Quantum Information and Computing- I

Dipan Kumar Ghosh
UM-DAE Centre for Excellence in Basic Sciences
Kalina, Mumbai 400098

April 21, 2016

1 Introduction

In this and the following lecture we will briefly review the postulates of Quantum Mechanics which form the cornerstone of our discussion on quantum computing. Postulates are a set of axioms which cannot be proved but are consistent with experimental observation. There are no standard on what constitutes the minimum set of such norms and it depends a lot on one's own point of view. Later, we will briefly review topics in linear algebra which will be useful for our course.

2 Postulates of Quantum Mechanics

Postulate - 1:
Physical states are represented by *state vectors* which provide a complete description of the physical system under discussion. These are vectors lie in an abstract space known as the **Hilbert space**. A Hilbert space is a *normed linear vector space* defined over a field of complex numbers. A state is denoted by a *ket* in Dirac notation. A physical state is described by a *ray* in the Hilbert space, which consists of vectors which differ from one another by a scalar multiplying factor, which, in general, is complex. Corresponding to a ket vector there exists a vector in its dual space known as a *bra* vector, the bra corresponding to the state $c \mid \psi\rangle$ is $c^*\langle\psi \mid$. we define an inner product of the vector $\mid \psi\rangle$ with a vector $\mid \varphi\rangle$ in the same space, which has the following properties:

1. Positivity : $\langle\psi \mid\psi\rangle \geq 0$, equality applies when $\mid \psi\rangle$ is a null vector.
2. linearity : $\langle\varphi \mid a\psi_1 + b\psi_2\rangle = a\langle\varphi \mid\psi_1\rangle + b\langle\varphi \mid\psi_2\rangle$.
3. The complex conjugate of $\langle\psi \mid\varphi\rangle$ is skew symmetric and is $\langle\psi \mid\varphi\rangle^* = \langle\varphi \mid\psi\rangle$, so that $\langle\psi \mid\psi\rangle$ is a real number. For all non-zero vectors $\mid \psi\rangle$, we can normalise $\langle\psi \mid\psi\rangle = 1$.

4. Schwartz Inequality: $| \langle\psi | \varphi\rangle |^2 \leq \langle\psi | \psi\rangle\langle\varphi | \varphi\rangle$

The dimension of the Hilbert space may be finite or infinite. For most of the time in QIC course we deal with space of finite dimensions. (One of the properties of the Hilbert space is that it is *complete* in norm but as we deal with finite dimensionality, this property is of no consequence to us.)
The set of states $\{| e_n\rangle\}$, $n = 1, 2, 3, \ldots d$, where d is the dimension of the space, is called a set of *basis vectors* if the set spans the linear vector space, i.e., if an arbitrary state $| \psi\rangle$ can be expressed as a linear superposition of the $| e_n\rangle$ s, i.e.,

$$| \psi\rangle = \sum_n \alpha_n | e_n\rangle \tag{1}$$

where the coefficients α_n are, in general, complex. We usually take the basis to be orthonormal, i.e.

$$\langle e_i | e_j\rangle = \delta_{i,j} \tag{2}$$

If the basis is so chosen, we get

$$\alpha_n = \langle e_n | \psi\rangle \tag{3}$$

According to the *Copenhagen Interpretation* of quantum mechanics, $| \alpha_n |^2$ is postulated to be the probability that the state $| \psi\rangle$ will be found in the state $| e_n\rangle$.
Postulate 2:

Physical observables such as position, momentum, energy etc. are represented as linear, self adjoint operators (called Hermitian operators) which have real eigenvalues. The operators map the Hilbert space onto itself, i.e., if $| \psi\rangle$ is a vector in the Hilbert space, the action of the operator $\hat{A}$ on this vector, i.e. $\hat{A} | \psi\rangle$ is another vector in the same space. Likewise, an operator $\hat{B}$ acts to the left on a bra vector, giving another bra vector $\langle\psi | \hat{B}$ on the dual space. The operators have the following properties:

1. Linearity: If α and β are complex scalars, we have

$$\hat{A}(\alpha | \psi\rangle + \beta | \varphi\rangle) = \alpha\hat{A} | \psi\rangle + \beta\hat{A} | \varphi\rangle$$

   (There is an exception to the above for the case of "time reversal" operator, where the coefficients on the right are replaced by their complex conjugates)

2. A unit operator $\hat{I}$ has the property $\hat{I} | \psi\rangle = | \psi\rangle$

3. A product of two operators $\hat{A}$ and $\hat{B}$ is also an operator , i.e., both $\hat{A}\hat{B}$ and $\hat{B}\hat{A}$ are operators. However, in general, $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. The operator product is associative, i.e. $\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C}$. The operator addition is both commutative and associative, i.e., $\hat{A} + \hat{B} = \hat{B} + \hat{A}$ and $\hat{A} + (\hat{B} + \hat{C}) = (\hat{A} + \hat{B}) + \hat{C}$. If there exists an operator $\hat{B}$ such that $\hat{A}\hat{B} = \hat{B}\hat{A} = \hat{I}$, then $\hat{B}$ is called the inverse of $\hat{A}$ and we write $\hat{B} = \hat{A}^{-1}$.

4. Adjoint of an Operator: We define adjoint of an operator $\hat{A}$ and denote it by $\hat{A}^{\dagger}\dagger$ through the relation

$$[\hat{A} \mid \psi\rangle]^* = \langle\psi \mid \hat{A}^{\dagger}$$

i.e. for an arbitrary vectors $\mid \varphi\rangle$ and $\mid \psi\rangle$ we have

$$\langle\varphi \mid \hat{A} \mid \psi\rangle = \langle\varphi \mid \hat{A}\psi\rangle = \langle\hat{A}^{\dagger}\varphi \mid \psi\rangle$$

i.e.

$$\langle\varphi \mid \hat{A} \mid \psi\rangle^* = \langle\psi \mid \hat{A}^{\dagger} \mid \varphi\rangle$$

It follows that if $\hat{A} =\mid \alpha\rangle\langle\beta \mid$ then $\hat{A}^{\dagger} =\mid \beta\rangle\langle\alpha \mid$. An operator is called self adjoint or hermitian if $\hat{A}^{\dagger} = \hat{A}$.
The hermitian conjugate of $\hat{A}\hat{B}$ is $(\hat{A}\hat{B})^{\dagger} = \hat{B}^{\dagger}\hat{A}^{\dagger}$. Note that an operator $\hat{A}\hat{B}$ is not hermitian even though $\hat{A}$ and $\hat{B}$ may be hermitian. $(\hat{A}\hat{B} + \hat{B}\hat{a})$ and $i(\hat{A}\hat{B} - \hat{B}\hat{A})$ are self adjoint.
According to the second postulate of quantum mechanics, a state $\mid \psi\rangle$ has a definite value λ for an observable represented by $\hat{A}$ if and only if $\mid \psi\rangle$ is an eigenstate of $\hat{A}$ with eigenvalue λ,

$$\hat{A} \mid \psi\rangle = \lambda \mid \psi\rangle$$

The eigenstates of $\hat{A}$ form a complete orthonormal basis for the Hilbert space $\mathcal{H}$. If $\{\hat{P}_n\}$ is orthogonal projection of $\hat{A}$ onto the eigen vector basis having eigenvalues λ_n, we can write

$$\hat{A} = \sum_n \lambda_n \hat{P}_n \tag{4}$$

The projection operators satisfy

$$\hat{P}_n\hat{P}_m = \delta_{m,n}\hat{P}_n$$
$$\hat{P}_n^{\dagger} = \hat{P}_n$$

Equation (**??**) is a statement of the **spectral theorem**. Suppose an operator $\hat{A}$ is such that $\hat{A} \mid \psi\rangle =\mid \varphi\rangle$ where $\mid \psi\rangle$ and $\mid \varphi\rangle$ have the same norm. We then have

$$\langle\psi \mid \hat{A}^{\dagger}\hat{A} \mid \psi\rangle = \langle\varphi \mid \varphi\rangle \equiv \langle\psi \mid \psi\rangle$$

so that

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

Such an operator is called a **unitary operator** which has a special place in quantum mechanics in general and in QIC in particular.
Postulate 3:
A quantum state evolves unitarily with time, its time evolution is given by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t} \mid \psi(t)\rangle = \hat{H} \mid \psi(t)\rangle$$

The corresponding equation for $\langle\psi(t\ |)$ is given by

$$-i\hbar\frac{\partial}{\partial t}\langle\psi(t)\ |= \langle\psi(t)\ |\ \hat{H}$$

It is seen that the product $\langle\psi(t)\ |\psi(t)\rangle$ remains unchanged with time and the evolution is unitary, as can be seen below. We define a time evolution operator $\hat{U}(t,t_0)$

$$|\ \psi(t)\rangle = \hat{U}(t,t_0)\ |\ \psi(t_0)\rangle$$

Note that $U(t_0,t_0) = \hat{I}$. We also have

$$|\ \psi(t)\rangle = \hat{U}(t,t_1)\ |\ \psi(t_1)\rangle = \hat{U}(t,t_1)\hat{U}(t_1,t_0)\ |\ \psi(t_0)\rangle$$

which gives

$$\hat{U}(t,t_0) = \hat{U}(t,t_1)\hat{U}(t_1,t_0)$$

Since $\langle\psi(t)\ |\psi(t)\rangle = \langle\psi(t_0)\ |\psi(t_0)\rangle$, we get

$$\langle\psi(t)\ |\psi(t)\rangle = \langle(\ |\ \psi(t_0)\hat{U}^\dagger(t,t_0)U(t,t_0)\ |\ \psi(t_0)\rangle$$

so that

$$\hat{U}^\dagger(t,t_0)\hat{U}(t,t_0) = \hat{I}$$

Consider the time evolution of a state from $t = t_0$ to $t = t_0 + \Delta t$. From Schrödinger equation we have

$$|\ \psi(t+\Delta t)\rangle = \left(1 - \frac{i}{\hbar}\hat{H}\Delta t\right)\ |\ \psi(t)\rangle$$

Since $\hat{H}$ is hermitian, $1 - \dfrac{i}{\hbar}\hat{H}\Delta t$ is unitary to order Δt. We can obtain a finite time evolution in small steps and if $\hat{H}$ has no explicit time dependence we can easily show that

$$\hat{U}(t,t_0) = \exp\left(-i\frac{\hat{H}}{\hbar}(t-t_0)\right)$$

Postulate 4:
The outcome of observation of A is an eigenvalue λ of the operator $\hat{A}$. Immediately after the measurement, the quantum state continues to be the eigenstate corresponding to this eigenvalue. If the state of the system before the measurement was $|\ \psi\rangle$, the measurement outcome will be λ_n with a probability

$$\langle\psi\ |\ P_n\ |\ \psi\rangle = \|\ P_n\ |\ \psi\rangle\ \|^2$$

The state immediately after the measurement is

$$\frac{P_n\ |\ \psi\rangle}{|\ \langle\psi\ |\ P_n\ |\ \psi\rangle\ |^{1/2}}$$

Note that there is an inherent dualism in the way a quantum state evolves with time. On one hand the linearity of the Schrödinger equation implies that the state develops unitarily when it is not being measured while on the other hand the measurement postulate is probabilistic in that it only assigns a probability of possible outcomes.

Postulate 5:
If $\hat{A}$ and $\hat{B}$ are two hermitian operators corresponding respectively to two classical observables a and b, then the commutator of $\hat{A}$ and $\hat{B}$ is given by

$$[\hat{A}, \hat{B}] = i\hbar\hat{C}$$

where $\hat{C}$ is an operator corresponding to a classical variable c which is given by the classical Poisson bracket of the variables a and b , $c = \{a, b\}$ and $\hbar$ is Planck's constant.

3 The Qubit

In both classical and quantum computation, the smallest unit of information is a *bit* which can take one of its two possible values. The quantum bit is usually called a **qubit**. We would distinguish it from a classical bit by calling the latter as a *cbit*. In the practical realisation of quantum computing, we will identify quits with physical states. However, it is convenient to treat them as abstract objects. This enables us to construct a mathematical framework of quantum computation, independent of a specific physical platform. The smallest non-trivial Hilbert space is two dimensional. The unit basis vector in this space are denoted by

$$| 0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad | 1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{5}$$

A general normalised vector in this space is $| \psi\rangle = a \mid 0\rangle + b \mid 1\rangle$, subject to $\mid a \mid^2 + \mid b \mid^2 = 1$. A qubit is *any* state $| \psi\rangle$ in this space. The basis vectors $| 0\rangle$ and $| 1\rangle$ are called the **computational basis states**.

If a measurement of the state $| \psi\rangle$ is made in the computational basis, we obtain the outcome $| 0\rangle$ with a probability $\mid a \mid^2$ and $| 1\rangle$ with a probability $\mid b \mid^2$. Except where either a or b is zero, a measurement disturbs the state which collapses to either $| 0\rangle$ or $| 1\rangle$. A single measurement cannot determine a or b. The way to get information about $| \psi\rangle$ is to prepare a large number of identical copies of the same system and make statistical inference from identical measurement. (We will see later that preparing identical copies is precluded by quantum no-cloning theorem.) Even then this would give us information only about the magnitudes of a and b not about their relative phases. We point out a few physical realisation of qubits:

1. We take $| 0\rangle$ to be a vertically polarised photon which is the state $| \updownarrow\rangle$ and $| 1\rangle$ to be a horizontally polarised photon $| \leftrightarrow\rangle$. An alternative could be to take a photon

polarised at 45° to the horizontal to be the state $|\,0\rangle$ and that polarised at 135° as the state $|\,1\rangle$. Representing the former at $|\,+\rangle$ and the latter as $|\,-\rangle$, we have

$$|\,+\rangle = \frac{1}{\sqrt{2}}(|\,\updownarrow\rangle + |\,\leftrightarrow\rangle)$$
$$|\,-\rangle = \frac{1}{\sqrt{2}}(|\,\updownarrow\rangle - |\,\leftrightarrow\rangle)$$

2. We could take the two states to be spin states of an electron in a magnetic field with respect to an arbitrarily defined z axis. We will discuss this at some length later.

3. We could take $|\,0\rangle$ to correspond to the ground state of an atomic system and $|\,1\rangle$ as the excited state of the same system.

It may be worth mentioning that it is possible to use quantum system with more than two states for information processing. A quantum system with three different basis states (e.g. spin one triplets) is called a **qutrit**. In general, a d dimensional system could be used with information contained in **qudits**.

3.1 Bloch Sphere Representation of Qubits

Let us recall some elementary facts about $S = \frac{1}{2}$ algebra. If $|\,+\rangle$ and $|\,-\rangle$ respectively represent the spin up and spin down states, we can express the spin operators as

$$S_x = \frac{\hbar}{2}\,(|\,+\rangle\langle -\,| + |\,-\rangle\langle +\,|)$$
$$S_y = \frac{i\hbar}{2}\,(|\,-\rangle\langle +\,| - |\,+\rangle\langle -\,|)$$
$$S_x = \frac{\hbar}{2}\,(|\,+\rangle\langle +\,| - |\,-\rangle\langle -\,|)$$

the corresponding matrix representations are

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$
$$S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Note that eigenvalues of $S_n = \vec{S}\cdot\hat{n}$ where $\hat{n}$ is any arbitrary direction is $\pm\frac{\hbar}{2}$. Here $\hat{n}$ is a real unit vector known as the **Bloch vector**. The comp onents of $\hat{n}$ in spherical polar coordinates is $\hat{n}(\theta,\varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, which gives the matrix S_n to be

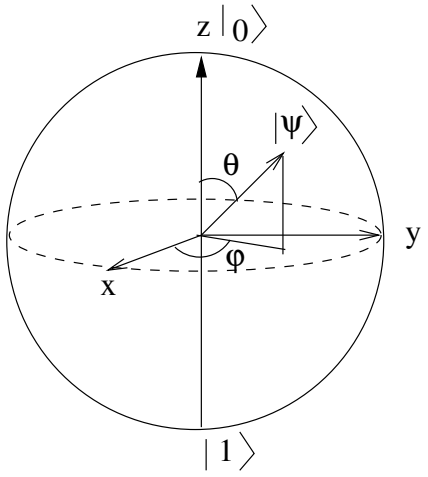
$$S_n = \vec{S}\cdot\hat{n} = \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta\cos\varphi - i\sin\theta\sin\varphi \\ \sin\theta\cos\varphi + i\sin\theta\sin\varphi & -\cos\theta \end{pmatrix}$$

Eigenvalues are seen to be $\frac{\hbar}{2}\lambda$ where

$$(\lambda^2 - \cos^2\theta) - [(\sin\theta\cos\varphi)^2 + (\sin\theta\sin\varphi)^2] = 0$$

which gives $\lambda = \pm 1$. The normalised eigenvector for the eigenvalue $+\frac{\hbar}{2}$ is

$$\mid \psi(\theta,\varphi)\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\varphi}\sin\frac{\theta}{2} \end{pmatrix} = \cos\frac{\theta}{2} \mid 0\rangle + e^{i\varphi}\sin\frac{\theta}{2} \mid 1\rangle$$



geometrically we can represent such an arbitrary state on the surface of a unit sphere called the **Bloch Sphere**. A point on this sphere has the spherical ordinate (θ,φ) and for each value of (θ,φ) it represents a ket $\psi(\theta,\varphi)$. It is easily seen that the north pole of the sphere ($\theta = \varphi = 0$) represents the state $\mid 0\rangle$ and the south pole ($\theta = 0, \varphi = \pi$) the state $\mid 1\rangle$. The point where the positive x axis intersects the sphere ($\theta = \pi/2, \varphi = 0$) is the state $\frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle)$ and the point where the negative x axis ($\theta = \pi/2, \varphi = \pi$) intersects the sphere is the state $\frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1)\rangle$.

How much information is contained in a qubit? In principle, an infinite amount because the binary expansion of θ may contain an infinite number of digits. However, whenever the qubit is observed (in computational basis) it will collapse to either 0 or 1. Suppose a state $\mid +\rangle$ or $\mid -\rangle$ is observed and we get the state $\mid 0\rangle$. After measurement, it would continue to be in this state. If, however, we do not measure the qubit, the system keeps such information. Suppose we measure the z component of the spin. If the original state was $\mid \psi(\theta,\varphi)\rangle$, the expectation value (as measured using a large number of identical copies) is

$$\langle\psi(\theta,\varphi) \mid \sigma_z \mid \psi(\theta,\varphi)\rangle = \begin{pmatrix} \cos\frac{\theta}{2} & e^{-i\varphi}\sin\frac{\theta}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\varphi}\sin\frac{\theta}{2} \end{pmatrix} = \cos\theta$$

However, this will not tell us anything about its component in the xy plane. Determination of $\mid \psi\rangle$ requires determination of three components. We can determine n_3 and n_1 from $\langle\sigma_z\rangle$ and $\langle\sigma_1\rangle$ but the sign of n_2 remains unknown.

3.2 Composite Systems:

Consider a composite system consisting of two sub-systems A and B. Let the Hilbert space of A be $\mathcal{H}_A$ and that of B be $\mathcal{H}_B$. The space of the composite system is $\mathcal{H}_A \otimes \mathcal{H}_B = \mathcal{H}_{AB}$. Suppose $\{\mid \alpha\rangle_A\}$ be a basis in $\mathcal{H}_A$ and $\{\mid \beta\rangle_B\}$ a basis for $\mathcal{H}_B$. We define the basis of $\mathcal{H}_{AB}$ to be the composite set $\mid \alpha, \beta\rangle_{AB}$. The orthonormality relationship for the basis is

$$_{AB}\langle\alpha', \beta' \mid \alpha, \beta\rangle_{AB} = \delta_{\alpha,\alpha'}\delta_{\beta,\beta'}$$

We define an operator in this composite space as $M_A \otimes N_B$ which acting on a state of the composite system

$$M_A \otimes N_B \mid \psi, \varphi\rangle_{AB} = M_A \mid \psi\rangle_A \otimes N_B \mid \varphi\rangle = \sum (M_A)_{\psi,\alpha} \mid \alpha\rangle (N_B)_{\varphi,\beta} \mid \beta\rangle$$

Let us look at two qubit system. The corresponding cbits are 00, 01, 10 and 11. The quantum states for two qubits is

$$\mid \psi\rangle = \alpha_{00} \mid 00\rangle + \alpha_{01} \mid 01\rangle + \alpha_{10} \mid 10\rangle + \alpha_{11} \mid 11\rangle$$

subject to $\mid \alpha_{00} \mid^2 + \mid \alpha_{01} \mid^2 + \mid \alpha_{10} \mid^2 \mid \alpha_{11} \mid^2 = 1$.

Suppose we measure the first qubit and get 0 (with probability $\mid \alpha_{00} \mid^2 + \mid \alpha_{01} \mid^2$). The post measurement state is

$$\mid \psi'\rangle = \frac{\alpha_{00} \mid 00\rangle + \alpha_{01} \mid 01\rangle}{\sqrt{\mid \alpha_{00} \mid^2 + \mid \alpha_{01} \mid^2}}$$

Note that not all two qubit states can be written as a product of single qubit states. One such state is

$$\mid \beta_{00}\rangle = \frac{\mid 00\rangle + \mid 11\rangle}{\sqrt{2}}$$

This is a member of a set known as **Bell states**, about which we will discuss at length letter. Suppose we measure the first qubit of the above state. If we get 0, the post measurement state is $\mid 00\rangle$ (which is obtained with a probability of 0.5). The result of the first qubit determines the second qubit as well. This shows that the mesurement outcomes are correlated.

4 Some results in Linear Algebra

In this section we will some important results in linear algebra.

Projection Operator:

We had defined a projection operator along the basis vector $\mid e_k\rangle$ to be given by

$$\mid e_k\rangle\langle e_k \mid$$

This operator, acting on an arbitrary state $\mid \psi\rangle$ projects the component along the ray $\mid e_k\rangle$ so that $\mid \psi\rangle - P_k \mid \psi\rangle$ is orthogonal to $\mid \psi\rangle$. Projection operator has the following properties:
(i) Idempotency : $P_k^2 = P_k$
(ii) Orthogonality : $P_k P_j = 0$ if $j \neq k$.
(iii) Completeness : $\sum_k P_k = I$

Example:

Let $\mid e_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ and $\mid e_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$. Express $\mid \psi\rangle = \begin{pmatrix}3\\5\end{pmatrix}$ in terms of the projection operators P_1 and P_2.
Solution:

$$P_1 = \mid e_1\rangle\langle e_1 \mid = \frac{1}{2}\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}$$

$$P_2 = \mid e_2\rangle\langle e_2 \mid = \frac{1}{2}\begin{pmatrix}1 & -1\end{pmatrix}\begin{pmatrix}1 & -1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix}$$

Consider $\mid \psi\rangle = \sum_k c_k \mid e_k\rangle = \begin{pmatrix}3\\2\end{pmatrix}$. We have,

$$P_1\begin{pmatrix}3\\2\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}3\\2\end{pmatrix} = \frac{5}{2}\begin{pmatrix}1\\1\end{pmatrix}$$

$$P_2\begin{pmatrix}3\\2\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix}\begin{pmatrix}3\\2\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}$$

Thus $\mid \psi\rangle = \frac{5}{2} \mid e_1\rangle + \frac{1}{2} \mid e_2\rangle$.

Spectral Theorem:
Spectral theorem is applicable for all normal matrices, i.e. those matrices which satisfy $AA^\dagger = A^\dagger A$. Clearly, hermitian matrices being normal, the theorem is applicable to them as well. In general, the eigenvalues of normal matrices need not be real and therefore, the converse is not necessarily true, i.e. all normal matrices are not hermitian. If a normal matrix has real eigenvalues, it is hermitian. Further, it can also be shown that for a normal matrix, the eigenvectors belonging to distinct eigenvalues are orthogonal. A normal matrix can be written in terms of its spectral decomposition

$$\hat{A} = \sum_i \lambda_i \mid \lambda_i\rangle\langle\lambda_i \mid$$

where $\{\mid \lambda_i\rangle\}$ are normalised eigenvectors of A labelled by the corresponding eigenvalues. This follows from the completeness relation $I = \sum_i \mid \lambda_i\rangle\langle\lambda_i \mid$. It is seen that

$$A = AI = \sum_i A \mid \lambda_i\rangle\langle\lambda_i \mid = \sum_i \lambda_i \mid \lambda_i\rangle\langle\lambda_i \mid$$

Note that $\mid \lambda_i\rangle\langle\lambda_i \mid$ is the projection operator in the direction of the vector $\mid \lambda_i\rangle$. Thus the effect of $\hat{A}$ in the one dimensional space spanned by $\{\lambda_i\}$ is equivalent to multipliertion by a scalar λ_i.

We will get an expression for the projection operators in the eigenspace of $\hat{A}$. Let $\hat{A}$ be a normal matrix and let $\{\lambda_\alpha\}$ and $\{\mid \lambda_{\alpha,p}\rangle;\ (1 \leq p \leq g_\alpha)$ be the eigenvalues and the corresponding eigenvectors which have a degeneracy g_α. If the space is n dimensional, we have, $\sum_\alpha 1 = n;\ \ \sum_\alpha g_\alpha = n$. We have the following expression for the projection operator for the space of λ_α

$$P_\alpha = \frac{\prod_{\beta\neq\alpha}(\hat{A} - \lambda_\beta \hat{I})}{\prod_{\gamma\neq\alpha}(\lambda_\alpha - \lambda_\gamma)} \tag{6}$$

The proof is straightforward. If P_α acts on $\mid \lambda_{\alpha,p}\rangle$, we get

$$\begin{aligned} P_\alpha \mid \lambda_{\alpha,p}\rangle &= \frac{\prod_{\beta\neq\alpha}(\hat{A} - \lambda_\beta \hat{I}) \mid \lambda_{\alpha,p}\rangle}{\prod_{\gamma\neq\alpha}(\lambda_\alpha - \lambda_\gamma)} \\ &= \frac{\prod_{\beta\neq\alpha}(\lambda_\alpha - \lambda_\beta)}{\prod_{\gamma\neq\alpha}(\lambda_\alpha - \lambda_\gamma)} \mid \lambda_{\alpha,p}\rangle = \mid \lambda_{\alpha,p}\rangle \end{aligned}$$

On the other hand, if it acts on a state belonging to a different eigenvalue ($\delta \neq \alpha$)

$$\begin{aligned} P_\alpha \mid \lambda_{\delta,q}\rangle &= \frac{\prod_{\beta\neq\alpha}(\hat{A} - \lambda_\beta \hat{I}) \mid \lambda_{\delta,q}\rangle}{\prod_{\gamma\neq\alpha}(\lambda_\alpha - \lambda_\gamma)} \\ &= \frac{\prod_{\beta\neq\alpha}(\lambda_\delta - \lambda_\beta)}{\prod_{\gamma\neq\alpha}(\lambda_\alpha - \lambda_\gamma)} \mid \lambda_{\delta,q}\rangle = 0 \end{aligned}$$

the vanishing of the numerator follows because one of the terms is $\beta = \delta$. Thus $P_\alpha = \sum_{p=1}^{g_\alpha} \mid \lambda_{\alpha,p}\rangle\langle\lambda_{\alpha,p} \mid$ is the projection operator onto the eigenspace of λ_α.

Example:

Consider σ_y again. We know the eigenvalues to be ± 1. Thus the projection operators can be written, using spectral theorem, as

$$P_1 = \frac{\sigma_y - (-I)}{1 - (-1)} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$$

$$P_2 = \frac{\sigma_y - (I)}{-1 - (+1)} = \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$$

It can be seen that $\sigma_y = 1 \times P_1 + (-1) \times P_2$.

Spectral decomposition is useful in evaluating function of a matrix. It may be observed

that functions of a matrix are defined by a power series expansion of the function. Since $A = \sum_\alpha \lambda_\alpha P_\alpha$, we have

$$A^n = \left(\sum_\alpha \lambda_\alpha P_\alpha\right)^n = \sum_\alpha \lambda_\alpha^n P_\alpha$$

the last relation follows by expanding the sum as a binomial series and observing that $P_\alpha P_\beta = P_\alpha \delta_{\alpha,\beta}$. Similarly, if A^{-1} exists then λ_α^{-1} is an eigenvalue of A^{-1}. This can be seen as follows.

$$\mid \lambda_{\alpha,p}\rangle = A^{-1}A \mid \lambda_{\alpha,p}\rangle = \lambda_\alpha A^{-1} \mid \lambda_{\alpha,p}\rangle$$

Thus

$$A^{-1} \mid \lambda_{\alpha,p}\rangle = \lambda_\alpha^{-1} \mid \lambda_{\alpha,p}\rangle$$

Consider σ_y again. We have,

$$\begin{aligned} e^{i\alpha\sigma_y} &= \sum_{k=0}^{\infty} \frac{(i\alpha\sigma_y)^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{(i\alpha)^k}{k!}(P_1 + (-1)P_2)^k \\ &= \sum_{k=0}^{\infty} \frac{(i\alpha)^k}{k!}(P_1 + (-1)P_2)^k \\ &= e^{i\alpha}P_1 + e^{-i\alpha}P_2 \\ &= \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \end{aligned}$$

Positive Operator:

An operator is said to be positive semi-definite if $\langle\psi \mid \hat{A}\mid \psi\rangle \geq 0$ for each $\mid \psi\rangle \in \mathcal{H}$ It is possible to show that a positive operator is hermitian which only has non-negative eigenvalues.

Polar Decomposition of an operator:

If the matrix representation of an operator $\hat{A}$ is non-singular, it becomes possible to express $\hat{A}$ as a product of a unitary operator U and a positive operator. There exist two such positive operators J and K, the former provides a left decomposition and the latter a right decomposition.

$$A = UJ = KU$$

where $J = \sqrt{A^\dagger A}$ and $K = \sqrt{AA^\dagger}$. Since $\sqrt{A^\dagger A}$ is a positive operator (and hence hermitian), it has a spectral decomposition

$$J = \sum_i \lambda_i \mid i\rangle\langle i \mid$$

with $\lambda_i \geq 0$. Let $\mid \psi\rangle = A \mid i\rangle$ so that $\langle\psi_i \mid \psi_i\rangle = \lambda_i^2$. Consider those λ_i for which $\lambda_i > 0$. Let $\mid e_i\rangle = \dfrac{\mid \psi\rangle}{\lambda_i} = \dfrac{A \mid i\rangle}{\lambda_i}$ so that $\mid e_i\rangle$ are normalised. They are also orthogonal because,

for $i \neq j$

$$\begin{aligned}\langle e_i \,|e_j\rangle &= \frac{1}{\lambda_i\lambda_j}\langle i \,|A^\dagger A|\, j\rangle \\ &= \frac{1}{\lambda_i\lambda_j}\langle i \,|J^2|\, j\rangle = 0\end{aligned}$$

The last relation follows because $|\, j\rangle$ is an eigenstate of J^2. Now we use Gram-Schmidt procedure to extend $\{|\, e_i\rangle\}$ so that it forms an orthonormal basis. Let $U = \sum_i \,|\, e_i\rangle i$. When $\lambda_i \neq 0$, we have

$$\begin{aligned}UJ \,|\, i\rangle &= \lambda_i U \,|\, i\rangle \\ &= \lambda_i \,|\, e_i\rangle \\ &=|\, \psi_i\rangle = A \,|\, i\rangle\end{aligned}$$

Thus $A = UJ,\ A^\dagger = JU^\dagger$ which gives $A^\dagger A = JU^\dagger UJ = J^2$. If A is invertible, J is also given by $U = AJ^{-1}$ which determines U uniquely. The process is called polar deposition because $\det A = \det U \det J$ is of the form $re^{i\theta}$.

Example:

Find the polar decomposition of

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

We have, $A^\dagger A = \begin{pmatrix} r^2 & 0 \\ 0 & r^2 \end{pmatrix}$, where $r = \sqrt{a^2+b^2}$. Thus $J = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}$. If $A = UJ$, we get

$$U = AJ^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

where $\cos\theta = \dfrac{a}{r}$ and $\sin\theta = \dfrac{b}{r}$.

Singular Value Decomposition:

Suppose A is an $m \times n$ matrix with complex entries. It is possible to decompose A as

$$A = U\Sigma V^\dagger$$

where U and V are unitary matrices i.e. $U^\dagger U = V^\dagger V = I$. U is a $m \times m$ square matrix, V is an $n \times n$ square matrix and Σ is an $m \times n$ matrix whose diagonal entries are positive semidefinite. Σ is known as singular value matrix. If A is a real matrix, U and V are orthogonal matrices.

The proof follows from the polar decomposition. Let $A = SJ$ where S is unitary . since J is positive, it is hermitian and is diagonalisable by a unitary matrix T. Let $J = T\Sigma T^\dagger$ where Σ is real and non-negative. Thus, $A = ST\Sigma T^\dagger$. Since S and T are unitary, we define $ST = U$ and $T = V$ from which we get

$$A = U\Sigma V^\dagger$$

We will illustrate it with an example. Using SVD consists of finding eigenvalues and eigenvectors of $A^\dagger A$. The eigenvectors of $A^\dagger A$ are the columns of V while those of $AA^\dagger$ are the columns of U. The diagonal entries of S are the square roots of the eigenvalues of $AA^\dagger$ or $A^\dagger A$, arranged in descending order.

Example: Obtain singular value decomposition of the matrix $A = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix}$.

Solution:
We have,

$$A^\dagger = \begin{pmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{pmatrix}$$

$$AA^T = \begin{pmatrix} 11 & 1 \\ 1 & 11 \end{pmatrix}$$

The eigenvalues of AA^T are 12 and 10. The normalised eigenvectors for $\lambda = 12$ are $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and that for $\lambda = 10$ is $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Thus the matrix V is given by

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

$$A^T A = \begin{pmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{pmatrix}$$

The eigenvalues of $A^T A$ are 12 ,10 and 0. The normalised eigenvectors for $\lambda = 12$ are $\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, that for $\lambda = 10$ is $\frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ and that for $\lambda = 0$ is $\frac{1}{\sqrt{30}}\begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix}$. Thus the matrix U is given by

$$U = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & -\frac{5}{\sqrt{30}} \end{pmatrix}$$

The singular value matrix Σ is given by

$$\Sigma = \begin{pmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \end{pmatrix}$$

where we have added a column with null entries to get the right dimension. The validity of SVD can be checked by direct multiplication.