,|\, label_2 \,\rangle\right)$$

for all kets $|\, label_1 \,\rangle$ and $|\, label_2 \,\rangle$.

In like manner, an (linear) operator or transformation on a bra space $\mathcal{H}^*$ is an element of

$$Hom_{\mathbb{C}}(\mathcal{H}^*, \mathcal{H}^*)$$

Moreover, each operator $\mathcal{O}$ on $\mathcal{H}$ can be identified with an operator, also denoted by $\mathcal{O}$, on $\mathcal{H}^*$ defined by

$$\langle\, label_1 \,|\longmapsto \langle\, label_1 \,|\, \mathcal{O}$$

where $\langle\, label_1 \,|\, \mathcal{O}$ is the bra defined by

$$\left(\langle\, label_1 \,|\, \mathcal{O}\right)\left(|\, label_2 \rangle\right) = \langle\, label_1 \,|\left(\mathcal{O} \,|\, label_2 \rangle\right)$$

(This is sometimes called Dirac's associativity law.) Hence, the expression

$$\langle\, label_1 \,|\, \mathcal{O} \,|\, label_2 \rangle$$

is unambiguous.

Remark 5. *Please note that*

$$\left(\mathcal{O} \,|\, label\rangle\right)^{\dagger} = \langle label|\, \mathcal{O}^{\dagger}$$

4.5. Quantum mechanics: Part II. Observables.

In quantum mechanics, an **observable** is simply a **Hermitian** (also called **self-adjoint**) operator on a Hilbert space $\mathcal{H}$, i.e., a linear operator $\mathcal{O}$ such that

$$\mathcal{O}^{\dagger} = \mathcal{O} \; .$$

A class of operators that play a fundamental role in quantum measurement are projection operators. A **projection operator** P on a Hilbert space $\mathcal{H}$ is a linear operator such that $P^2 = P$. Let V_P denote the sub-Hilbert space of all kets lying in the image of P. Then P is called the **projector for the subspace** V_P, and P is said to **project** $\mathcal{H}$ onto V_P. Projection operators are indeed very special observables, for:

Proposition 1. *Every projection operator is an observable.*

Projection operators can be naturally expressed in terms of Dirac notation. For let P be a projection operator for a subspace V_P of the Hilbert space $\mathcal{H}$. Then, if

$$\{ |label_0\rangle, |label_1\rangle, \dots, |label_{n-1}\rangle \}$$

is an orthonormal basis of V_P, the projection operator P can be written as

$$P = \sum_{j=0}^{n-1} |label_j\rangle \langle label_j| \,,$$

where $|label_j\rangle \langle label_j|$ denotes, for each j, the outer product of the unit length ket $|label_j\rangle$ with the unit length bra $\langle label_j|$. Moreover, $P_j = |label_j\rangle \langle label_j|$ is for each j the projection operator for the subspace spanned by the ket $|label_j\rangle$.

An **eigenvalue** a of an operator A is a complex number for which there exists a ket $|label\rangle$ in $\mathcal{H}$ such that

$$A\,|label\rangle = a\,|label\rangle \ .$$

The ket $|label\rangle$ is called an **eigenket** of A corresponding to the eigenvalue a. The **eigenspace** V_a of the eigenvalue a is the subspace of $\mathcal{H}$ of all eigenkets corresponding to the eigenvalue a, i.e., the space

$$V_a = \{ \ |label\rangle \mid A\,|label\rangle = a\,|label\rangle \ \}$$

We denote the **projection operator for the eigenspace** V_a by P_a.

THEOREM 1. *The eigenvalues a_j of an observable $\mathcal{O}$ are all real numbers. Moreover, eigenkets corresponding to distinct eigenvalues are orthogonal.*

The observables and unitary transformations of quantum mechanics are examples of a larger class of linear operators, called normal operators.

DEFINITION 1. *A linear operator A is said to be **normal** if it commutes with its adjoint, i.e.,*

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

The normal linear operators form a class of linear operators with many notable properties, one of the most important of which is the spectral decomposition theorem[6].

THEOREM 2 (Spectral Decomposition). *Let $a_0, a_1, \dots, a_{n-1}$ be the eigenvalues of a linear operator A on the Hilbert space $\mathcal{H}$. Then A is a normal operator if and only if it can be written in the form*

$$A = \sum_{j=0}^{n-1} a_j P_{a_j} \ ,$$

[6] A proof of the spectral decomposition theorem can be found in almost any suitably advanced book on linear algebra, such as for example [**45**].

where P_{a_j} denotes the projection operator for the eigenspace V_{a_j} corresponding to the eigenvalue a_j. Moreover, a linear operator A is normal if and only if the Hilbert space $\mathcal{H}$ is a direct sum of the eigenspaces of A, i.e., if and only if

$$\mathcal{H} = V_{a_0} \oplus V_{a_1} \oplus \cdots \oplus V_{a_{n-1}} \ ,$$

where the eigenspaces are mutually orthogonal.

Remark 6. *Hence, the projection operators P_{a_0}, P_{a_1}, ... , $P_{a_{n-1}}$ are* ***mutually orthogonal****, i.e., $P_{a_i}P_{a_j} = 0$ if $i \neq j$, and* ***complete****, i.e.,*

$$\sum_{j=0}^{n-1} P_{a_j} = 1 \ ,$$

where "1" denotes the identity linear transformation on the Hilbert space $\mathcal{H}$. Moreover, for all kets $|\psi\rangle$ in $\mathcal{H}$, we have the decomposition

$$|\psi\rangle = \sum_{j=0}^{n-1} P_{a_j} |\psi\rangle \ .$$

Thus, if A is a normal linear operator, then there is an orthonormal basis of the underlying Hilbert space $\mathcal{H}$ consisting entirely of eigenkets of A. In other words, the operator A can be **diagonalized**, i.e., written as a diagonal matrix as follows:

For each j, let

$$\left\{ e_{j1}, e_{j2}, \dots, e_{jm_j} \right\}$$

be an orthonormal basis of the eigenspace V_{a_j}. Then

$$\left\{ e_{01}, e_{02}, \dots, e_{0m_0}, \quad e_{11}, e_{12}, \dots, e_{1m_0}, \quad \dots, \quad e_{(n-1)1}, e_{(n-1)2}, \dots, e_{(n-1)m_{n-1}} \right\}$$

is an orthonormal basis of the Hilbert space $\mathcal{H}$. In terms of this basis, the matrix representation of the normal operator A is the $n \times n$ diagonal matrix with the following diagonal:

$$\left(\underbrace{a_0, \dots, a_0,}_{m_0} \quad \underbrace{a_1, \dots, a_1,}_{m_1} \quad \dots, \quad \underbrace{a_{n-1}, \dots, a_{n-1}}_{m_{n-1}} \right)$$

Definition 2. *Let A be a normal linear operator on a Hilbert space $\mathcal{H}$. Then an eigenvalue a is said to be* ***degenerate*** *if the corresponding eigenspace V_a is of dimension greater than 1. Otherwise, the eigenvalue a is said to be* ***nondegenerate****. A linear operator A is said to be* ***nondegenerate*** *it all its eigenvalues are nondegenerate. Otherwise, it is said to be* ***degenerate****.*

Notational Convention. *Let a_0, a_1, ... , a_{n-1} be the eigenvalues of a normal operator A. Let V_{a_0}, V_{a_1}, ... , $V_{a_{n-1}}$ denote the corresponding eigenspaces of respective dimensions n_{a_0}, n_{a_1}, ... , $n_{a_{n-1}}$. If an eigenvalue a_j is degenerate, then we will frequently denote an orthonormal basis of the eigenspace V_{a_j} simply by*

$$|a_j, 0\rangle, |a_j, 1\rangle, \dots, |a_j, n_{a_j} - 1\rangle \ .$$

On the other hand, if a_j is a nondegenerate eigenvalue of A, then we will denote a normalized (i.e., unit length) eigenket corresponding to a_j simply by

$$\left| a_j \right\rangle \ .$$

In terms of this notation, the spectral decomposition of A is

$$A = \sum_{j=0}^{n-1} a_j \sum_{k=0}^{n_{a_k}-1} \left| a_j, k \right\rangle \left\langle a_j, k \right| \ .$$

If A is nondegenerate, the spectral decomposition can be written more simply as

$$A = \sum_{j=0}^{n-1} a_j \left| a_j \right\rangle \left\langle a_j \right| \ .$$

Example 1. *Let $\mathcal{H}$ be the Hilbert space with orthonormal basis*

$$\{ \ \left|0\right\rangle, \left|1\right\rangle, \left|2\right\rangle, \left|3\right\rangle \ \} \ ,$$

and let $\mathcal{O}$ be the observable which, when written as a matrix in the above orthonormal basis, is given.

$$\mathcal{O} = \begin{pmatrix} 0 & 0 & -2i & 0 \\ 0 & 0 & 0 & 2i \\ 2i & 0 & 0 & 0 \\ 0 & -2i & 0 & 0 \end{pmatrix}$$

The observable $\mathcal{O}$ is degenerate. Its eigenvalues and corresponding eigenspace orthonormal bases (using the notation convention found on page 13) are given in the table below:

Eigenvalue	***Orthonormal Basis of Eigenspace $\mathbf{V}_{\mathbf{a}_j}$***
$a_0 = 2$	$\left\|+2, 0\right\rangle = \frac{1}{\sqrt{2}}\left(\left\|0\right\rangle + i\left\|2\right\rangle\right) = \frac{1}{\sqrt{2}}\left(1, 0, i, 0\right)^T$ $\left\|+2, 1\right\rangle = \frac{1}{\sqrt{2}}\left(\left\|1\right\rangle - i\left\|3\right\rangle\right) = \frac{1}{\sqrt{2}}\left(0, 1, 0, -i\right)^T$
$a_1 = -2$	$\left\|-2, 0\right\rangle = \frac{1}{\sqrt{2}}\left(\left\|0\right\rangle - i\left\|2\right\rangle\right) = \frac{1}{\sqrt{2}}\left(1, 0, -i, 0\right)^T$ $\left\|-2, 1\right\rangle = \frac{1}{\sqrt{2}}\left(\left\|1\right\rangle + i\left\|3\right\rangle\right) = \frac{1}{\sqrt{2}}\left(0, 1, 0, i\right)^T$

Its spectral decomposition is

$$\begin{aligned}\mathcal{O} &= a_0 P_0 + a_1 P_1\\ &= (+2)\Big(|{+2},0\rangle \langle {+2},0| \;+\; |{+2},1\rangle \langle {+2},1| \Big)\\ &\qquad\qquad + (-2)\Big(|{-2},0\rangle \langle {-2},0| \;+\; |{-2},1\rangle \langle {-2},1| \Big)\\ &= 2\begin{pmatrix} \frac{1}{2} & 0 & -\frac{i}{2} & 0\\ 0 & \frac{1}{2} & 0 & \frac{i}{2}\\ \frac{i}{2} & 0 & \frac{1}{2} & 0\\ 0 & -\frac{i}{2} & 0 & \frac{1}{2}\end{pmatrix} + (-2)\begin{pmatrix} \frac{1}{2} & 0 & \frac{i}{2} & 0\\ 0 & \frac{1}{2} & 0 & -\frac{i}{2}\\ -\frac{i}{2} & 0 & \frac{1}{2} & 0\\ 0 & \frac{i}{2} & 0 & \frac{1}{2}\end{pmatrix}\end{aligned}$$

In terms of the eigenket orthonormal basis

$$\Big\{ |{+2},0\rangle\,, \quad |{+2},1\rangle\,, \quad |{-2},0\rangle\,, \quad |{-2},1\rangle \Big\}\ ,$$

the matrix representation of the observable becomes the diagonal matrix

$$\mathcal{O} = \begin{pmatrix} a_0 & 0 & 0 & 0\\ 0 & a_0 & 0 & 0\\ 0 & 0 & a_1 & 0\\ 0 & 0 & 0 & a_1\end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & -2 & 0\\ 0 & 0 & 0 & -2\end{pmatrix}$$

Finally, as an illustration of ***Remark 6****, we have*

$$\begin{aligned}P_0 + P_1 &= \Big(|{+2},0\rangle \langle {+2},0| \;+\; |{+2},1\rangle \langle {+2},1| \Big)\\ &\qquad\qquad + \Big(|{-2},0\rangle \langle {-2},0| \;+\; |{-2},1\rangle \langle {-2},1| \Big)\\ &= \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} = 1\ ,\end{aligned}$$

where the last "1" denotes the identity operator.

Example 2. *The Pauli spin matrices*

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

are examples of observables that frequently appear in quantum mechanics and quantum computation. Their eigenvalues and eigenkets are given in the following table:

Pauli Matrices	***Eigenvalue/Eigenket***	
$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$+1$	$\frac{\vert 0\rangle + \vert 1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
	-1	$\frac{\vert 0\rangle - \vert 1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$
$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$	$+1$	$\frac{\vert 0\rangle + i\vert 1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$
	-1	$\frac{\vert 0\rangle - i\vert 1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$
$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$+1$	$\vert 0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
	-1	$\vert 1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Thus, the spectral decomposition of the Pauli operator σ_1 is:

$$\sigma_1 = (+1)\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)\left(\frac{\langle 0| + \langle 1|}{\sqrt{2}}\right) + (-1)\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)\left(\frac{\langle 0| - \langle 1|}{\sqrt{2}}\right) ,$$

which, in terms of the notation convention found on page 13, can be written as

$$\sigma_1 = (+1)\,|{+1}\rangle\,\langle{+1}| \quad + \quad (-1)\,|{-1}\rangle\,\langle{-1}| \ .$$

Please note that, as an illustration of **Remark *6*,** *we have*

$$|{+1}\rangle\,\langle{+1}| \quad + \quad |{-1}\rangle\,\langle{-1}| = 1 \ ,$$

where " 1 " denotes the identity operator.

4.6. Quantum mechanics: Part III. Quantum measurement — General principles.

According to quantum measurement theory, measurement is defined by the following rubric:

Quantum Measurement Rubric: *The **measurement of an observable** A of a quantum system $\mathcal{Q}$ in the state $|\psi\rangle$ produces the eigenvalue a_i as the measured result with probability*

$$Prob\,(\text{Value} \quad a_i \quad \text{is} \quad \text{observed}) = \left\| P_{a_i}\,|\psi\rangle \right\|^2 = \langle \psi \mid P_{a_i} \mid \psi \rangle \ ,$$

and forces the quantum system $\mathcal{Q}$ into the state

$$\frac{P_{a_i}\,|\psi\rangle}{\sqrt{\langle \psi \mid P_{a_i} \mid \psi \rangle}} \ ,$$

which lies in the corresponding eigenspace V_{a_i}.

Since quantum measurement is such a hotly debated topic among physicists, we (in self-defense) quote P.A.M. Dirac[**28**]:

> "A measurement always causes the (quantum mechanical) system to jump into an eigenstate of the dynamical variable that is being measured."

If $\mathcal{Q}$ is a quantum system in state $|\psi\rangle$, then the measurement of the system $\mathcal{Q}$ with respect to the observable A can be diagrammatically represented as follows:

$$\boxed{\begin{array}{ccccc} & \text{First} & & \text{Second} & \\ |\psi\rangle = \sum_i P_{a_i}|\psi\rangle & \text{Meas. of } A & \frac{P_{a_j}|\psi\rangle}{\sqrt{\langle\psi|P_{a_j}|\psi\rangle}} & \text{Meas. of } A & \frac{P_{a_j}|\psi\rangle}{\sqrt{\langle\psi|P_{a_j}|\psi\rangle}} \\ & \Longrightarrow & & \Longrightarrow & \\ & Prob = \left\langle \psi \mid P_{a_j} \mid \psi \right\rangle & & Prob = 1 & \end{array}}$$

Please note that the measured value is the eigenvalue a_j with probability $\left\langle \psi \mid P_{a_j} \mid \psi \right\rangle$. If the same measurement is repeated on the quantum system $\mathcal{Q}$ after the first measurement, then the result of the second measurement is no longer probabilistic. It produces the previously measured eigenvalue a_j, and the state of $\mathcal{Q}$ remains the same, i.e., $\frac{P_{a_j}|\psi\rangle}{\sqrt{\langle\psi|P_{a_j}|\psi\rangle}}$.

Example 1. (Cont.) *Let the Hilbert space $\mathcal{H}$ and the observable $\mathcal{O}$ be as in* ***Example 1****. Let $\mathcal{Q}$ be a quantum system in the state $|\psi\rangle \in \mathcal{H}$ given by*

$$|\psi\rangle = \frac{1}{3}\left(2|0\rangle + 2i|1\rangle - |3\rangle\right) = \frac{1}{3}\left(2, 2i, 0, -1\right)^T$$

Then measurement of the quantum system $\mathcal{Q}$ with respect to the observable $\mathcal{O}$ is summarized in the table below:

Index **j**	**Probability** $\left\langle \psi \mid \mathbf{P}_{\mathbf{a_j}} \mid \psi \right\rangle$	**Eigenvalue** $\mathbf{a_j}$	**Resulting State** $\frac{\mathbf{P}_{\mathbf{a_j}}|\psi\rangle}{\sqrt{\langle\psi|\mathbf{P}_{\mathbf{a_j}}|\psi\rangle}}$
0	$\frac{5}{18}$	$\mathbf{+2}$	$\frac{1}{\sqrt{10}}\left(\ 2|0\rangle + i|1\rangle + 2i|2\rangle + |3\rangle\ \right)$
1	$\frac{13}{18}$	$\mathbf{-2}$	$\frac{1}{\sqrt{26}}\left(\ 2|0\rangle + 3i|1\rangle - 2i|2\rangle - 3|3\rangle\ \right)$

Example 3. *Let $\mathcal{H}$ be a Hilbert space with orthonormal basis*

$$\{\ |0\rangle, |1\rangle, |2\rangle, \ldots, |n-1\rangle\ \}\ ,$$

and let $\mathcal{O}$ be the observable

$$\mathcal{O} = |0\rangle\langle 0|\ \ .$$

The operator is a degenerate observable. Its eigenvalues and corresponding eigenkets are given in the table below:

Eigenvalue	***Orthonormal Basis of Eigenspace*** $\mathbf{V}_{\mathbf{a}_j}$
$a_0 = 1$	$\|0\rangle$
$a_1 = 0$	$\begin{cases} \|0,0\rangle = \|1\rangle \\ \|0,1\rangle = \|2\rangle \\ \vdots \\ \|0,n-1\rangle = \|n\rangle \end{cases}$

Let $\mathcal{Q}$ *be a quantum system in the state*

$$|\psi\rangle = \sum_{j=0}^{n-1} \alpha_j \, |j\rangle \ \ ,$$

where

$$\sum_{j=0}^{n-1} |\alpha_j|^2 = 1 \ \ ,$$

If the quantum system $\mathcal{Q}$ *is measured with respect to the observable*

$$\mathcal{O} = |0\rangle \, \langle 0| \ \ ,$$

then the result of this measurement is given in the table below:

Index $\mathbf{j}$	***Probability*** $\left\langle \psi \mid \mathbf{P}_{\mathbf{a_j}} \mid \psi \right\rangle$	***Eigenvalue*** $\mathbf{a_j}$	***Resulting State*** $\frac{\mathbf{P}_{\mathbf{a_j}} \|\psi\rangle}{\sqrt{\left\langle \psi \| \mathbf{P}_{\mathbf{a_j}} \| \psi \right\rangle}}$
$\mathbf{0}$	$\left\|\alpha_{\mathbf{0}}^{\mathbf{2}}\right\|$	$+\mathbf{1}$	$\|\mathbf{0}\rangle$
$\mathbf{1}$	$\sum_{\mathbf{j=1}}^{\mathbf{n-1}} \|\alpha_{\mathbf{j}}\|^{\mathbf{2}}$	$\mathbf{0}$	$\frac{\sum_{\mathbf{j=1}}^{\mathbf{n-1}} \alpha_{\mathbf{j}} \|\mathbf{j}\rangle}{\sqrt{\sum_{\mathbf{j=1}}^{\mathbf{n-1}} \|\alpha_{\mathbf{j}}\|^{\mathbf{2}}}}$

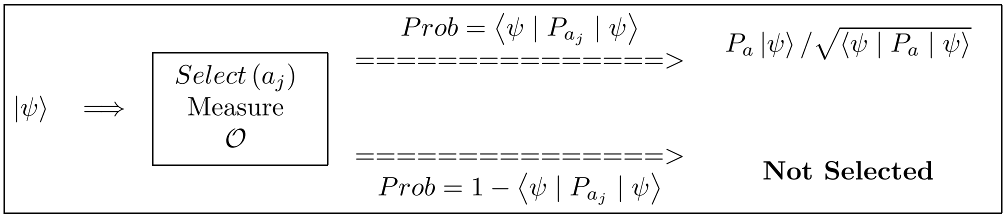
It should be mentioned that Julian Schwinger's **algebra of measurement operators**, a slightly different perspective on quantum measurement, can be useful in describing quantum measurement. It is very briefly described as follows:

There are quantum measuring devices that block all states resulting from a measurement of an observable $\mathcal{O}$ except for states corresponding to a selected eigenvalue a' of $\mathcal{O}$. A measurement, in this instance, is called a **selective measurement** (or **filtration**) of the observable $\mathcal{O}$ with respect to the selected eigenvalue a'.

Let $|\psi\rangle$ be the state of a quantum system $\mathcal{Q}$ before a selective measurement of the observable $\mathcal{O}$ with respect to the eigenvalue a'. After selective measurement, the quantum system $\mathcal{Q}$ is discarded if the resulting state

$$P_a \, |\psi\rangle \, / \sqrt{\langle \psi \mid P_a \mid \psi\rangle}$$

does not lie in the selected eigenspace $V_{a'}$, i.e., if $a \neq a'$, and accepted if it does, i.e., if $a = a'$. If a_0, a_1, ... , a_{n-1} are all the distinct eigenvalues of an observable $\mathcal{O}$, then a selective measurement with respect to a chosen eigenvalue a_j can be diagrammatically depicted as:


$|\psi\rangle \Longrightarrow$
$Select\,(a_j)$ Measure $\mathcal{O}$
$Prob = \langle \psi \mid P_{a_j} \mid \psi \rangle$
================>
$P_a \, |\psi\rangle \, / \sqrt{\langle \psi \mid P_a \mid \psi\rangle}$
================>
$Prob = 1 - \langle \psi \mid P_{a_j} \mid \psi \rangle$
Not Selected

Remark 7. *The above selective measurement is the same as a selective measurement of a_j with respect to the observable P_{a_j}. For this reason, the projection operator P_{a_j} is sometimes called a* ***selective measurement operator****.*

Readers interested in learning more about selective measurement should refer to the many references on Julian Schwinger's algebra of measurement operators, such as [**86**] and [**36**].

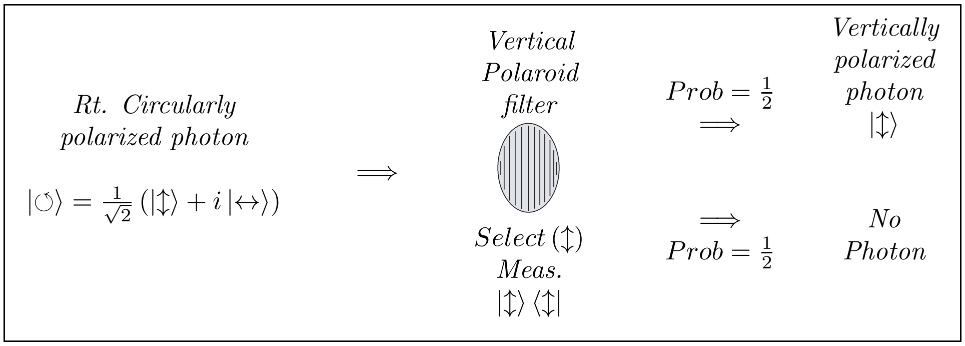
In the above discussion, we have focused mainly on von Neumann measure theory, and just briefly touched upon Schwinger's theory of measurement. The most general theory of quantum measurement is that of **positive operator valued measures** (**POVM**s) (also known as, **probability operator valued measures**). For more information on POVM theory, please refer to such books as, for example, [**23**], [**48**], [**77**], and [**81**].

4.7. Polarized light: Part III. Three examples of quantum measurement.

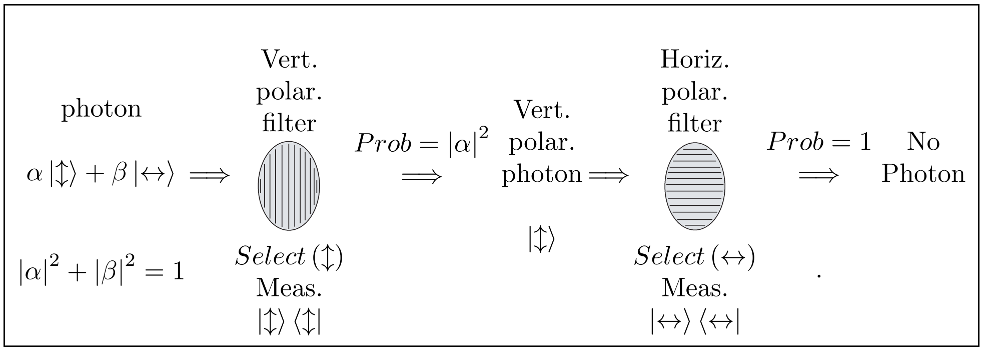
We can now apply the above general principles of quantum measurement to polarized light. Three examples are given below:[7]

[7] The last two examples can easily be verified experimentally with at most three pair of polarized sunglasses.

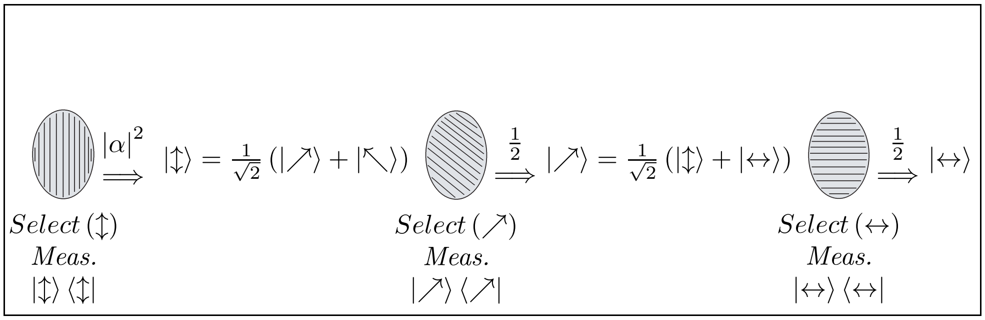
EXAMPLE 4.


Rt. Circularly polarized photon
$|\circlearrowleft\rangle = \frac{1}{\sqrt{2}}\left(|\updownarrow\rangle + i\,|\leftrightarrow\rangle\right)$
$\Longrightarrow$
Vertical Polaroid filter
Select $(\updownarrow)$ *Meas.*
$|\updownarrow\rangle\langle\updownarrow|$
$Prob = \frac{1}{2}$ $\Longrightarrow$
Vertically polarized photon $|\updownarrow\rangle$
$\Longrightarrow$ $Prob = \frac{1}{2}$
No Photon

EXAMPLE 5. *A vertically polarized filter followed by a horizontally polarized filter.*


photon
$\alpha\,|\updownarrow\rangle + \beta\,|\leftrightarrow\rangle \Longrightarrow$
$|\alpha|^2 + |\beta|^2 = 1$
Vert. polar. filter
Select $(\updownarrow)$ Meas.
$|\updownarrow\rangle\langle\updownarrow|$
$Prob = |\alpha|^2$ $\Longrightarrow$
Vert. polar. photon $\Longrightarrow$
$|\updownarrow\rangle$
Horiz. polar. filter
Select $(\leftrightarrow)$ Meas.
$|\leftrightarrow\rangle\langle\leftrightarrow|$
$Prob = 1$ $\Longrightarrow$
No Photon
.

EXAMPLE 6. *But if we insert a diagonally polarized filter (by* 45^o *off the vertical) between the two polarized filters in the above example, we have:*


$\xrightarrow{|\alpha|^2}$
$|\updownarrow\rangle = \frac{1}{\sqrt{2}}\left(|\nearrow\rangle + |\nwarrow\rangle\right)$
$\xrightarrow{\frac{1}{2}}$
$|\nearrow\rangle = \frac{1}{\sqrt{2}}\left(|\updownarrow\rangle + |\leftrightarrow\rangle\right)$
$\xrightarrow{\frac{1}{2}}$ $|\leftrightarrow\rangle$
Select $(\updownarrow)$ *Meas.* $|\updownarrow\rangle\langle\updownarrow|$
Select $(\nearrow)$ *Meas.* $|\nearrow\rangle\langle\nearrow|$
Select $(\leftrightarrow)$ *Meas.* $|\leftrightarrow\rangle\langle\leftrightarrow|$

where the input to the first filter is $\alpha\,|\updownarrow\rangle + \beta\,|\leftrightarrow\rangle$.

4.8. A Rosetta stone for Dirac notation: Part III. Expected values.

In this section we prove the following proposition which is an almost immediate consequence of the spectral decomposition theorem:

Proposition 2. *Let $|\psi\rangle$ be the state of a quantum system $\mathcal{Q}$, and let A be an observable of $\mathcal{Q}$. Then The* ***average value (expected value)*** *of a measurement with respect to an observable A of a quantum system in a state $|\psi\rangle$ is:*

$$\langle A\rangle = \langle\psi|\, A\, |\psi\rangle$$

Proof. Let a_0, a_1, ... , a_{n-1} be all the distinct eigenvalues of the observable A, and let P_{a_0}, P_{a_1}, ... , $P_{a_{n-1}}$ be the corresponding projection operators. Then by the spectral decomposition theorem, we have

$$A = \sum_{j=0}^{n-1} a_j P_{a_j} \ .$$

So,

$$\langle A\rangle = \langle\psi \mid A \mid \psi\rangle = \left\langle \psi \mid \sum\nolimits_{j=0}^{n-1} a_j P_{a_j} \mid \psi \right\rangle = \sum\nolimits_{j=0}^{n-1} a_j \left\langle \psi \mid P_{a_j} \mid \psi \right\rangle \ .$$

But

$$\left\langle \psi \mid P_{a_j} \mid \psi \right\rangle = Prob\left(\text{Observing } a_j \text{ on input } \ |\psi\rangle\right) \ .$$

Hence,

$$\langle A\rangle = \sum_{j=0}^{n-1} a_j Prob\left(\text{Observing } a_j \text{ on input } \ |\psi\rangle\right)$$

is indeed the average observed eigenvalue. □

4.9. Quantum Mechanics: Part IV. The Heisenberg uncertainty principle.

There is, surprisingly enough, a limitation of what we can observe in the quantum world.

From classical probability theory, we know that one yardstick of uncertainty is the **standard deviation**, which measures the average fluctuation about the mean. Thus, the **uncertainty** involved in the measurement of a quantum observable A is defined as the standard deviation of the observed eigenvalues. This standard deviation is given by the expression

$$Uncertainty(A) = \sqrt{\left\langle (\triangle A)^2 \right\rangle}$$

where

$$\triangle A = A - \langle A\rangle$$

Two observables A and B are said to be **compatible** if they commute, i.e., if

$$AB = BA.$$

Otherwise, they are said to be **incompatible**.

Let $[A, B]$, called the **commutator** of A and B, denote the expression

$$[A, B] = AB - BA$$

In this notation, two operators A and B are compatible if and only if $[A, B] = 0$.

The following principle is one expression of how quantum mechanics places limits on what can be observed:

Heisenberg's Uncertainty Principle[8]

$$\left\langle (\triangle A)^2 \right\rangle \left\langle (\triangle B)^2 \right\rangle \geq \frac{1}{4} \left| \langle [A, B] \rangle \right|^2$$

Thus, if A and B are incompatible, i.e., do not commute, then, by measuring A more precisely, we are forced to measure B less precisely, and vice versa. We can not simultaneously measure both A and B to unlimited precision. Measurement of A somehow has an impact on the measurement of B, and vice versa.

4.10. Quantum mechanics: Part V. Dynamics of closed quantum systems: Unitary transformations, the Hamiltonian, and Schrödinger's equation.

An operator U on a Hilbert space $\mathcal{H}$ is **unitary** if

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

Unitary operators are of central importance in quantum mechanics for many reasons. We list below only two:

- Closed quantum mechanical systems transform only via unitary transformations
- Unitary transformations preserve quantum probabilities

Let $|\psi(t)\rangle$ denote the state as a function of time t of a closed quantum mechanical system $\mathcal{Q}$. Then the dynamical behavior of the state of $\mathcal{Q}$ is determined by the **Schrödinger equation**

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle \ ,$$

where $\hbar$ denotes **Planck's constant** divided by 2π, and where H denotes an observable of $\mathcal{Q}$ called the **Hamiltonian**. The Hamiltonian is the quantum mechanical analog of the Hamiltonian of classical mechanics. In classical physics, the Hamiltonian is the total energy of the system.

REMARK 8. *The dynamical behavior of non-closed quantum systems is much more complex. (See* [**77**, Chapter 8]*.)*

[8] We have assumed units have been chosen such that $\hbar = 1$.

4.11. The mathematical perspective.

From the mathematical perspective, Schrödinger's equation is written as:

$$\frac{\partial}{\partial t}U(t) = -\frac{i}{\hbar}H(t)U(t) \ ,$$

where

$$\left|\psi(t)\right\rangle = U\left|\psi(0)\right\rangle \ ,$$

and where $-\frac{i}{\hbar}H(t)$ is a skew-Hermitian operator lying in the Lie algebra of the unitary group. The solution is given by a multiplicative integral, called the **path-ordered integral**,

$$U(t) = {}_{t}\mathcal{P}_{0}\, e^{-\frac{i}{\hbar}H(t)dt},$$

which is taken over the path $-\frac{i}{\hbar}H(t)$ in the Lie algebra of the unitary group. The path-ordered integral is given by:

$$\begin{aligned}{}_{t}\mathcal{P}_{0}\, e^{-\frac{i}{\hbar}H(t)dt} &= \lim_{n\to\infty}\prod_{k=n}^{0} e^{-\frac{i}{\hbar}H(k\frac{t}{n})\frac{t}{n}} \\ &= \lim_{n\to\infty}\left[e^{-\frac{i}{\hbar}H(n\cdot\frac{t}{n})\frac{t}{n}}\cdot e^{-\frac{i}{\hbar}H((n-1)\cdot\frac{t}{n})\frac{t}{n}}\cdot\ \ldots\ \cdot e^{-\frac{i}{\hbar}H(1\cdot\frac{t}{n})}\cdot e^{-\frac{i}{\hbar}H(0\cdot\frac{t}{n})\frac{t}{n}}\right]\end{aligned}$$

Remark 9. *The standard notation for the above path-ordered integral is*

$$\mathbf{P}\exp\left(-\frac{i}{\hbar}\int_{0}^{t}H(t)dt\right)$$

We prefer the elongated "P" notation for multiplicative integrals because it is similar to the elongated "S" notation for additive integrals.

If the Hamiltonian $H(t) = H$ is independent of time, then all matrices commute and the above path-ordered integral simplifies to

$${}_{t}\mathcal{P}_{0}\, e^{-\frac{i}{\hbar}Hdt} = e^{\int_0^t -\frac{i}{\hbar}Hdt} = e^{-\frac{i}{\hbar}Ht}$$

Thus, in this case, $U(t)$ is nothing more than a one parameter subgroup of the unitary group.

5. The Density Operator

5.1. Introducing the density operator.

John von Neumann suggested yet another way of representing the state of a quantum system.

Let $|\psi\rangle$ be a unit length ket (i.e., $\langle\ \psi \mid \psi\ \rangle = 1$) in the Hilbert space $\mathcal{H}$ representing the state of a quantum system[9]. The **density operator** ρ associated with the state ket $|\psi\rangle$ is defined as the outer product of the ket $|\psi\rangle$ (which can be thought of as a column vector) with the bra $\langle\psi|$ (which can be thought of as a row vector), i.e.,

$$\rho = |\psi\rangle \langle\psi|$$

The density operator formalism has a number of advantages over the ket state formalism. One advantage is that the density operator can also be used to represent hybrid quantum/classical states, i.e., states which are a classical statistical mixture of quantum states. Such hybrid states may also be thought of as quantum states for which we have incomplete information.

For example, consider a quantum system which is in the states (each of unit length)

$$|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle$$

with probabilities

$$p_1, p_2, \ldots, p_n$$

respectively, where

$$p_1 + p_2 + \ldots + p_n = 1$$

(Please note that the states $|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle$ need not be orthogonal.) Then the density operator representation of this state is defined as

$$\rho = p_1 |\psi_1\rangle \langle\psi_1| + p_2 |\psi_2\rangle \langle\psi_2| + \ldots + p_n |\psi_n\rangle \langle\psi_n|$$

If a density operator ρ can be written in the form

$$\rho = |\psi\rangle \langle\psi| \ ,$$

it is said to represent a **pure ensemble** . Otherwise, it is said to represent a **mixed ensemble** .

[9] Please recall that each of the kets in the set $\{\ \lambda |\psi\rangle \mid \lambda \in \mathbb{C},\ \lambda \neq 0\ \}$ represent the same state of a quantum system. Hence, we can always (and usually do) represent the state of a quantum system as a unit normal ket, i.e., as a ket such that $\langle\ \psi \mid \psi\ \rangle = 1$.

5.2. Properties of density operators.

It can be shown that all density operators are positive semi-definite Hermitian operators of trace 1, and vice versa. As a result, we have the following crisp mathematical definition:

Definition 3. *An linear operator on a Hilbert space $\mathcal{H}$ is a* ***density operator*** *if it is a positive semi-definite Hermitian operator of trace 1.*

It can be shown that a density operator represents a pure ensemble if and only if $\rho^2 = \rho$, or equivalently, if and only if $Trace(\rho^2) = 1$. For all ensembles, both pure and mixed, $Trace(\rho^2) \leq 1$.

From standard theorems in linear algebra, we know that, for every density operator ρ, there exists a unitary matrix U which **diagonalizes** ρ, i.e., such that $U\rho U^\dagger$ is a diagonal matrix. The diagonal entries in this matrix are, of course, the eigenvalues of ρ. These are non-negative real numbers which all sum to 1.

Finally, if we let $\mathcal{D}$ denote the set of all density operators for a Hilbert space $\mathcal{H}$, then $i\mathcal{D}$ is a convex subset of the Lie algebra of the unitary group associated with $\mathcal{H}$.

5.3. Quantum measurement in terms of density operators.

Let $a_0, a_1, \ldots, a_{n-1}$ denote all the distinct eigenvalues of an observable A, and let $P_{a_0}, P_{a_1}, \ldots, P_{a_{n-1}}$ be the corresponding projection operators. Finally, let $\mathcal{Q}$ be a quantum system with state given by the density operator ρ.

If the quantum system $\mathcal{Q}$ is measured with respect to the observable A, then with probability

$$p_i = Trace\left(P_{a_i}\rho\right)$$

the resulting measured eigenvalue is a_i, and the resulting state of $\mathcal{Q}$ is given by the density operator

$$\rho_i = \frac{P_{a_i}\rho P_{a_i}}{Trace\left(P_{a_i}\rho\right)} \ .$$

Moreover, for an observable A, the averaged observed eigenvalue expressed in terms of the density operator is:

$$\langle A \rangle = trace(\rho A)$$

Thus, we have extended the definition of $\langle A \rangle$ so that it applies to mixed as well as pure ensembles, i.e., generalized the following formula to mixed ensembles:

$$\langle A \rangle = \langle \psi \mid A \mid \psi \rangle = trace\left(\left|\psi\right\rangle\left\langle\psi\right| A\right) = trace(\rho A) \ .$$

EXERCISE 1. *Let the Hilbert space $\mathcal{H}$ and the observable $\mathcal{O}$ be as defined in* ***Example 1*** *on page 14. Let $\mathcal{Q}$ be a quantum system with state given by the density operator*

$$\rho = \begin{pmatrix} \frac{3}{8} & 0 & 0 & -\frac{3}{8} \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{3}{8} & 0 & 0 & \frac{3}{8} \end{pmatrix} .$$

Assume that the quantum system $\mathcal{Q}$ is measured with respect to the observable $\mathcal{O}$.

For $j = 0, 1$, find the probability $Prob\,(Observing\ a_j)$ of observing the eigenvalue a_j, and find the corresponding state ρ_j of the measured $\mathcal{Q}$.

5.4. Some examples of density operators.

For example, consider the following mixed ensemble of the polarization state of a photon:

EXAMPLE 7.

Ket	$\lvert\updownarrow\rangle$	$\lvert\nearrow\rangle$
Prob.	$\frac{3}{4}$	$\frac{1}{4}$

In terms of the basis $\lvert\leftrightarrow\rangle$, $\lvert\updownarrow\rangle$ of the two dimensional Hilbert space $\mathcal{H}$, the density operator ρ of the above mixed ensemble can be written as:

$$\begin{aligned} \rho &= \tfrac{3}{4}\lvert\updownarrow\rangle\langle\updownarrow\rvert + \tfrac{1}{4}\lvert\nearrow\rangle\langle\nearrow\rvert \\ &= \tfrac{3}{4}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} + \tfrac{1}{4}\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\ &= \tfrac{3}{4}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \tfrac{1}{8}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{7}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} \end{pmatrix} \end{aligned}$$

EXAMPLE 8. *The following two* ***preparations*** *produce mixed ensembles with the same density operator:*

Ket	$\lvert\updownarrow\rangle$	$\lvert\leftrightarrow\rangle$
Prob.	$\frac{1}{2}$	$\frac{1}{2}$

and

Ket	$\lvert\nearrow\rangle$	$\lvert\nwarrow\rangle$
Prob.	$\frac{1}{2}$	$\frac{1}{2}$

For, for the left preparation, we have

$$\begin{aligned} \rho &= \tfrac{1}{2}\lvert\updownarrow\rangle\langle\updownarrow\rvert + \tfrac{1}{2}\lvert\leftrightarrow\rangle\langle\leftrightarrow\rvert \\ &= \tfrac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} + \tfrac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix} \\ &= \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

And for the right preparation, we have

$$\begin{aligned}
\rho &= \tfrac{1}{2}\left|\nearrow\right\rangle\left\langle\nearrow\right| + \tfrac{1}{2}\left|\nwarrow\right\rangle\left\langle\nwarrow\right| \\
&= \tfrac{1}{2}\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \end{pmatrix} + \tfrac{1}{2}\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \end{pmatrix} \\
&= \tfrac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \tfrac{1}{4}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

There is no way of physically distinguishing the above two mixed ensembles which were prepared in two entirely different ways. For the density operator represents all that can be known about the state of the quantum system.

5.5. The partial trace of a linear operator.

In order to deal with a quantum system composed of many quantum subsystems, we need to define the partial trace.

Let

$$\mathcal{O} : \mathcal{H} \longrightarrow \mathcal{H} \in Hom_{\mathbb{C}}(\mathcal{H}, \mathcal{H})$$

be a linear operator on the Hilbert space $\mathcal{H}$.

Since Hilbert spaces are free algebraic objects, it follows from standard results in abstract algebra[10] that

$$Hom_{\mathbb{C}}(\mathcal{H}, \mathcal{H}) \cong \mathcal{H} \otimes \mathcal{H}^* \ ,$$

where we recall that

$$\mathcal{H}^* = Hom_{\mathbb{C}}(\mathcal{H}, \mathbb{C}) \ .$$

It follows that such an operator $\mathcal{O}$ can be written in the form

$$\mathcal{O} = \sum_{\alpha} a_\alpha \left|h_\alpha\right\rangle \otimes \left\langle k_\alpha\right| \ ,$$

where the kets $\left|h_\alpha\right\rangle$ lie in $\mathcal{H}$ and the bras $\left\langle k_\alpha\right|$ lie in $\mathcal{H}^\dagger$.

Thus, the standard **trace** of a linear operator

$$Trace : Hom_{\mathbb{C}}(\mathcal{H}, \mathcal{H}) \longrightarrow \mathbb{C}$$

is nothing more than a contraction, i.e.,

$$Trace(\mathcal{O}) = \sum_{\alpha} a_\alpha \left\langle\, k_\alpha \mid h_\alpha \,\right\rangle \ ,$$

i.e., a replacement of each outer product $\left|h_\alpha\right\rangle \otimes \left\langle k_\alpha\right|$ by the corresponding bracket

[10] See for example [**61**].

$\langle\, k_\alpha \mid h_\alpha \,\rangle$.

We can generalize the *Trace* as follows:

Let $\mathcal{H}$ now be the tensor product of Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$, … ,$\mathcal{H}_n$, i.e.,

$$\mathcal{H} = \bigotimes_{j=1}^{n} \mathcal{H}_j \ .$$

It follows once again from standard results in abstract algebra that

$$Hom_{\mathbb{C}}\left(\mathcal{H}, \mathcal{H}\right) \cong \bigotimes_{j=1}^{n} \left(\mathcal{H}_j \otimes \mathcal{H}_j^*\right) \ .$$

Hence, the operator $\mathcal{O}$ can be written in the form

$$\mathcal{O} = \sum_\alpha a_\alpha \bigotimes_{j=1}^{n} \left| h_{\alpha,j} \right\rangle \otimes \left\langle k_{\alpha,j} \right| \ ,$$

where, for each j, the kets $\left| h_{\alpha,j} \right\rangle$ lie in $\mathcal{H}_j$ and the bras $\left\langle k_{\alpha,j} \right|$ lie in $\mathcal{H}_j^*$ for all α.

Next we note that, for every subset $\mathcal{I}$ of the set of indices $\mathcal{J} = \{1, 2, \ldots, n\}$, we can define the **partial trace** over $\mathcal{I}$, written

$$Trace_{\mathcal{I}} : Hom_{\mathbb{C}}\left(\bigotimes_{j\in\mathcal{J}} \mathcal{H}_j, \bigotimes_{j\in\mathcal{J}} \mathcal{H}_j \right) \longrightarrow Hom_{\mathbb{C}}\left(\bigotimes_{j\in\mathcal{J}-\mathcal{I}} \mathcal{H}_j, \bigotimes_{j\in\mathcal{J}-\mathcal{I}} \mathcal{H}_j \right) \ ,$$

as the contraction on the indices $\mathcal{I}$, i.e.,

$$Trace_{\mathcal{I}}\left(\mathcal{O}\right) = \sum_\alpha a_\alpha \left(\prod_{j\in\mathcal{I}} \langle\, k_{\alpha,j} \mid h_{\alpha,j} \,\rangle \right) \bigotimes_{j\in\mathcal{J}-\mathcal{I}} \mid h_{\alpha,j} \,\rangle \langle\, k_{\alpha,j} \mid \ .$$

For example, let $\mathcal{H}_1$ and $\mathcal{H}_0$ be two dimensional Hilbert spaces with selected orthonormal bases $\{\left|0_1\right\rangle, \left|1_1\right\rangle\}$ and $\{\left|0_0\right\rangle, \left|1_0\right\rangle\}$, respectively. Thus, $\{\left|0_1 0_0\right\rangle, \left|0_1 1_0\right\rangle, \left|1_1 0_0\right\rangle, \left|1_1 1_0\right\rangle\}$ is an orthonormal basis of $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_0$.

Let $\rho \in Hom_{\mathbb{C}}\left(\mathcal{H}, \mathcal{H}\right)$ be the operator

$$\begin{aligned} \rho &= \left(\frac{\left|0_1 0_0\right\rangle - \left|1_1 1_0\right\rangle}{\sqrt{2}} \right) \otimes \left(\frac{\left\langle 0_1 0_0\right| - \left\langle 1_1 1_0\right|}{\sqrt{2}} \right) \\ &= \frac{1}{2}\left(\left|0_1 0_0\right\rangle \left\langle 0_1 0_0\right| - \left|0_1 0_0\right\rangle \left\langle 1_1 1_0\right| - \left|1_1 1_0\right\rangle \left\langle 0_1 0_0\right| + \left|1_1 1_0\right\rangle \left\langle 1_1 1_0\right| \right) \end{aligned}$$

which in terms of the basis $\{\left|0_1 0_0\right\rangle, \left|0_1 1_0\right\rangle, \left|1_1 0_0\right\rangle, \left|1_1 1_0\right\rangle\}$ can be written as the matrix

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

where the rows and columns are listed in the order $\left|0_1 0_0\right\rangle$, $\left|0_1 1_0\right\rangle$, $\left|1_1 0_0\right\rangle$, $\left|1_1 1_0\right\rangle$

The partial trace $Trace_0$ with respect to $\mathcal{I} = \{0\}$ of ρ is

$$\begin{aligned}\rho_1 &= Trace_0\left(\rho\right)\\ &= \frac{1}{2}Trace_0\left(\left|0_1 0_0\right\rangle\left\langle 0_1 0_0\right| - \left|0_1 0_0\right\rangle\left\langle 1_1 1_0\right| - \left|1_1 1_0\right\rangle\left\langle 0_1 0_0\right| + \left|1_1 1_0\right\rangle\left\langle 1_1 1_0\right|\right)\\ &= \frac{1}{2}\left(\left\langle 0_0 \,|\, 0_0\right\rangle\left|0_1\right\rangle\left\langle 0_1\right| - \left\langle 1_0 \,|\, 0_0\right\rangle\left|0_1\right\rangle\left\langle 1_1\right| - \left\langle 0_0 \,|\, 1_0\right\rangle\left|1_1\right\rangle\left\langle 0_1\right| + \left\langle 1_0 \,|\, 1_0\right\rangle\left|1_1\right\rangle\left\langle 1_1\right|\right)\\ &= \frac{1}{2}\left(\left|0_1\right\rangle\left\langle 0_1\right| - \left|0_1\right\rangle\left\langle 1_1\right|\right)\end{aligned}$$

which in terms of the basis $\{\left|0_1\right\rangle, \left|1_1\right\rangle\}$ becomes

$$\rho_1 = Trace_0(\rho) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} ,$$

where the rows and columns are listed in the order $\left|0_1\right\rangle$, $\left|1_1\right\rangle$.

5.6. Multipartite quantum systems.

One advantage density operators have over kets is that they provide us with a means for dealing with multipartite quantum systems.

Definition 4. *Let $\mathcal{Q}_1$, $\mathcal{Q}_2$, ... , $\mathcal{Q}_n$ be quantum systems with underlying Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$, ... , $\mathcal{H}_n$, respectively. The global quantum system $\mathcal{Q}$ consisting of the quantum systems $\mathcal{Q}_1$, $\mathcal{Q}_2$, ... , $\mathcal{Q}_n$ is called a **multipartite quantum system**. Each of the quantum systems $\mathcal{Q}_j$ $(j = 1, 2, \ldots, n)$ is called a **constituent "part"** of $\mathcal{Q}$. The underlying Hilbert space $\mathcal{H}$ of $\mathcal{Q}$ is the tensor product of the Hilbert spaces of the constituent "parts," i.e.,*

$$\mathcal{H} = \bigotimes_{j=1}^{n} \mathcal{H}_j \ .$$

If the density operator ρ is the state of a multipartite quantum system $\mathcal{Q}$, then the state of each constituent "part" $\mathcal{Q}_j$ is the density operator ρ_j given by the partial trace

$$\rho_j = Trace_{\mathcal{J}-\{j\}}\left(\rho\right) \ ,$$

where $\mathcal{J} = \{1, 2, \ \ldots \ , n\}$ is the set of indices.

Obviously, much more can be said about the states of multipartite systems and their constituent parts. However, we will forego that discussion until after we have had an opportunity introduce the concepts of quantum entanglement and von Neumann entropy.

5.7. Quantum dynamics in density operator formalism.

Under a unitary transformation U, a density operator ρ transforms according to the rubric:

$$\rho \longmapsto U \rho U^{\dagger}$$

Moreover, in terms of the density operator, Schrödinger's equation[11] becomes:

$$i\hbar \frac{\partial \rho}{\partial t} = [H, \rho] \ ,$$

where $[H, \rho]$ denotes the **commutator** of H and ρ, i.e.,

$$[H, \rho] = H\rho - \rho H$$

5.8. The mathematical perspective.

From the mathematical perspective, one works with $i\rho$ instead of ρ because $i\rho$ lies in the Lie algebra of the unitary group. Thus, the density operator transforms under a unitary transformation U according to the rubric:

$$i\rho \longmapsto Ad_U (i\rho) \ ,$$

where Ad_U denotes the **big adjoint representation**[12], i.e., the representation

$$i\rho \longmapsto U \left(i\rho \right) U^{-1}$$

From the mathematical perspective, Schrödinger's equation is in this case more informatively written as:

$$\frac{\partial (i\rho)}{\partial t} = ad_{-\frac{i}{\hbar}H} (i\rho) \ ,$$

where $ad_{-\frac{i}{\hbar}H}$ denotes the **little adjoint representation**[13] , i.e.,

$$ad_{-\frac{i}{\hbar}H} \left(i\rho \right) = \left[-\frac{i}{\hbar} H, i\rho \right] = \left(-\frac{i}{\hbar} H \right) \left(i\rho \right) - \left(i\rho \right) \left(-\frac{i}{\hbar} H \right) \ .$$

Thus, the solution to the above form of Schrödinger's equation is given by the path ordered integral:

$$\rho = \left({}_{t}\mathcal{S}_{0} \, e^{-\frac{1}{\hbar}\left(ad_{iH(t)} \right) dt} \right) \rho_0$$

where ρ_0 denotes the density operator at time $t = 0$.

[11] Schrödinger's equation determines the dynamics of closed quantum systems. However, non-closed quantum systems are also of importance in quantum computation and quantum information theory. See for example the Schumacher's work on superoperators [**85**], or [**77**, Chapter 8].

[12] For a more in depth discussion of the big adjoint representation, see[**67**].

[13] For a more in depth discussion of the little adjoint representation, see [**67**].

6. The Heisenberg model of quantum mechanics

Consider a computing device with inputs and outputs. Assume we have no knowledge of the internal workings of the device. We are allowed to probe the device with inputs and observe the corresponding outputs. But we are given no information as to how the device performs its calculation. We call such a device a **blackbox** computing device.

For such blackboxes, we say that two theoretical models for blackboxes are **equivalent** provided both predict the same input/output behavior. We may prefer one model over the other for various reasons, such as simplicity, aesthetics, or whatever meets our fancy. But the basic fact is that each of the two equivalent models is just as "correct" as the other.

In like manner, two theoretical models of the quantum world are said to be **equivalent** if they both predict the same results in regard to quantum measurements.

Up to this point, we have been describing the Schrödinger model of quantum mechanics, frequently called the **Schrödinger picture**. Heisenberg proposed yet another model, called the **Heisenberg picture**. Both models have been proven to be equivalent.

In the Heisenberg picture, state kets remain stationary with time, but observables move with time. While state kets, and hence density operators, remain fixed with respect to time, the observables A change dynamically as:

$$A \longmapsto U^{\dagger} A U$$

under a unitary transformation $U = U(t)$, where the unitary transformation is determined by the equation

$$i\hbar \frac{\partial U}{\partial t} = HU$$

It follows that the equation of motion of observables is according to the following equation

$$i\hbar \frac{\partial A}{\partial t} = [A, H]$$

One advantage the Heisenberg picture has over the Schrödinger picture is that the equations appearing in it are similar to those found in classical mechanics.

In summary, we have the following table which contrasts the two pictures:

	Schrödinger Picture	Heisenberg Picture
State ket	Moving $\|\psi_0\rangle \longmapsto \|\psi\rangle = U\|\psi_0\rangle$	Stationary $\|\psi_0\rangle$
Density Operator	Moving $\rho_0 \longmapsto \rho = U\rho_0 U^\dagger = Ad_U\left(\rho_0\right)$	Stationary ρ_0
Observable	Stationary A_0	Moving $A_0 \longmapsto A = U^\dagger A_0 U = Ad_{U^\dagger}\left(A_0\right)$
Observable Eigenvalues	Stationary a_j	Stationary a_j
Observable Frame	Stationary $A_0 = \sum_j a_j \|a_j\rangle_0 \langle a_j\|_0$	Moving $A_0 = \sum_j a_j \|a_j\rangle_0 \langle a_j\|_0$ $\longmapsto$ $A_t = \sum_j a_j \|a_j\rangle_t \langle a_j\|_t$ where $\|a_j\rangle_t = U^\dagger \|a_j\rangle_0$
Dynamical Equations	$i\hbar\frac{\partial U}{\partial t} = H^{(S)}U$ $i\hbar\frac{\partial}{\partial t}\|\psi\rangle = H^{(S)}\|\psi\rangle$	$i\hbar\frac{\partial U}{\partial t} = H^{(H)}U$ $i\hbar\frac{\partial A}{\partial t} = \left[A, H^{(H)}\right]$
Measurement	Measurement of observable A_0 produces eigenvalue a_j with probability $\left\|\left(\langle a_j\|_0\right)\|\psi\rangle\right\|^2 = \left\|\left(\langle a_j\|_0\right)\|\psi\rangle\right\|^2$	Measurement of observable A produces eigenvalue a_j with probability $\left\|\left(\langle a_j\|_t\right)\|\psi_0\rangle\right\|^2 = \left\|\left(\langle a_j\|_0\right)\|\psi\rangle\right\|^2$

where

$$H^{(H)} = U^\dagger H^{(S)} U$$

It follows that the Schrödinger Hamiltonian $H^{(S)}$ and the Heisenberg Hamiltonian are related as follows:

$$\frac{\partial H^{(S)}}{\partial t} = U\frac{\partial H^{(H)}}{\partial t}U^\dagger,$$

where terms containing $\frac{\partial U}{\partial t}$ and $\frac{\partial U^\dagger}{\partial t}$ have cancelled out as a result of the Schrödinger equation.

We should also mention that the Schrödinger and Heisenberg pictures can be transformed into one another via the mappings:

$S \longrightarrow H$	$H \longrightarrow S$
$\left\|\psi^{(S)}\right\rangle \longmapsto \left\|\psi^{(H)}\right\rangle = U^{\dagger}\left\|\psi^{(S)}\right\rangle$	$\left\|\psi^{(H)}\right\rangle \longmapsto \left\|\psi^{(S)}\right\rangle = U\left\|\psi^{(H)}\right\rangle$
$\rho^{(S)} \longmapsto \rho^{(H)} = U^{\dagger}\rho^{(S)}U$	$\rho^{(H)} \longmapsto \rho^{(S)} = U\rho^{(H)}U^{\dagger}$
$A^{(S)} \longmapsto A^{(H)} = U^{\dagger}A^{(S)}U$	$A^{(H)} \longmapsto A^{(S)} = UA^{(H)}U^{\dagger}$
$A^{(S)} \longmapsto A^{(H)} = U^{\dagger}A^{(S)}U$	$A^{(H)} \longmapsto A^{(S)} = UA^{(H)}U^{\dagger}$

Obviously, much more could be said on this topic.

For quantum computation from the perspective of the Heisenberg model, please refer to the work of Deutsch and Hayden[**26**], and also to Gottesman's "study of the ancient Hittites" :-) [**34**].

7. Quantum entanglement

7.1. The juxtaposition of two quantum systems.

Let $\mathcal{Q}_1$ and $\mathcal{Q}_2$ be two quantum systems that have been separately prepared respectively in states $|\psi_1\rangle$ and $|\psi_2\rangle$, and that then have been united without interacting. Because $\mathcal{Q}_1$ and $\mathcal{Q}_2$ have been separately prepared without interacting, their states $|\psi_1\rangle$ and $|\psi_2\rangle$ respectively lie in distinct Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. Moreover, because of the way in which $\mathcal{Q}_1$ and $\mathcal{Q}_2$ have been prepared, no physical prediction relating to one of these quantum systems depends in any way whatsoever on the other quantum system.

The global quantum system $\mathcal{Q}$ consisting of the two quantum systems $\mathcal{Q}_1$ and $\mathcal{Q}_2$ as prepared above is called a **juxtaposition** of the quantum systems $\mathcal{Q}_1$ and $\mathcal{Q}_2$. The state of the global quantum system $\mathcal{Q}$ is the tensor product of the states $|\psi_1\rangle$ and $|\psi_2\rangle$. In other words, the state of $\mathcal{Q}$ is:

$$|\psi_1\rangle \otimes |\psi_2\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$$

7.2. An example: An n-qubit register $\mathcal{Q}$ consisting of the juxtaposition of n qubits.

Let $\mathcal{H}$ be a two dimensional Hilbert space, and let $\{|0\rangle, |1\rangle\}$ denote an arbitrarily selected orthonormal basis[14]. Let $\mathcal{H}_{n-1}$, $\mathcal{H}_{n-2}$, ... ,$\mathcal{H}_0$ be distinct Hilbert spaces, each isomorphic to $\mathcal{H}$, with the obvious induced orthonormal bases

$$\{\ |0_{n-1}\rangle\ ,\ |1_{n-1}\rangle\ \}\ ,\ \{\ |0_{n-2}\rangle\ ,\ |1_{n-2}\rangle\ \}\ ,\ \ldots\ ,\ \{\ |0_0\rangle\ ,\ |1_0\rangle\ \}$$

respectively.

[14] We obviously have chosen to label the basis elements in a suggestive way.

Consider n qubits $\mathcal{Q}_{n-1}$, $\mathcal{Q}_{n-2}$, ... , $\mathcal{Q}_0$ separately prepared in the states

$$\frac{1}{\sqrt{2}}\left(|0_{n-1}\rangle + |1_{n-1}\rangle\right), \ \frac{1}{\sqrt{2}}\left(|0_{n-2}\rangle + |1_{n-2}\rangle\right), \ \ldots \ , \ \frac{1}{\sqrt{2}}\left(|0_0\rangle + |1_0\rangle\right),$$

respectively. Let $\mathcal{Q}$ denote the global system consisting of the separately prepared (without interacting) qubits $\mathcal{Q}_{n-1}$, $\mathcal{Q}_{n-2}$, ... , $\mathcal{Q}_0$. Then the state $|\psi\rangle$ of $\mathcal{Q}$ is the tensor product:

$$\begin{aligned}|\psi\rangle &= \frac{1}{\sqrt{2}}\left(|0_{n-1}\rangle + |1_{n-1}\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(|0_{n-2}\rangle + |1_{n-2}\rangle\right) \otimes \ldots \otimes \frac{1}{\sqrt{2}}\left(|0_0\rangle + |1_0\rangle\right) \\ &= \left(\frac{1}{\sqrt{2}}\right)^n \left(|0_{n-1}0_{n-2}\ldots 0_1 0_0\rangle + |0_{n-1}0_{n-2}\ldots 0_1 1_0\rangle + \ \ldots \ + |1_{n-1}1_{n-2}\ldots 1_1 1_0\rangle\right)\end{aligned}$$

which lies in the Hilbert space

$$\mathcal{H} = \mathcal{H}_{n-1} \otimes \mathcal{H}_{n-2} \otimes \ \ldots \ \otimes \mathcal{H}_0 .$$

Notational Convention: We will usually omit subscripts whenever they can easily be inferred from context.

Thus, the global system $\mathcal{Q}$ consisting of the n qubits $\mathcal{Q}_{n-1}$, $\mathcal{Q}_{n-2}$, ... , $\mathcal{Q}_0$ is in the state

$$|\psi\rangle = \left(\frac{1}{\sqrt{2}}\right)^n \left(|00\ldots 00\rangle + |00\ldots 01\rangle + \ \ldots \ + |11\ldots 11\rangle\right) \in \bigotimes_0^{n-1} \mathcal{H}$$

The reader should note that the n-qubit register $\mathcal{Q}$ is a superposition of kets with labels consisting of all the binary n-tuples. If each binary n-tuple $b_{n-1}b_{n-2}\ldots b_0$ is identified with the integer

$$b_{n-1}2^{n-1} + b_{n-2}2^{n-2} + \ldots + b_0 2^0 \ ,$$

i.e., if we interpret each binary n-tuple as the radix 2 representation of an integer, then we can rewrite the state as

$$|\psi\rangle = \left(\frac{1}{\sqrt{2}}\right)^n \left(|0\rangle + |1\rangle + |2\rangle + \ \ldots \ + |2^n - 1\rangle\right).$$

In other words, this n-qubit register contains all the integers from 0 to $2^n - 1$ in superposition. But most importantly, it contains all the integers 0 to $2^n - 1$ *simultaneously*!

This is an example of the massive parallelism that is possible within quantum computation. However, there is a downside. If we observe (measure) the register, then all the massive parallelism disappears. On measurement, the quantum world selects for us one and only one of the 2^n integers. The probability of observing any particular one of the integers is $\left|\left(1/\sqrt{2}\right)^n\right|^2 = (\frac{1}{2})^n$. The selection of which integer is observed is unfortunately not made by us, but by the quantum world.

Thus, harnessing the massive parallelism of quantum mechanics is no easy task! As we will see, a more subtle approach is required.

7.3. An example of the dynamic behavior of a 2-qubit register.

We now consider the previous n-qubit register for $n = 2$. In terms of the bases described in the previous section, we have:

$$\left\{\begin{array}{ccccccc} |0\rangle & = & |00\rangle & = & \begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} & = & \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \\ |1\rangle & = & |01\rangle & = & \begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} & = & \begin{pmatrix}0\\1\\0\\0\end{pmatrix} \\ & & & = & & & \\ |2\rangle & = & |10\rangle & = & \begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} & = & \begin{pmatrix}0\\0\\1\\0\end{pmatrix} \\ |3\rangle & = & |11\rangle & = & \begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} & = & \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \end{array}\right.$$

Let us assume that the initial state $|\psi\rangle_{t=0}$ of our 2-qubit register is

$$|\psi\rangle_{t=0} = \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \otimes |0\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |10\rangle\right) = \frac{1}{\sqrt{2}}\left(|0\rangle - |2\rangle\right) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}$$

Let us also assume that from time $t = 0$ to time $t = 1$ the dynamical behavior of the above 2-qubit register is determined by a constant Hamiltonian H, which when written in terms of the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\} = \{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$ is given by

$$H = \frac{\pi\hbar}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix},$$

where the rows and the columns are listed in the order $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, i.e., in the order $|0\rangle$, $|1\rangle$, $|2\rangle$, $|3\rangle$.

Then, as a consequence of Schrödinger's equation, the Hamiltonian H determines a unitary transformation

$$U_{CNOT} = {}_{t}\mathcal{S}_{0}\, e^{-\frac{i}{\hbar}Hdt} = e^{\int_0^1 -\frac{i}{\hbar}Hdt} = e^{-\frac{i}{\hbar}H}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \left|0\right\rangle \left\langle 0\right| + \left|1\right\rangle \left\langle 1\right| + \left|2\right\rangle \left\langle 3\right| + \left|3\right\rangle \left\langle 2\right|$$

which moves the 2-qubit register from the initial state $\left|\psi\right\rangle_{t=0}$ at time $t = 0$ to $\left|\psi\right\rangle_{t=1} = U_{CNOT} \left|\psi\right\rangle_{t=0}$ at time $t = 1$. Thus,

$$\left|\psi\right\rangle_{t=1} = U_{CNOT} \left|\psi\right\rangle_{t=0} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \left(\left|00\right\rangle - \left|11\right\rangle\right) = \frac{1}{\sqrt{2}} \left(\left|0\right\rangle - \left|3\right\rangle\right)$$

The resulting state (called an **EPR pair** of qubits for reasons we shall later explain) can no longer be written as a tensor product of two states. Consequently, we no longer have the juxtaposition of two qubits.

Somehow, the resulting two qubits have in some sense "lost their separate identities." Measurement of any one of the qubits immediately impacts the other.

For example, if we measure the 0-th qubit (i.e., the right-most qubit), the EPR state in some sense "jumps" to one of two possible states. Each of the two possibilities occurs with probability $\frac{1}{2}$, as indicated in the table below:

<table>
<tr><td colspan="3">$\frac{1}{\sqrt{2}} \left(\left|0_1 0_0\right\rangle - \left|1_1 1_0\right\rangle\right)$</td></tr>
<tr><td>↙↙↙</td><td>Meas.
0-th
Qubit</td><td>↘↘↘</td></tr>
<tr><td>$Prob = \frac{1}{2}$
$\left|0_1 0_0\right\rangle$</td><td colspan="2">$Prob = \frac{1}{2}$
$\left|1_1 1_0\right\rangle$</td></tr>
</table>

So we see that a measurement of one of the qubits causes a change in the other.

7.4. Definition of quantum entanglement.

The above mentioned phenomenon is so unusual and so non-classical that it warrants a name.

Definition 5. *Let $\mathcal{Q}_1$, $\mathcal{Q}_2$, ... , $\mathcal{Q}_n$ be quantum systems with underlying Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$, ... , $\mathcal{H}_n$, respectively. Then the global quantum system $\mathcal{Q}$ consisting of the quantum systems $\mathcal{Q}_1$, $\mathcal{Q}_2$, ... , $\mathcal{Q}_n$ is said to be* ***entangled*** *if its state $|\psi\rangle \in \mathcal{H} = \bigotimes_{j=1}^n \mathcal{H}_j$ can not be written in the form*

$$|\psi\rangle = \bigotimes_{j=1}^{n} |\psi_j\rangle \ ,$$

where each ket $|\psi_j\rangle$ lies in the Hilbert space $\mathcal{H}_j$ for, $j = 1, 2, \ldots, n$. We also say that such a state $|\psi\rangle$ is ***entangled****.*

Thus, the state

$$|\psi\rangle_{t=1} = \frac{1}{\sqrt{2}} \left(|00\rangle - |11\rangle \right)$$

of the 2-qubit register of the previous section is entangled.

Remark 10. *In terms of density operator formalism, a pure ensemble ρ is entangled if it can not be written in the form*

$$\rho = \bigotimes_{j=1}^{n} \rho_j \ ,$$

where the ρ_j's denote density operators.

Please note that we have defined entanglement only for pure ensembles. For mixed ensembles, entanglement is not well understood[15]. As a result, the "right" definition of entanglement of mixed ensembles is still unresolved. We give one definition below:

Definition 6. *A density operator ρ on a Hilbert space $\mathcal{H}$ is said to be entangled with respect to the Hilbert space decomposition*

$$\mathcal{H} = \bigotimes_{j=1}^{n} \mathcal{H}_j$$

if it can not be written in the form

$$\rho = \sum_{k=1}^{\ell} \lambda_k \left(\bigotimes_{j=1}^{n} \rho_{(j,k)} \right) \ ,$$

for some positive integer ℓ, where the λ_k's are positive real numbers such that

$$\sum_{k=1}^{\ell} \lambda_k = 1 \ .$$

and where each $\rho_{(j,k)}$ is a density operator on the Hilbert space $\mathcal{H}_j$.

[15] Quantum entanglement is not even well understood for pure ensembles.

Readers interested in pursuing this topic further should refer to the works of Bennett , the Horodecki's, Nielsen, Smolin, Wootters , and others[**8**], [**51**], [**65**], [**76**], [**1**].

7.5. Einstein, Podolsky, Rosen's (EPR's) grand challenge to quantum mechanics.

Albert Einstein was skeptical of quantum mechanics, so skeptical that he together with Podolsky and Rosen wrote a joint paper[**29**] appearing in 1935 challenging the very foundations of quantum mechanics. Their paper hit the scientific community like a bombshell. For it delivered a direct frontal attack at the very heart and center of quantum mechanics.

At the core of their objection was quantum entanglement. Einstein and his colleagues had insightfully recognized the central importance of this quantum phenomenon.

Their argument centered around the fact that quantum mechanics violated either the **principle of non-locality**[16] or the **principle of reality**[17] . They argued that, as a result, quantum mechanics must be incomplete, and that quantum entanglement could be explained by missing **hidden variables**.

For many years, no one was able to conceive of an experiment that could determine which of the two theories, i.e., quantum mechanics or EPR's hidden variable theory, was correct. In fact, many believed that the two theories were not distinguishable on physical grounds.

It was not until Bell developed his famous inequalities [**5**],[**6**], [**15**], that a physical criterion was found to distinguish the two theories. Bell developed inequalities which, if violated, would clearly prove that quantum mechanics is correct, and hidden variable theories are not. Many experiments were performed[18]. Each emphatically supported quantum mechanics, and clearly demonstrated the incorrectness of hidden variable theory. Quantum mechanics was the victor!

7.6. Why did Einstein, Podolsky, Rosen (EPR) object?

But why did Einstein and his colleagues object so vehemently to quantum entanglement?

As a preamble to our answer to this question, we note that Einstein and his colleagues were convinced of the validity of the following two physical principles:

1) The **principle of local interactions** , i.e., that all the known forces of nature are local interactions,
2) The **principle of non-locality**, i.e., that spacelike separated regions of spacetime are physically independent of one another.

[16] We will later explain the principle of non-locality. See also [**15**].

[17] For an explanation of the principle of reality as well as the principle of non-localty, please refer, for example, to [**81**], [**15**].

[18] See for example [**2**].

Their conviction in regard to principle 1) was based on the fact that all four known forces of nature, i.e., gravitational, electromagnetic, weak, and strong forces, are **local interactions**. By this we mean:

i) They are mediated by another entity, e.g., graviton, photon, etc.
ii) They propagate no faster than the speed c of light
iii) Their strength drops off with distance

Their conviction in regard to principle 2) was based on the following reasoning:

Two points in spacetime $P_1 = (x_1, y_1, z_1, t_1)$ and $P_2 = (x_2, y_2, z_2, t_2)$ are separated by a **spacelike distance** provided the distance between (x_1, y_1, z_1) and (x_2, y_2, z_2) is greater than $c\,|t_2 - t_1|$, i.e.,

$$Distance\left((x_1, y_1, z_1), (x_2, y_2, z_2)\right) > c\,|t_2 - t_1| \ ,$$

where c denotes the speed of light. In other words, no signal can travel between points that are said to be separated by a spacelike distance unless the signal travels faster than the speed of light. But because of the basic principles of relativity, such superluminal communication is not possible.

Hence we have:

The principle of non-locality: Spacelike separated regions of spacetime are physically independent. In other words, spacelike separated regions can not influence one another.

7.7. EPR's objection.

We now are ready to explain why Einstein and his colleagues objected so vehemently to quantum entanglement. We explain Bohm's simplified version of their argument.

Consider a two qubit quantum system that has been prepared by **Alice**[19] in her laboratory in the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|0_1 0_0\rangle - |1_1 1_0\rangle\right) \ .$$

After the preparation, she decides to keep qubit #1 in her laboratory, but enlists Captain James T. Kirk of the Starship Enterprise to transport qubit #0 to her friend **Bob**[20] who is at some far removed distant part of the universe, such as at a Federation outpost orbiting about the double star Alpha Centauri in the constellation Centaurus.

After Captain Kirk has delivered qubit #0, Alice's two qubits are now separated by a spacelike distance. Qubit #1 is located in her Earth based laboratory. Qubits #0 is located with Bob at a Federation outpost orbiting Alpha Centauri. But the

[19] Alice is a well known personality in quantum computation, quantum cryptography, and quantum information theory.

[20] Bob is another well known personality in quantum computation, quantum cryptography, and quantum information theory.

two qubits are still entangled, even in spite of the fact that they are separated by a spacelike distance.

If Alice now measures qubit #1 (which is located in her Earth based laboratory), then the principles of quantum mechanics force her to conclude that instantly, without any time lapse, both qubits are "effected." As a result of the measurement, both qubits will be either in the state $|0_1 0_0\rangle$ or the state $|1_1 1_0\rangle$, each possibility occurring with probability $1/2$.

This is a non-local "interaction." For,

- The "interaction" occurred without the presence of any force. It was not mediated by anything.
- The measurement produced an instantaneous change, which was certainly faster than the speed of light.
- The strength of the "effect" of the measurement did not drop off with distance.

No wonder Einstein was highly skeptical of quantum entanglement. Yet puzzlingly enough, since no information is exchanged by the process, the principles of general relativity are not violated. As a result, such an "effect" can not be used for superluminal communication.

For a more in-depth discussion of the EPR paradox and the foundations of quantum mechanics, the reader should refer to [**15**].

7.8. Quantum entanglement: The Lie group perspective.

Many aspects of quantum entanglement can naturally be captured in terms of Lie groups and their Lie algebras.

Let

$$\mathcal{H} = \mathcal{H}_{n-1} \otimes \mathcal{H}_{n-2} \otimes \ldots \otimes \mathcal{H}_0 = {}_{n-1}\bigotimes{}_0 \mathcal{H}_j$$

be a decomposition of a Hilbert space $\mathcal{H}$ into the tensor product of the Hilbert spaces $\mathcal{H}_{n-1}$, $\mathcal{H}_{n-2}$, ... ,$\mathcal{H}_0$. Let $\mathbb{U} = \mathbb{U}(\mathcal{H})$, $\mathbb{U}_{n-1} = \mathbb{U}(\mathcal{H}_{n-1})$, $\mathbb{U}_{n-2} = \mathbb{U}(\mathcal{H}_{n-2})$, ... ,$\mathbb{U}_0 = \mathbb{U}(\mathcal{H}_0)$, denote respectively the Lie groups of all unitary transformations on $\mathcal{H}$, $\mathcal{H}_{n-1}$, $\mathcal{H}_{n-2}$, ... ,$\mathcal{H}_0$. Moreover, let $\mathbf{u} = \mathbf{u}(\mathcal{H})$, $\mathbf{u}_{n-1} = \mathbf{u}_{n-1}(\mathcal{H}_{n-1})$, $\mathbf{u}_{n-2} = \mathbf{u}_{n-2}(\mathcal{H}_{n-2})$, ... , $\mathbf{u}_0 = \mathbf{u}_0(\mathcal{H}_0)$ denote the corresponding Lie algebras.

Definition 7. *The* ***local subgroup*** $\mathbb{L} = \mathbb{L}(\mathcal{H})$ *of* $\mathbb{U} = \mathbb{U}(\mathcal{H})$ *is defined as the subgroup*

$$\mathbb{L} = \mathbb{U}_{n-1} \otimes \mathbb{U}_{n-2} \otimes \ldots \otimes \mathbb{U}_0 = {}_{n-1}\bigotimes{}_0 \mathbb{U}_j \ .$$

The elements of $\mathbb{L}$ *are called* ***local unitary transformations*** *. Unitary transformations which are in* $\mathbb{U}$ *but not in* $\mathbb{L}$ *are called* ***global unitary transformations****. The corresponding lie algebra*

$$\ell = \mathbf{u}_{n-1} \boxplus \mathbf{u}_{n-2} \boxplus \ldots \boxplus \mathbf{u}_0$$

*is called the **local Lie algebra**, where '$\boxplus$' denotes the **Kronecker sum***[21].

Local unitary transformations can not entangle quantum systems with respect to the above tensor product decomposition. However, global unitary transformations are those unitary transformations which can and often do produce interactions which entangle quantum systems. This leads to the following definition:

DEFINITION 8. *Two states $|\psi_1\rangle$ and $|\psi_2\rangle$ in $\mathcal{H}$ are said to be **locally equivalent** (or, of the **same entanglement type**) , written*

$$|\psi_1\rangle \underset{local}{\sim} |\psi_2\rangle \ \ ,$$

if there exists a local unitary transformation $U \in \mathbb{L}$ such that

$$U\,|\psi_1\rangle = |\psi_2\rangle \ \ .$$

*The equivalence classes of local equivalence $\underset{local}{\sim}$ are called the **entanglement classes of** $\mathcal{H}$. Two density operators ρ_1 and ρ_2, (and hence, the corresponding two skew Hermitian operators $i\rho_1$ and $i\rho_2$ lying in $\mathbf{u}$) are said to be **locally equivalent** (or, of the **same entanglement type**), written*

$$\rho_1 \underset{local}{\sim} \rho_2 \ ,$$

if there exists a local unitary transformation $U \in \mathbb{L}$ such that

$$Ad_U(\rho_1) = \rho_2 \ ,$$

*where Ad_U denotes the big adjoint representation, i.e., $Ad_U(i\rho) = U(i\rho)U^\dagger$. The equivalence classes under this relation are called **entanglement classes** of the Lie algebra $\mathbf{u}(\mathcal{H})$.*

Thus, the entanglement classes of the Hilbert space $\mathcal{H}$ are just the **orbits** of the group action of $\mathbb{L}(\mathcal{H})$ on $\mathcal{H}$. In like manner, the entanglement classes of the Lie algebra $\mathbf{u}(\mathcal{H})$ are the **orbits** of the big adjoint action of $\mathbb{L}(\mathcal{H})$ on $\mathbf{u}(\mathcal{H})$. Two states are entangled in the same way if and only if they lie in the same entanglement class, i.e., the same orbit.

For example, let us assume that Alice and Bob collectively possess two qubits $\mathcal{Q}_{AB}$ which are in the entangled state

$$|\psi_1\rangle = \frac{|0_B 0_A\rangle + |1_B 1_A\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix} \ ,$$

and moreover that Alice possesses the qubit labeled A, but not the qubit labeled B, and that Bob holds qubit B, but not qubit A. Let us also assume that Alice and Bob are also separated by a spacelike distance. As a result, they can only apply local unitary transformations to the qubits that they possess.

[21] The Kronecker sum $A \boxplus B$ is defined as

$$A \boxplus B = A \otimes \mathbf{1} + \mathbf{1} \otimes B \ ,$$

where $\mathbf{1}$ denotes the identity transformation.

Alice could, for example, apply the local unitary transformation

$$U_A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

to her qubit to move Alice's and Bob's qubits A and B respectively into the state

$$|\psi_2\rangle = \frac{|0_B 1_A\rangle - |1_B 0_A\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} ,$$

Bob also could accomplish the same by applying the local unitary transformation

$$U_B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

to his qubit.

By local unitary transformations, Alice and Bob can move the state of their two qubits to any other state within the same entanglement class. But with local unitary transformations, there is no way whatsoever that Alice and Bob can transform the two qubits into a state lying in a different entanglement class (i.e., a different orbit), such as

$$|\psi_3\rangle = |0_B 0_A\rangle \, .$$

The only way Alice and Bob could transform the two qubits from state $|\psi_1\rangle$ to the state $|\psi_3\rangle$ is for Alice and Bob to come together, and make the two qubits interact with one another via a global unitary transformation such as

$$U_{AB} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

The main objective of this approach to quantum entanglement is to determine when two states lie in the same orbit or in different orbits? In other words, what is needed is a complete set of invariants, i.e., invariants that completely specify all the orbits (i.e., all the entanglement classes). We save this topic for another lecture[**67**].

At first it would seem that state kets are a much better vehicle than density operators for the study of quantum entanglement. After all, state kets are much simpler mathematical objects. So why should one deal with the additional mathematical luggage of density operators?

Actually, density operators have a number of advantages over state kets. The most obvious advantage is that density operators certainly have an upper hand over state kets when dealing with mixed ensembles. But their most important advantage is that the orbits of the adjoint action are actually manifolds, which have a very rich and pliable mathematical structure. Needless to say, this topic is beyond the scope of this paper.

Remark 11. *It should also be mentioned that the mathematical approach discussed in this section by no means captures every aspect of the physical phenomenon of quantum entanglement. The use of ancilla and of classical communication have not been considered. For an in-depth study of the relation between quantum entanglement and classical communication (including catalysis), please refer to* [**1**, Chapter 5]*,* [**76**]*, and to the article by Popescu and Rohrlich in* [**66**].

In regard to describing the locality of unitary operations, we will in Section 10 of this paper have need for a little less precision than that given in the above definitions. So we give the following (unfortunately rather technical) definitions:

Definition 9. *Let $\mathcal{H}$, $\mathcal{H}_{n-1}$, $\mathcal{H}_{n-2}$, $\ldots$,$\mathcal{H}_0$ be as stated above. Let $\mathcal{P} = \{B_\alpha\}$ be a* ***partition*** *of the set of indices $\{0,1,2,\ldots,n-1\}$, i.e., $\mathcal{P}$ is a collection of disjoint subsets B_α of $\{0,1,2,\ldots,n-1\}$, called* ***blocks****, such that $\bigcup_\alpha B_\alpha = \{0,1,2,\ldots,n-1\}$. Then the $\mathcal{P}$-**tensor product decomposition** of $\mathcal{H}$ is defined as*

$$\mathcal{H} = \bigotimes_{B_\alpha \in \mathcal{P}} \mathcal{H}_{B_\alpha} \ ,$$

where

$$\mathcal{H}_{B_\alpha} = \bigotimes_{j \in B_\alpha} \mathcal{H}_j \ ,$$

*for each block B_α in $\mathcal{P}$. Also the subgroup of $\mathcal{P}$-**local unitary transformations** $\mathbb{L}_{\mathcal{P}}(\mathcal{H})$ is defined as the subgroup of local unitary transformations of $\mathcal{H}$ corresponding to the $\mathcal{P}$-tensor decomposition of $\mathcal{H}$.*

We define the ***fineness of a partition*** *$\mathcal{P}$, written $fineness(\mathcal{P})$, as the maximum number of indices in a block of $\mathcal{P}$. We say that a unitary transformation U of $\mathcal{H}$ is* ***sufficiently local*** *if there exists a partition $\mathcal{P}$ with sufficiently small $fineness(\mathcal{P})$ (e.g., $fineness(\mathcal{P}) \le 3$) such that $U \in \mathbb{L}_{\mathcal{P}}(\mathcal{H})$.*

Remark 12. *The above lack of precision is needed because there is no way to know what kind (if any) of quantum computing devices will be implemented in the future. Perhaps we will at some future date be able to construct quantum computing devices that locally manipulate more than 2 or 3 qubits at a time?*

8. Entropy and quantum mechanics

8.1. Classical entropy, i.e., Shannon Entropy.

Let $\mathcal{S}$ be a probability distribution on a finite set $\{s_1, s_2, \ldots, s_n\}$ of elements called **symbols** given by

$$\mathrm{Prob}\,(s_j) = p_j \ ,$$

where $\sum_{j=1}^n p_j = 1$. Let s denote the random variable (i.e., **finite memoryless stochastic source**) that produces the value s_j with probability p_j.

Definition 10. *The* ***classical entropy*** *(also called the* ***Shannon entropy****) $H(S)$ of a probability distribution $\mathcal{S}$ (or of the source s) is defined as:*

$$H(\mathcal{S}) = H(s) = -\sum_{j=1}^{n} p_j \lg(p_j) \ ,$$

where 'lg' denotes the log *to the base 2 .*

Classical entropy $H(\mathcal{S})$ is a measure of the uncertainty inherent in the probability distribution $\mathcal{S}$. Or in other words, it is the measure of the uncertainty of an observer before the source s "outputs" a symbol s_j.

One property of such classical **stochastic sources** we often take for granted is that the output symbols s_j are completely distinguishable from one another. We will see that this is not necessarily the case in the strange world of the quantum.

8.2. Quantum entropy, i.e., Von Neumann entropy.

Let $\mathcal{Q}$ be a quantum system with state given by the density operator ρ.

Then there are many **preparations**

Preparation

$\|\psi_1\rangle$	$\|\psi_2\rangle$	$\ldots$	$\|\psi_n\rangle$
p_1	p_2	$\ldots$	p_n

which will produce the same state ρ. These preparations are classical stochastic sources with classical entropy given by

$$H = -\sum p_j \lg(p_j) \ .$$

Unfortunately, the classical entropy H of a preparation does not necessarily reflect the uncertainty in the resulting state ρ. For two different preparations $\mathcal{P}_1$ and $\mathcal{P}_2$, having different entropies $H(\mathcal{P}_1)$ and $H(\mathcal{P}_2)$, can (and often do) produce the same state ρ. The problem is that the states of the preparation may not be completely physically distinguishable from one another. This happens when the states of the preparation are not orthogonal. (Please refer to the Heisenberg uncertainty principle.)

John von Neumann found that the true measure of quantum entropy can be defined as follows:

Definition 11. *Let $\mathcal{Q}$ be a quantum system with state given by the density operator ρ. Then the* ***quantum entropy*** *(also called the* ***von Neumann entropy****) of $\mathcal{Q}$, written $S(\mathcal{Q})$, is defined as*

$$S(\mathcal{Q}) = -Trace\,(\rho \lg \rho) \ ,$$

where '$\lg \rho$' denotes the log to the base 2 of the operator ρ.

Remark 13. *The operator* $\lg \rho$ *exists and is an analytic map* $\rho \longmapsto \lg \rho$ *given by the power series*

$$\lg \rho = \frac{1}{\ln 2} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(\rho - I)^n}{n}$$

provided that ρ *is sufficiently close to the identity operator* I*, i.e., provided*

$$\|\rho - I\| < 1 \ ,$$

where

$$\|A\| = \sup_{v \in \mathcal{H}} \frac{\|Av\|}{\|v\|} \ .$$

It can be shown that this is the case for all positive definite Hermitian operators of trace 1.

For Hermitian operators ρ *of trace* 1 *which are not positive definite, but only positive semi-definite (i.e., which have a zero eigenvalue), the logarithm* $\lg(\rho)$ *does not exist. However, there exists a sequence* $\rho_1, \rho_2, \rho_3, \ldots$ *of positive definite Hermitian operators of trace* 1 *which converges to* ρ*, i.e., such that*

$$\rho = \lim_{k \longrightarrow \infty} \rho_k$$

It can then be shown that the limit

$$\lim_{k \longrightarrow \infty} \rho_k \lg \rho_k$$

exists.

Hence, $S(\rho)$ *is defined and exists for all density operators* ρ*.*

Quantum entropy is a measure of the uncertainty at the quantum level. As we shall see, it is very different from the classical entropy that arises when a measurement is made.

One important feature of quantum entropy $S(\rho)$ is that it is invariant under the big adjoint action of unitary transformations, i.e.,

$$S\left(\ Ad_U(\rho) \ \right) = S\left(U\rho U^{\dagger}\right) = S(\rho) \ .$$

It follows that, for closed quantum systems, it is a **dynamical invariant.** As the state ρ moves according to Schrödinger's equation, the quantum entropy $S(\rho)$ of ρ remains constant. It does not change unless measurement is made, or, as we shall see, unless we ignore part of the quantum system.

Because of unitary invariance, the quantum entropy can be most easily computed by first diagonalizing ρ with a unitary transformation U, i.e.,

$$U\rho U^{\dagger} = \triangle(\overrightarrow{\lambda}) \ ,$$

where $\triangle(\overrightarrow{\lambda})$ denotes the diagonal matrix with diagonal $\overrightarrow{\lambda} = (\lambda_1, \lambda_2, \ \ldots \ , \lambda_n)$.

Once ρ has been diagonalized , we have

$$\begin{aligned} S(\rho) &= -Trace\left(\triangle(\overrightarrow{\lambda})\lg\triangle(\overrightarrow{\lambda})\right) \\ &= -Trace\,(\ \triangle(\lambda_1\lg\lambda_1,\ \lambda_2\lg\lambda_2,\ \ldots\ ,\ \lambda_n\lg\lambda_n)\ ) \\ &= -\sum_{j=1}^{n}\lambda_j\lg\lambda_j\ , \end{aligned}$$

where the λ_j's are the eigenvalues of ρ, and where $0\lg 0\equiv 0$.

Please note that, because ρ is positive semi-definite Hermitian of trace 1, all the eigenvalues of ρ are non-negative real numbers such that

$$\sum_{j=1}^{n}\lambda_j = 1\ .$$

As an immediate corollary we have that the quantum entropy of a pure ensemble must be zero, i.e.,

$$\boxed{\rho\text{ pure ensemble}\Longrightarrow S(\rho)=0}$$

There is no quantum uncertainty in a pure ensemble. However, as expected, there is quantum uncertainty in mixed ensembles.

8.3. How is quantum entropy related to classical entropy?

But how is classical entropy H related to quantum entropy S?

Let A be an observable of the quantum system $\mathcal{Q}$. Then a measurement with respect to A of $\mathcal{Q}$ produces an eigenvalue a_i with probability

$$p_i = Trace\,(P_{a_i}\rho)\ \ ,$$

where P_{a_i} denotes the projection operator for the eigenspace of the eigenvalue a_i. For example, if a_i is a non-degenerate eigenvalue, then $P_{a_i} = |a_i\rangle\langle a_i|$.

In other words, measurement with respect to A of a quantum system $\mathcal{Q}$ in state ρ can be identified with a classical stochastic source with the eigenvalues a_i as output symbols occurring with probability p_i. We denote this classical stochastic source simply by (ρ, A) .

The two entropies $S(\rho)$ and $H(\rho, A)$ are by no means the same. One is a measure of quantum uncertainty before measurement, the other a measure of the classical uncertainty that results from measurement. The quantum entropy $S(\rho)$ is usually a lower bound for the classical entropy, i.e.,

$$S(\rho)\le H(\rho, A)\ .$$

If A is a complete observable (hence, non-degenerate), and if A is compatible with ρ, i.e., $[\rho, A]=0$, then $S(\rho)=H(\rho, A)$.

8.4. When a part is greater than the whole, then Ignorance = uncertainty.

Let $\mathcal{Q}$ be a multipartite quantum system with constituent parts $\mathcal{Q}_{n-1}, \ldots, \mathcal{Q}_1$, $\mathcal{Q}_0$, and let the density operator ρ denote the state of $\mathcal{Q}$. Then from section 5.6 of this paper we know that the state ρ_j of each constituent "part" $\mathcal{Q}_j$ is given by the partial trace over all degrees of freedom except $\mathcal{Q}_j$, i.e., by

$$\rho_j = \underset{\substack{0 \le k \le n-1 \\ k \ne j}}{Trace}(\rho) \quad .$$

By applying the above partial trace, we are focusing only on the quantum system $\mathcal{Q}_j$, and literally ignoring the remaining constituent "parts" of $\mathcal{Q}$. By taking the partial trace, we have done nothing physical to the quantum system. We have simply ignored parts of the quantum system.

What is surprising is that, by intentionally ignoring "part" of the quantum system, we can in some cases create more quantum uncertainty. This happens when the constituent "parts" of $\mathcal{Q}$ are quantum entangled.

For example, let $\mathcal{Q}$ denote the bipartite quantum system consisting of two qubits $\mathcal{Q}_1$ and $\mathcal{Q}_0$ in the entangled state

$$|\Psi_{\mathcal{Q}}\rangle = \frac{|0_1 0_0\rangle - |1_1 1_0\rangle}{\sqrt{2}} \ .$$

The corresponding density operator $\rho_{\mathcal{Q}}$ is

$$\begin{aligned}\rho_{\mathcal{Q}} &= \frac{1}{2}\left(|0_1 0_0\rangle\langle 0_1 0_0| - |0_1 0_0\rangle\langle 1_1 1_0| - |1_1 1_0\rangle\langle 0_1 0_0| + |1_1 1_0\rangle\langle 1_1 1_0|\right) \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}\end{aligned}$$

Since $\rho_{\mathcal{Q}}$ is a pure ensemble, there is no quantum uncertainty, i.e.,

$$S(\rho_{\mathcal{Q}}) = 0 \ .$$

Let us now focus on qubit #0 (i.e., $\mathcal{Q}_0$). The resulting density operator ρ_0 for qubit #0 is obtained by tracing over $\mathcal{Q}_1$, i.e.,

$$\rho_0 = Trace_1(\rho_{\mathcal{Q}}) = \frac{1}{2}\left(\ |0\rangle\langle 0| + |1\rangle\langle 1|\ \right) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ .$$

Hence, the quantum uncertainty of qubit #0 is

$$S(\rho_0) = 1 \ .$$

Something most unusual, and non-classical, has happened. Simply by ignoring part of the quantum system, we have increased the quantum uncertainty. The quantum uncertainty of the constituent "part" $\mathcal{Q}_0$ is greater than that of he whole quantum system $\mathcal{Q}$. This is not possible in the classical world, i.e., not possible for Shannon entropy. (For more details, see [**17**].)

9. There is much more to quantum mechanics

There is much more to quantum mechanics. For more in-depth overviews, there are many outstanding books. Among such books are [**15**], [**18**], [**28**], [**31**], [**43**], [**47**], [**48**], [**52**], [**54**], [**72**], [**78**], [**75**], [**81**], [**83**], [**84**], and many more.

Part 3. Part of a Rosetta Stone for Quantum Computation

10. The Beginnings of Quantum Computation – Elementary Quantum Computing Devices

We begin this section with some examples of quantum computing devices. By a **quantum computing device**[22] we mean a unitary transformation U that is the composition of finitely many sufficiently local unitary transformations, i.e.,

$$U = U_{n-1}U_{n-2}\dots U_1U_0,$$

where U_{n-1}, U_{n-2}, ... ,U_1 ,U_0 are sufficiently local[23] unitary transformations. Each U_j is called a **computational step** of the device.

Our first examples will be constructed by embedding classical computing devices within the realm of quantum mechanics. We will then look at some other quantum computing devices that are not the embeddings of classical devices.

10.1. Embedding classical (memoryless) computation in quantum mechanics.

One objective in this section is to represent[24] classical computing devices as unitary transformations. Since unitary transformations are invertible, i.e., reversible, it follows that the only classical computing devices that can be represented as such transformations must of necessity be reversible devices. Hence, the keen interest in reversible computation[25].

For a more in depth study of reversible computation, please refer to the work of Bennett and others.

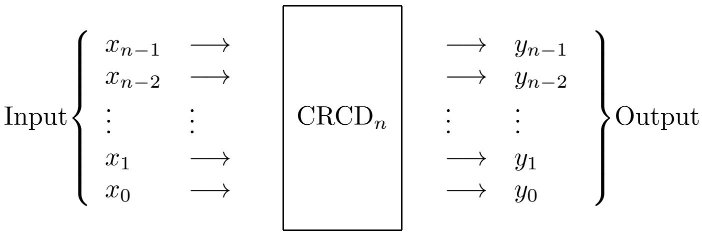
[22] Unfortunately, Physicists have "stolen" the akronym QCD. :-)

[23] See Definition 9 in Section 7.8 of this paper for a definition of the term 'sufficiently local'.

[24] Double meaning is intended.

[25] For references on reversible computation, see [**49**, Chapter 5] and [**77**, Chapter 3].

10.2. Classical reversible computation without memory.



Each **classical** n-input/n-output (binary memoryless) **reversible computing device** (**CRCD**$_n$) can be identified with a bijection

$$\pi : \{0,1\}^n \longrightarrow \{0,1\}^n$$

on the set $\{0,1\}^n$ of all binary n-tuples. Thus, we can in turn identify each CRCD_n with an element of the permutation group S_{2^n} on the 2^n symbols

$$\{ \ |\overrightarrow{a}\rangle \quad | \quad \overrightarrow{a} \in \{0,1\}^n \ \} \ .$$

Let

$$\mathcal{B}_n = \mathcal{B}\left\langle x_0, x_1, \ \ldots \ , x_{n-1} \right\rangle$$

denote the **free Boolean ring** on the symbols $x_0, x_1, \ \ldots \ , x_{n-1}$. Then the binary n-tuples $\overrightarrow{a} \in \{0,1\}^n$ are in one-to-one correspondence with the **minterms** of $\mathcal{B}_n$, i.e.,

$$\overrightarrow{a} \longleftrightarrow x^{\overrightarrow{a}} = \prod_{j=0}^{n-1} x_j^{a_j} \ ,$$

where

$$\begin{cases} x_j^0 &= \overline{x}_j \\ \\ x_j^1 &= x_j \end{cases}$$

Since there is a one-to-one correspondence between the automorphisms of $\mathcal{B}_n$ and the permutations on the set of minterms, it follows that CRCD_n's can also be identified with the **automorphism group** $Aut\left(\mathcal{B}_n\right)$ of the free Boolean ring $\mathcal{B}_n$.

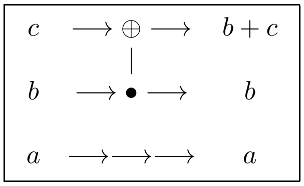
Moreover, since the set of binary n-tuples $\{0,1\}^n$ is in one-to-one correspondence with the set of integers $\{0,1,2, \ \ldots \ , 2^n-1\}$ via the radix 2 representation of integers, i.e.,

$$\left(b_{n-1}, b_{n-2}, \ \ldots \ , b_1, b_0\right) \longleftrightarrow \sum_{j=0}^{n-1} b_j 2^j \ ,$$

we can, and frequently do, identify binary n-tuples with integers.

For example, consider the Controlled-NOT gate, called **CNOT** , which is defined by the following **wiring diagram**:

CNOT =



,

where '$\bullet$' and '$\oplus$' denote respectively a **control bit** and a **target bit**, and where '$a+b$' denotes the exclusive 'or' of bits a and b. This corresponds to the permutation $\pi = (26)(37)$, i.e.,

$$\left\{ \begin{array}{llcll} |0\rangle = & |000\rangle & \longmapsto & |000\rangle & = |0\rangle \\ |1\rangle = & |001\rangle & \longmapsto & |001\rangle & = |1\rangle \\ |2\rangle = & |010\rangle & \longmapsto & |110\rangle & = |6\rangle \\ |3\rangle = & |011\rangle & \longmapsto & |111\rangle & = |7\rangle \\ & & & & \\ |4\rangle = & |100\rangle & \longmapsto & |100\rangle & = |4\rangle \\ |5\rangle = & |101\rangle & \longmapsto & |101\rangle & = |5\rangle \\ |6\rangle = & |110\rangle & \longmapsto & |010\rangle & = |2\rangle \\ |7\rangle = & |111\rangle & \longmapsto & |011\rangle & = |3\rangle \end{array} \right. ,$$

where we have used the following indexing conventions:

$$\left\{ \begin{array}{l} \text{First=Right=Bottom} \\ \\ \text{Last=Left=Top} \end{array} \right.$$

As another example, consider the **Toffoli** gate , which is defined by the following wiring diagram:

$$\textbf{Toffoli} = \boxed{\begin{array}{ccccc} c & \longrightarrow & \oplus & \longrightarrow & c+ab \\ & & | & & \\ b & \longrightarrow & \bullet & \longrightarrow & b \\ & & | & & \\ a & \longrightarrow & \bullet & \longrightarrow & a \end{array}} \ ,$$

where 'ab' denotes the logical 'and' of a and b. As before, '$+$' denotes exclusive 'or'. This gate corresponds to the permutation $\pi = (67)$.

In summary, we have:

$$\boxed{\{\ CRCD_n\ \} = S_2^n = Aut\left(\mathcal{B}_n\right)}$$

10.3. Embedding classical irreversible computation within classical reversible computation.

A classical 1-input/n-output (binary memoryless) irreversible computing device can be thought of as a Boolean function $f = f(x_{n-2}, \ldots, x_1, x_0)$ in $\mathcal{B}_{n-1} = \mathcal{B}\left\langle x_0, x_1, \ldots, x_{n-2}\right\rangle$. Such irreversible computing devices can be transformed into reversible computing devices via the monomorphism

$$\iota : \mathcal{B}_{n-1} \longrightarrow Aut(\mathcal{B}_n),$$

where $\iota(f)$ is the automorphism in $Aut(\mathcal{B}_n)$ defined by

$$(x_{n-1}, x_{n-2}, \ldots, x_1, x_0) \longmapsto (x_{n-1} \oplus f, x_{n-2}, \ldots, x_1, x_0)\,,$$

and where '$\oplus$' denotes exclusive 'or'. Thus, the image of each Boolean function f is a product of disjoint transpositions in S_{2^n}.

As an additive group (ignoring ring structure), $\mathcal{B}_{n-1}$ is the abelian group $\bigoplus_{j=0}^{2^{(n-1)}-1} \mathbb{Z}_2$, where $\mathbb{Z}_2$ denotes the cyclic group of order two.

Classical Binary Memoryless Computation is summarized in the table below:

Summary Classical Binary Memoryless Computation
$\mathcal{B}_{n-1} = \bigoplus_{j=0}^{2^{(n-1)}-1} \mathbb{Z}_2 \xrightarrow{\iota} S_{2^n} = Aut(\mathcal{B}_n)$

10.4. The unitary representation of reversible computing devices.

It is now a straight forward task to represent CRCD$_n$'s as unitary transformations. We simply use the **standard unitary representation**

$$\boxed{\nu : S_2^n \longrightarrow \mathbb{U}(2^n;\mathbb{C})}$$

of the symmetric group S_{2^n} into the group of $2^n \times 2^n$ unitary matrices $\mathbb{U}(2^n;\mathbb{C})$. This is the representation defined by

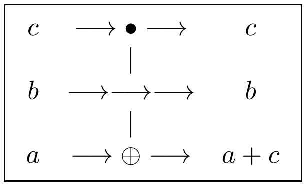
$$\pi \longmapsto \left(\delta_{k,\pi k}\right)_{2^n \times 2^n} \ ,$$

where $\delta_{k\ell}$ denotes the Kronecker delta, i.e.,

$$\delta_{k\ell} = \begin{cases} 1 & \text{if } k = \ell \\ 0 & \text{otherwise} \end{cases}$$

We think of such unitary transformations as quantum computing devices.

For example, consider the controlled-NOT gate $\mathbf{CNOT}' = (45)(67) \in S_8$ given by the wiring diagram

$\mathbf{CNOT}' =$ 

This corresponds to the unitary transformation

$$U_{\mathbf{CNOT}'} = \nu(\mathbf{CNOT}') = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Moreover, consider the Toffoli gate **Toffoli**$'$ $= (57) \in S_8$ given by the wiring diagram

$$\mathbf{Toffoli}' = \boxed{\begin{array}{ccc} c & \longrightarrow \bullet \longrightarrow & c \\ & | & \\ b & \longrightarrow \oplus \longrightarrow & b+ac \\ & | & \\ a & \longrightarrow \bullet \longrightarrow & a \end{array}}$$

This corresponds to the unitary transformation

$$U_{\mathbf{Toffoli}'} = \nu(\mathbf{Toffoli}') = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Abuse of Notation and a Caveat: Whenever it is clear from context, we will use the name of a CRCD$_n$ to also refer to the unitary transformation corresponding to the CRCD$_n$. For example, we will denote $\nu(CNOT)$ and $\nu(Toffoli)$ simply by $CNOT$ and $Toffoli$. Moreover we will also use the wiring diagram of a CRCD$_n$ to refer to the unitary transformation corresponding to the CRCD$_n$. For quantum computation beginners, this can lead to some confusion. Be careful!

10.5. Some other simple quantum computing devices.

After CRCD$_n$'s are embedded as quantum computing devices, they are no longer classical computing devices. After the embedding, they suddenly have acquired much more computing power. Their inputs and outputs can be a superposition of many states. They can entangle their outputs. It is misleading to think of their input qubits as separate, for they could be entangled.

As an illustration of this fact, please note that the quantum computing device **CNOT**$''$ given by the wiring diagram

$$\mathbf{CNOT}'' = \boxed{\begin{array}{ccc} b & \longrightarrow \bullet \longrightarrow & a+b \\ & | & \\ a & \longrightarrow \oplus \longrightarrow & a \end{array}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

is far from classical. It is more than a permutation. It is a linear operator that respects quantum superposition.

For example, **CNOT**$''$ can take two non-entangled qubits as input, and then produce two entangled qubits as output. This is something no classical computing device can do. For example,

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes |0\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |10\rangle\right) \longmapsto \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right)$$

For completeness, we list two other quantum computing devices that are embeddings of CRCD$_n$'s, **NOT** and **SWAP**:

$$\mathbf{NOT} = \boxed{a \quad \longrightarrow \boxed{\mathbf{NOT}} \longrightarrow \quad a+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1$$

and

$$\mathbf{SWAP} = \boxed{\begin{array}{ccccccccccc} b & \longrightarrow & \bullet & \longrightarrow & \longrightarrow & \oplus & \longrightarrow & \longrightarrow & \bullet & \longrightarrow & a \\ & & | & & & | & & & | & & \\ a & \longrightarrow & \oplus & \longrightarrow & \longrightarrow & \bullet & \longrightarrow & \longrightarrow & \oplus & \longrightarrow & b \end{array}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

10.6. Quantum computing devices that are not embeddings.

We now consider quantum computing devices that are not embeddings of CRCD$_n$'s.

The **Hadamard** gate **H** is defined as:

$$\mathbf{H} = \boxed{\longrightarrow \mathbf{H} \longrightarrow} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} .$$

Another quantum gate is the **square root of NOT**, i.e., $\sqrt{\mathbf{NOT}}$, which is given by

$$\sqrt{\mathbf{NOT}} = \boxed{\longrightarrow \sqrt{\mathbf{NOT}} \longrightarrow} = \frac{1-i}{2} \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix} = \frac{1+i}{2} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} .$$

There is also the **square root of swap** $\sqrt{\mathbf{SWAP}}$ which is defined as:

$$\sqrt{\mathbf{SWAP}} = \boxed{\longrightarrow \sqrt{\mathbf{SWAP}} \longrightarrow} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1+i}{2} & \frac{1-i}{2} & 0 \\ 0 & \frac{1-i}{2} & \frac{1+i}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} .$$

Three frequently used unary quantum gates are the rotations:

$$\boxed{\longrightarrow \boxed{e^{i\theta\sigma_1}} \longrightarrow} = \begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix} = e^{i\theta\sigma_1}$$

$$\boxed{\longrightarrow \boxed{e^{i\theta\sigma_2}} \longrightarrow} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = e^{i\theta\sigma 2}$$

$$\boxed{\longrightarrow \boxed{e^{i\theta\sigma_3}} \longrightarrow} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} = e^{i\theta\sigma_3}$$

10.7. The implicit frame of a wiring diagram.

Wiring diagrams have the advantage of being a simple means of describing some rather complicated unitary transformations. However, they do have their drawbacks, and they can, if not used with care, be even misleading.

One problem with wiring diagrams is that they are not frame (i.e., basis) independent descriptions of unitary transformations. Each wiring diagram describes a unitary transformation using an implicitly understood basis.

For example, consider $\mathbf{CNOT}''$ given by the wiring diagram:

$$\mathbf{CNOT}'' = \boxed{\begin{array}{ccccc} b & \longrightarrow & \bullet & \longrightarrow & a+b \\ & & | & & \\ a & \longrightarrow & \oplus & \longrightarrow & a \end{array}} \; .$$

The above wiring diagram defines $\mathbf{CNOT}''$ in terms of the implicitly understood basis

$$\left\{ |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} .$$

This wiring diagram suggests that qubit #1 controls qubit #0, and that qubit #1 is not effected by qubit #0. But this is far from the truth. For, $\mathbf{CNOT}''$ transforms

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

into

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} ,$$

where we have used our indexing conventions

$$\left\{ \begin{array}{l} \text{First=Right=Bottom} \\ \\ \text{Last=Left=Top} \end{array} \right. .$$

In fact, in the basis

$$\left\{ |0'\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}},\ |1'\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} \right\}$$

the wiring diagram of the same unitary transformation $\mathbf{CNOT}''$ is:

$$\boxed{\begin{array}{ccccc} b & \longrightarrow & \oplus & \longrightarrow & a+b \\ & & | & & \\ a & \longrightarrow & \bullet & \longrightarrow & a \end{array}}$$

The roles of the target and control qubits appeared to have switched!

11. The No-Cloning Theorem

In this section, we prove the no-cloning theorem of Dieks[**24**], Wootters and Zurek [**93**]. The theorem states that there can be no device that produces exact replicas or copies of a quantum state.

In mathematical terms, a device which replicates quantum states, i.e., a quantum replicator, is defined as follows:

DEFINITION 12. *Let $\mathcal{H}$ be a Hilbert space. Then a* ***quantum replicator*** *for $\mathcal{H}$ consists of an auxiliary Hilbert space $\mathcal{H}_A$, a fixed state $|\psi_0\rangle \in \mathcal{H}_A$ (called the* ***initial state of replicator****), and a unitary transformation*

$$U : \mathcal{H}_A \otimes \mathcal{H} \otimes \mathcal{H} \longrightarrow \mathcal{H}_A \otimes \mathcal{H} \otimes \mathcal{H}$$

such that, for some fixed state $|blank\rangle \in \mathcal{H}$,

$$U\,|\psi_0\rangle\,|a\rangle\,|blank\rangle = |\psi_a\rangle\,|a\rangle\,|a\rangle \ \ ,$$

for all states $|a\rangle \in \mathcal{H}$, where $|\psi_a\rangle \in \mathcal{H}_A$ (called the ***replicator state after replication*** *of $|a\rangle$) depends on $|a\rangle$.*

THEOREM 3 (No-Cloning). *Let $\mathcal{H}$ be a Hilbert space of dimension greater than one. Then a quantum replicator for $\mathcal{H}$ does not exist.*

The proof of the no-cloning theorem, i.e., the proof that quantum replicators do not exist, is an amazingly simple application of the linearity of quantum mechanics. The key idea is that copying is an inherently nonlinear transformation, while the unitary transformations of quantum mechanics are inherently linear. Ergo, copying can not be a unitary transformation.

More specifically, the proof goes as follows:

PROOF. Since a quantum state is determined by a ket up to a multiplicative non-zero complex number, we can without loss of generality assume that $|\psi_0\rangle$, $|a\rangle$, $|blank\rangle$ are all of unit length. From unitarity, it follows that $|\psi_a\rangle$ is also of unit length.

Let $|a\rangle$, $|b\rangle$ be two kets of unit length in $\mathcal{H}$ such that

$$0 < |\,\langle\ a \mid b\ \rangle\,| < 1\ .$$

Then

$$\begin{cases} U\,|\psi_0\rangle\,|a\rangle\,|blank\rangle & = & |\psi_a\rangle\,|a\rangle\,|a\rangle \\ U\,|\psi_0\rangle\,|b\rangle\,|blank\rangle & = & |\psi_b\rangle\,|b\rangle\,|b\rangle \end{cases}$$

Hence,

$$\begin{aligned}\langle blank|\,\langle a|\,\langle\psi_0|\,U^\dagger U\,|\psi_0\rangle\,|b\rangle\,|blank\rangle &= \langle blank|\,\langle a|\,\langle\ \psi_0 \mid \psi_0\ \rangle\,|b\rangle\,|blank\rangle \\ &= \langle\ a \mid b\ \rangle\end{aligned}$$

On the other hand,

$$\begin{aligned}\langle blank|\,\langle a|\,\langle\psi_0|\,U^\dagger U\,|\psi_0\rangle\,|b\rangle\,|blank\rangle &= \langle a|\,\langle a|\,\langle\ \psi_a\mid\psi_b\ \rangle\,|b\rangle\,|b\rangle \\ &= \langle\ a\mid b\ \rangle^2\,\langle\ \psi_a\mid\psi_b\ \rangle\end{aligned}$$

Thus,

$$\langle\ a\mid b\ \rangle^2\,\langle\ \psi_a\mid\psi_b\ \rangle = \langle\ a\mid b\ \rangle\ \ .$$

And so,

$$\langle\ a\mid b\ \rangle\,\langle\ \psi_a\mid\psi_b\ \rangle = 1\ .$$

But this equation can not be satisfied since

$$|\langle\ a\mid b\ \rangle| < 1$$

and

$$|\langle\ \psi_a\mid\psi_b\ \rangle| \le \|\ |\psi_a\rangle\ \|\,\|\ |\psi_b\rangle\ \| = 1$$

Hence, a quantum replicator cannot exist. □

Exercise 2. *Although it is not possible to clone* ***all*** *states in $\mathcal{H}$ (emphasis on key word "all"), it is nonetheless still possible to clone all states of a subset of $\mathcal{H}$ consisting of mutually orthogonal states.*

Let $\{|0\rangle,\ldots,|n-1\rangle\}$ be an orthonormal basis of the Hilbert space $\mathcal{H}$. Construct a unitary transformation U such that $U:|k\rangle\,|blank\rangle\longmapsto|k\rangle\,|k\rangle$, for $0\le k<n$.

Exercise 3. *Is cloning possible on a one dimensional Hilbert space $\mathcal{H}$? Please explain your answer.*

12. Quantum teleportation

We now give a brief description of quantum teleportation, a means possibly to be used by future quantum computers to bus qubits from one location to another.

The no-cloning theorem emphatically states that qubits cannot be copied. However, ... qubits can be teleported, as has been demonstrated in laboratory settings. Such a mechanism could be used to bus qubits from one computer location to another. It could be used to create devices called **quantum repeaters**.

But what do we mean by teleportation?

Teleportation is the transferring of an object from one location to another by a process that:

1) Firstly dissociates (i.e., destroys) the object to obtain information. – The object to be teleported is first scanned to extract sufficient information to reassemble the original object.
2) Secondly transmits the acquired information from one location to another.
3) Lastly reconstructs the object at the new location from the received information. – An exact replica is re-assembled at the destination out of locally available materials.

Two key effects of teleportation should be noted:

1) The original object is destroyed during the process of teleportation. Hence, the no-cloning theorem is not violated.
2) An exact replica of the original object is created at the intended destination.

Scotty of the Starship Enterprise was gracious enough to loan me the following teleportation manual. So I am passing it on to you.

Quantum Teleportation Manual

Step. 1 .(Location A) Preparation: At location A, construct an EPR pair of qubits (qubits #2 and #3) in $\mathcal{H}_2 \otimes \mathcal{H}_3$.

$$\begin{tikzcd}
\left|00\right\rangle \arrow[r, mapsto] & \boxed{\begin{matrix}\text{Unitary}\\ \text{Matrix}\end{matrix}} \arrow[r, mapsto] & \frac{\left|01\right\rangle - \left|10\right\rangle}{\sqrt{2}} \\
\mathcal{H}_2 \otimes \mathcal{H}_3 \arrow[rr] & & \mathcal{H}_2 \otimes \mathcal{H}_3
\end{tikzcd}$$

Step 2. Transport: Physically transport entangled qubit #3 from location A to location B.

Step 3. : The qubit to be teleported, i.e., qubit #1, is delivered to location A in an unknown state

$$a\left|0\right\rangle + b\left|1\right\rangle$$

As a result of Steps 1 to 3, we have:

- Locations A and B share an EPR pair, i.e.
  - The qubit to be teleported, i.e., qubit #1, is at Location A
  - Qubit #2 is at Location A
  - Qubit #3 is at Location B
  - Qubits #2 & #3 are entangled
- The current state $\left|\Phi\right\rangle$ of all three qubits is:

$$\left|\Phi\right\rangle = \left(a\left|0\right\rangle + b\left|1\right\rangle\right)\left(\frac{\left|01\right\rangle - \left|10\right\rangle}{\sqrt{2}}\right) \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$$

To better understand what is about to happen, we re-express the state $|\Phi\rangle$ of the three qubits in terms of the following basis (called the **Bell basis**) of $\mathcal{H}_1 \otimes \mathcal{H}_2$:

$$\begin{cases} |\Psi_A\rangle & = & \frac{|10\rangle - |01\rangle}{\sqrt{2}} \\ \\ |\Psi_B\rangle & = & \frac{|10\rangle + |01\rangle}{\sqrt{2}} \\ \\ |\Psi_C\rangle & = & \frac{|00\rangle - |11\rangle}{\sqrt{2}} \\ \\ |\Psi_D\rangle & = & \frac{|00\rangle + |11\rangle}{\sqrt{2}} \end{cases}$$

The result is:

$$|\Phi\rangle \quad = \quad \begin{array}{c} \frac{1}{2}[\quad |\Psi_A\rangle\left(-a\,|0\rangle - b\,|1\rangle\right) \\ \\ +\,|\Psi_B\rangle\left(-a\,|0\rangle + b\,|1\rangle\right) \\ \\ +\,|\Psi_C\rangle\left(a\,|1\rangle + b\,|0\rangle\right) \\ \\ +\,|\Psi_D\rangle\left(a\,|1\rangle - b\,|0\rangle\right) \quad] \end{array} \quad ,$$

where, as you might have noticed, we have written the expression in a suggestive way.

REMARK 14. *Please note that since the completion of Step 3, we have done nothing physical. We have simply performed some algebraic manipulations of the expression representing the state $|\Phi\rangle$ of the three qubits.*

Let $U : \mathcal{H}_1 \otimes \mathcal{H}_2 \longrightarrow \mathcal{H}_1 \otimes \mathcal{H}_2$ be the unitary transformation defined by

$$\begin{cases} |\Psi_A\rangle & \longmapsto & |00\rangle \\ |\Psi_B\rangle & \longmapsto & |01\rangle \\ |\Psi_C\rangle & \longmapsto & |10\rangle \\ |\Psi_D\rangle & \longmapsto & |11\rangle \end{cases}$$

Step 4. (Location A): [26]Apply the local unitary transformation $U \otimes I : \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3 \longrightarrow \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ to the three qubits (actually more precisely, to qubits #1 and #2). Thus, under $U \otimes I$ the state $|\Phi\rangle$ of all

[26] Actually, there is no need to apply the unitary transformation U. We could have instead made a complete Bell state measurement, i.e., a measurement with respect to the compatible observables $|\Psi_A\rangle\langle\Psi_A|$, $|\Psi_B\rangle\langle\Psi_B|$, $|\Psi_C\rangle\langle\Psi_C|$, $|\Psi_D\rangle\langle\Psi_D|$. We have added the additional step 4 to make quantum teleportation easier to understand for quantum computation beginners.

three qubits becomes

$$|\Phi'\rangle = \begin{array}{l} \frac{1}{2}[\quad |00\rangle\left(-a\,|0\rangle - b\,|1\rangle\right) \\ \quad + |01\rangle\left(-a\,|0\rangle + b\,|1\rangle\right) \\ \quad + |10\rangle\left(a\,|1\rangle + b\,|0\rangle\right) \\ \quad + |11\rangle\left(a\,|1\rangle - b\,|0\rangle\right) \quad] \end{array}$$

Step 5. (Location A): Measure qubits #1 and #2 to obtain two bits of classical information. The result of this measurement will be one of the bit pairs $\{00, 01, 10, 11\}$.

Step 6.: Send from location A to location B (via a classical communication channel) the two classical bits obtained in Step 5.

As an intermediate summary, we have:

1) Qubit #1 has been disassembled, and
2) The information obtained during disassembly (two classical bits) has been sent to location B.

Step 7. (Location B): The two bits (i,j) received from location A are used to select from the following table a unitary transformation $U^{(i,j)}$ of $\mathcal{H}_3$, (i.e., a local unitary transformation $I_4 \otimes U^{(i,j)}$ on $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$)

Rec. Bits	$U^{(i,j)}$	Future effect on qubit #3
00	$U^{(00)} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$-a\,|0\rangle - b\,|1\rangle \longmapsto a\,|0\rangle + b\,|1\rangle$
01	$U^{(01)} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$-a\,|0\rangle + b\,|1\rangle \longmapsto a\,|0\rangle + b\,|1\rangle$
10	$U^{(10)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$a\,|1\rangle + b\,|0\rangle \longmapsto a\,|0\rangle + b\,|1\rangle$
11	$U^{(11)} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$a\,|1\rangle - b\,|0\rangle \longmapsto a\,|0\rangle + b\,|1\rangle$

Step 8. (Location B): The unitary transformation $U^{(i,j)}$ selected in Step 7 is applied to qubit #3.

As a result, qubit #3 is at location B and has the original state of qubit #1 when qubit #1 was first delivered to location A, i.e., the state

$$a\,|0\rangle + b\,|1\rangle$$

It is indeed amazing that no one knows the state of the quantum teleported qubit except possibly the individual that prepared the qubit. Knowledge of the actual state of the qubit is not required for teleportaton. If its state is unknown before the teleportation, it remains unknown after the teleportation. All that we know is that the states before and after the teleportation are the same.

13. There is much more to quantum computation

Needles to say, there is much more to quantum computation. I hope that you found this introductory paper useful. For further reading on quantum computation, we refer the reader to the many informative books on subject, such as [**1**], [**41**], [**50**], [**66**], and [**77**].

References

[1] Alber, G., T. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Rotteler, H. Weinfurther, R. Werner, and A. Zeilinger, **"Quantum Information: An Introduction to Basic Theorectical Concepts and Experiments,"** Springer-Verlag, (2001).

[2] Aspect, A., J. Dalibard, and G. Roger, Phys. Rev. Lett., 49, (1982), p 1804.

[3] Barenco, A, C.H. Bennett, R. Cleve, D.P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J.A. Smolin, and H. Weinfurter, **Elementary gates for quantum computation**, Phys. Rev. A, **52**, (1995), pp 3475 - 3467.

[4] Beardon, Alan F., "The Geometry of Discrete Groups," Springer-Verlag, (1983).

[5] Bell, J.S., Physics, 1, (1964), pp 195 - 200.

[6] Bell, J.S., "**Speakable and Unspeakable in Quantum Mechanics**," Cambridge University Press (1987).

[7] Bennett, C.H. et al., Phys. Rev. Lett. **70**, (1995), pp 1895.

[8] Bennett, C.H., D.P. DiVincenzo, J.A. Smolin, and W.K. Wootters, Ohys. Rev. A, **54**, (1996), pp 3824.

[9] Berman, Gennady, Gary D. Doolen, Ronnie Mainieri, and Vladimir I. Tsifrinovich, "**Introduction to Quantum Computation**," World Scientific (1999).

[10] Bernstein, Ethan, and Umesh Vazirani, **Quantum complexity theory**, Siam J. Comput., Vol. 26, No.5 (1997), pp 1411 - 1473.

[11] Brandt, Howard E., **Qubit devices and the issue of quantum decoherence**, Progress in Quantum Electronics, Vol. 22, No. 5/6, (1998), pp 257 - 370.

[12] Brandt, Howard E., **Qubit devices**, in "**Quantum Computation**," this AMS Proceedings of Symposia in Applied Mathematics (PSAPM).

[13] Brassard, Gilles, and Paul Bratley, "**Algorithmics: Theory and Practice,"** Printice-Hall, (1988).

[14] Brooks, Michael (Ed.), "**Quantum Computing and Communication**s," Springer-Verlag (1999).

[15] Bub, Jeffrey, "**Interpreting the Quantum World**," Cambridge University Press (1997).

[16] Cartan, Henri, and Samuel Eilenberg, "**Homological Algebra**," Princeton University Press, Princeton, New Jersey, (1956)

[17] Cerf, Nicholas J. and Chris Adami, "**Quantum information theory of entanglement and measurement**," in **Proceedings of Physics and Computation, PhysComp'96**, edited by J. Leao T. Toffoli, pp 65 - 71. See also http://xxx.lanl.gov/abs/quant-ph/9605039 .

[18] Cohen-Tannoudji, Claude, Bernard Diu, and Frank Laloë, "**Quantum Mechanics**," Volumes 1 & 2, John Wiley & Sons (1977)

[19] D'Espagnat, Bernard, "**Veiled Reality: Analysis of Present Day Quantum Mechanical Concepts**," Addison-Wesley (1995)

[20] D'Espagnat, Bernard, "**Conceptual Foundations of Quantum Mechanics**," (Second Edition), Addison-Wesley (1988)

[21] Cormen, Thomas H., Charles E. Leiserson, and Ronald L. Rivest, "**Introduction to Algorithms**," McGraw-Hill, (1990).

[22] Cox, David, John Little, and Donal O'Shea, "**Ideals, Varieties, and Algorithms**," (second edition), Springer-Verlag, (1996).

[23] Davies, E.B., "**Quantum Theory of Open Systems**," Academic Press, (1976).

[24] Dieks, D., Phys. Lett., **92**, (1982), p 271.

[25] Deutsch, David, "**The Fabric of Reality**," Penguin Press, New York (1997).

[26] Deutsch, David, and Patrick Hayden, **Information flow in entangled quantum systems**, http://xxx.lanl.gov/abs/quant-ph/9906007.

[27] Deutsch, David, **Quantum theory, the Church-Turing principle and the universal quantum computer**, Proc. Royal Soc. London A, **400**, (1985), pp 97 - 117.

[28] Dirac, P.A.M., "**The Principles of Quantum Mechanics**," (Fourth edition). Oxford University Press (1858).

[29] Einstein, A., B. Podolsky, N. Rosen, **Can quantum, mechanical description of physical reality be considered complete?**, Phys. Rev. **47**, 777 (1935); D. Bohm "**Quantum Theory**", Prentice-Hall, Englewood Cliffs, NJ (1951).

[30] Ekert, Artur K.and Richard Jozsa, **Quantum computation and Shor's factoring algorithm**, Rev. Mod. Phys., 68,(1996), pp 733-753.

[31] Feynman, Richard P., Robert B. Leighton, and Matthew Sands, "**The Feyman Lectures on Physics: Vol. III. Quantum Mechanics**," Addison-Wesley Publishing Company, Reading, Massachusetts (1965).

[32] Feynman, Richard P., "**Feynman Lectures on Computation**," (Edited by Anthony J.G. Hey and Robin W. Allen), Addison-Wesley, (1996).

[33] Gilmore, Robert, "**Alice in Quantumland**," Springer-Verlag (1995).

[34] Gottesman, Daniel, **The Heisenberg representation of quantum computers**, http://xxx.lanl.gov/abs/quant-ph/9807006.

[35] Gottesman, Daniel, **An introduction to quantum error correction**, in "**Quantum Computation**," this AMS Proceedings of the Symposia in Applied Mathematics (PSAPM). (http://xxx.lanl.gov/abs/quant-ph/0004072)

[36] Gottfried, "**Quantum Mechanics: Volume I. Fundamentals**," Addison-Wesley (1989).

[37] Grover, Lov K., **Quantum computer can search arbitrarily large databases by a single querry**, Phys. Rev. Letters (1997), pp 4709-4712.

[38] Grover, Lov K., **A framework for fast quantum mechanical algorithms**, http://xxx.lanl.gov/abs/quant-ph/9711043.

[39] Grover, L., Proc. 28th Annual ACM Symposium on the Theory of Computing, ACM Press, New Yorkm (1996), pp 212 - 219.

[40] Grover, L., Phys. Rev. Lett. 78, (1997), pp 325 - 328.

[41] Gruska, Jozef, "**Quantum Computing**," McGraw-Hill, (1999)

[42] Gunther, Ludwig, "An Axiomatic Basis for Quantum Mechanics: Volume I. Derivation of Hilbert Space Structure," Springer-Verlag (1985).

[43] Haag, R., "**Local Quantum Physics: Fields, Particles, Algebras**," (2nd revised edition), Springer-Verlag.

[44] Halmos, Paul R., **"Lectures on Boolean Algebras,"** Van Nostrand, (1967).

[45] Halmos, Paul R., **"Finite-Dimensional Vector Spaces,"** Van Nostrand, (1958).

[46] Hardy, G.H., and E.M. Wright, "**An Introduction to the Theory of Numbers**," Oxford Press, (1965).

[47] Heisenberg, Werner, "**The Physical Principles of Quantum Theory**," translated by Eckart and Hoy, Dover.

[48] Helstrom, Carl W., "**Quantum Detection and Estimation Theory**," Academic Press (1976).

[49] Hey, Anthony J.G. (editor), "**Feynman and Computation**," Perseus Books, Reading, Massachusetts, (1998).

[50] Hirvensalo, Mika, **"Quantum Computing,"** Springer-Verlag, (2001).

[51] Horodecki, O., M. Horodecki, and R. Horodecki, Phys. Rev. Lett. **82**, (1999), pp 1056.

[52] Holevo, A.S., "**Probabilistic and Statistical Aspects of Quantum Theory**," North-Holland, (1982).

[53] Hoyer, Peter, **Efficient quantum transforms**, http://xxx.lanl.gov/abs/quant-ph/9702028.

[54] Jauch, Josef M., "**Foundations of Quantum Mechanics**," Addison-Wesley Publishing Company, Reading, Massachusetts (1968).

[55] Jozsa, Richard, **Searching in Grover's Algorithm**, http://xxx.lanl.gov/abs/quant-ph/9901021.

[56] Jozsa, Richard, **Quantum algorithms and the Fourier transform**, quant-ph preprint archive 9707033 17 Jul 1997.

[57] Jozsa, Richard, Proc. Roy. Soc. London Soc., Ser. A, 454, (1998), 323 - 337.

[58] Kauffman, Louis H., **Quantum topology and quantum computing**, in "**Quantum Computation**," this AMS Proceedings of the Symposia in Applied Mathematics (PSAPM).

[59] Kitaev, A., **Quantum measurement and the abelian stabiliser problem,** (1995), http://xxx.lanl.gov/abs/quant-ph/9511026.

[60] Kitaev, Alexei, **Quantum computation with anyons**, this AMS Proceedings of the Symposia in Applied Mathematics (PSAPM).

[61] Lang, Serge, "**Algebra**," Addison- Wesley (1971).

[62] Lenstra, A.K., and H.W. Lenstra, Jr., eds., "**The Development of the Number Field Sieve**," Lecture Notes in Mathematics, Vol. 1554, Springer-Velag, (1993).

[63] Lenstra, A.K., H.W. Lenstra, Jr., M.S. Manasse, and J.M. Pollard, **The number field sieve**. Proc. 22nd Annual ACM Symposium on Theory of ComputingACM, New York, (1990), pp 564 - 572. (See exanded version in Lenstra & Lenstra, (1993), pp 11 - 42.)

[64] LeVeque, William Judson, "**Topics in Number Theory: Volume I**," Addison-Wesley, (1958).

[65] Linden, N., S. Popescu, and A. Sudbery, **Non-local properties of multi-particle density matrices**, http://xxx.lanl.gov/abs/quant-ph/9801076.

[66] Lo, Hoi-Kwong, Tim Spiller & Sandu Popescu(editors), "**Introduction to Quantum Computation & Information**," World Scientific (1998).

[67] Lomonaco, Samuel J., Jr., "**A entangled tale of quantum entanglement,** in "**Quantum Computation**," this AMS Proceedings of the Symposia in Applied Mathematics (PSAPM). (http://xxx.lanl.gov/abs/quant-ph/0101120)

[68] Lomonaco, Samuel J., Jr., **A quick glance at quantum cryptography**, Cryptologia, Vol. 23, No. 1, January,1999, pp 1-41. (http://xxx.lanl.gov/abs/quant-ph/9811056)

[69] Lomonaco, Samuel J., Jr., **A talk on quantum cryptography: How Alice Outwits Eve**, in "**Coding Theory and Cryptography: From Enigma and Geheimsschreiber to Quantum Theory**," edited by David Joyner, Springer-Verlag, (2000), pp 144 - 174. Also in this AMS Proceedings of the Symposia in Applied Mathematics (PSAPM).

[70] Lomonaco, Samuel J., Jr., **Shor's quantum factoring algorithm**, in "**Quantum Computation**," this AMS Proceedings of the Symposia in Applied Mathematics (PSAPM). (http://xxx.lanl.gov/abs/quant-ph/0010034)

[71] Lomonaco, Samuel J., Jr., **Grover's quantum search algorithm**, in "**Quantum Computation**," this AMS Proceedings of the Symposia in Applied Mathematics (PSAPM). (http://xxx.lanl.gov/abs/quant-ph/0010040)

[72] Mackey, George W., "**Mathematical Foundations of Quantum Mechanics**," Addison-Wesley (1963).

[73] Milburn, Gerald J., "**The Feynman Processor**," Perseus Books, Reading, Massachusetts (1998)

[74] Miller, G.L., **Riemann's hypothesis and tests for primality**, J. Comput. System Sci., 13, (1976), pp 300 - 317.

[75] von Neumann, John, "**Mathematical Foundations of Quantum Mechanics**," Princeton University Press.

[76] Nielsen, M.A., **Continuity bounds on entanglement**, Phys. Rev. A, Vol. 61, 064301, pp 1-4.

[77] Nielsen, Michael A., and Isaac L. Chuang, **"Quantum Computation and Quantum Information,"** Cambridge University Press, (2000).

[78] Omnès, Roland, **"An Interpretation of Quantum Mechanics,"** Princeton University Press, Princeton, New Jersey, (1994).

[79] Omnès, Roland, "**Understanding Quantum Mechanics**," Princeton University Press (1999).

[80] Penrose, Roger, "**The Large, the Small and the Human Mind**," Cambridge University Press, (1997).

[81] Peres, Asher, **"Quantum Theory: Concepts and Methods,"** Kluwer Academic Publishers, Boston, (1993).

[82] Raymond, Pierre, "**Field Theory: A Modern Primer**," Addison-Wesley (1989).

[83] Piron, C., "**Foundations of Quantum Physics**," Addison-Wesley, (1976).

[84] Sakurai, J.J., "**Modern Quantum Mechanics**," (Revised edition), Addison-Wesley Publishing Company, Reading, Massachusetts (1994).

[85] Schumacher, Benjamin, **Sending entanglement through noisy quantum channels**, (22 April 1996), http://xxx.lanl.gov/abs/quant-ph/9604023.

[86] Schwinger, Julian, **"Quantum Mechanics: Symbolism of Quantum Measurement,"** Springer-Verlag, (2001).

[87] Shor, Peter W., **Polynomial time algorithms for prime factorization and discrete logarithms on a quantum computer**, SIAM J. on Computing, 26(5) (1997), pp 1484 - 1509. (http://xxx.lanl.gov/abs/quant-ph/9508027)

[88] Shor, Peter W., **Introduction to quantum algorithms**, in "**Quantum Computation**," this AMS Proceedings of the Symposia in Applied Mathematics (PSAPM). (http://xxx.lanl.gov/abs/quant-ph/0005003)

[89] Stinson, Douglas R., "**Cryptography: Theory and Practice**," CRC Press, Boca Raton, (1995).

[90] Vazirani, Umesh, **Quantum complexity theory**, in "**Quantum Computation**," this AMS Proceedings of the Symposia in Applied Mathematics (PSAPM).

[91] Williams, Collin P., and Scott H. Clearwater, "**Explorations in Quantum Computation**," Springer-Verlag (1997).

[92] Williams, Colin, and Scott H. Clearwater, "**Ultimate Zero and One**," Copernicus, imprint by Springer-Verlag, (1998).

[93] Wootters, W.K., and W.H. Zurek, **A single quantum cannot be cloned,** Nature, Vol. 299, 28 October 1982, pp 982 - 983.

University of Maryland Baltimore County, Baltimore, MD 21250
E-mail address: Lomonaco@UMBC.EDU
URL: http://www.csee.umbc.edu/~lomonaco