Document II
MIXED STATES*

Pramod S. Joag

November 18, 2018

Contents

1   Preamble

In the last section of the first write up we saw that, in an overwhelmingly large number of situations, we may have only a partial information on the preparation of a state of a quantum system. This partial information can at best be used to specify a list of pure states $\{|\psi_k\rangle, k = 1, \ldots, N\}$ with the corresponding probabilities $p_k, k = 1, \ldots, N$, where p_k is the probability that the system is in the pure state $|\psi_k\rangle$ at some initial time t when the system is prepared. The system is in exactly one of the pure states $\{|\psi_k\rangle, k = 1, \ldots, N\}$ but we lack the knowledge of exactly which state out of $\{|\psi_k\rangle, k = 1, \ldots, N\}$ the system is in. The partial information on the preparation can only give the probability of the system being in the corresponding pure state. In other words, the state of the system is specified as a proper or classical mixture of pure states defined by the ensemble [1] $\{|\psi_k\rangle, p_k; k = 1, \ldots, N\}$. A given ensemble of pure states, that is, a probability distribution over a list of pure states, corresponds to a physical process, either realized naturally or in a laboratory, which we have called above the preparation or the preparation procedure of the classical mixture. The state specified in this way, as a classical mixture of pure states, is called a mixed state. The preparation leading to a given mixed state is not unique. As we shall see, there are infinitely many possible preparations leading to the same mixed state.

The question is : how can we incorporate mixed states in the quantum formalism? We cannot use a linear combination of pure states forming the ensemble, that is, $\sum_k p_k |\psi_k\rangle$, because such a linear combination amounts to the superposition of states, giving a new pure state in the state space. Thus we require a new theoretical structure to deal with mixed states. We proceed to build this structure.

2   Prerequisites

We start by learning some basic theoretical structures required for incorporating mixed states into quantum formalism.

2.1   Projection operators

The first piece we require is the so called projection operator, or a projector for short. Consider a n dimensional Hilbert space or the state space of a quantum system. An arbitrary normalized state $|\psi\rangle$ can be expanded as the superposition of states of an orthonormal basis $\{|k\rangle; k = 1, \ldots, n\}$ as

$$|\psi\rangle = \sum_{k=1}^{n} |k\rangle (\langle k|\psi\rangle). \tag{1}$$

---

[1] By ensemble we mean a large collection of identical non-interacting systems with a fraction p_k in the pure state $|\psi_k\rangle$. I emphasize that here we are not concerned with the state of the ensemble as a whole, that is, the state of all the systems making the ensemble put together.

This expansion is analogous to that of a unit vector $\hat{\mathbf{k}}$ in an n dimensional space in terms of an orthogonal basis of unit vectors $\hat{\mathbf{i}} = \hat{\mathbf{1}}, \hat{\mathbf{2}}, \ldots, \hat{\mathbf{n}}$, as

$$\hat{\mathbf{k}} = \sum_{\hat{\mathbf{i}}} (\hat{\mathbf{k}} \cdot \hat{\mathbf{i}}) \hat{\mathbf{i}}. \tag{2}$$

with $(\hat{\mathbf{k}} \cdot \hat{\mathbf{i}})$ as the projection of $\hat{\mathbf{k}}$ along the basis vector $\hat{\mathbf{i}}$. Thus in Eq.(1), the scalar product $\langle k|\psi \rangle$ can be viewed as the result of projecting the state vector $|\psi\rangle$ along the basis vector $|k\rangle$. What we need, then, is the operator which projects a state $|\psi\rangle$ along a state say $|\alpha\rangle$ in the sense of Eq.(1) and Eq.(2). This is the operator $|\alpha\rangle\langle\alpha|$ defined by its action on the state via

$$(|\alpha\rangle\langle\alpha|)|\psi\rangle = |\alpha\rangle(\langle\alpha|\psi\rangle). \tag{3}$$

We now establish some important properties of the projection operator.

- Linearity : We have,

$$(|\alpha\rangle\langle\alpha|)(a|\psi\rangle + b|\phi\rangle) = a|\alpha\rangle(\langle\alpha|\psi\rangle) + b|\alpha\rangle(\langle\alpha|\phi\rangle), \tag{4}$$

  and

$$(a|\alpha\rangle\langle\alpha| + b|\beta\rangle\langle\beta|)|\psi\rangle = a|\alpha\rangle(\langle\alpha|\psi\rangle) + b|\beta\rangle(\langle\beta|\psi\rangle), \tag{5}$$

  where a and b are complex numbers.

- Idempotency : Consider

$$
\begin{aligned}
(|\alpha\rangle\langle\alpha|)^2|\psi\rangle &= (|\alpha\rangle\langle\alpha|)(|\alpha\rangle\langle\alpha|)|\psi\rangle \\
&= (|\alpha\rangle\langle\alpha|)[|\alpha\rangle(\langle\alpha|\psi\rangle)] \\
&= |\alpha\rangle(\langle\alpha|\psi\rangle)
\end{aligned}
$$

  since $|\alpha\rangle$ is normalized so that $\langle\alpha|\alpha\rangle = 1$. This leads to the operator identity

$$(|\alpha\rangle\langle\alpha|)^2 = |\alpha\rangle\langle\alpha|. \tag{6}$$

  This result is analogous to what we already know from elementary geometry : Projecting a vector twice along a direction is equivalent to projecting it only once.

- Orthogonal projectors : Consider two orthogonal states $|\alpha\rangle, |\beta\rangle, \quad \langle\alpha|\beta\rangle = 0$. The corresponding projectors form a pair of orthogonal projection operators. Successive application of a pair of orthogonal projectors on any state nullifies that state.

$$(|\alpha\rangle\langle\alpha|)(|\beta\rangle\langle\beta|)|\psi\rangle = (|\alpha\rangle\langle\alpha|)[|\beta\rangle(\langle\beta|\psi\rangle)] = 0|\psi\rangle = 0,$$

  since $\langle\alpha|\beta\rangle = 0$. Thus if Q and P are a pair of orthogonal projectors their product (in any order) is a zero operator.

$$QP = 0. \tag{7}$$

We check that the operators $|\alpha\rangle\langle\alpha|$ and $I - |\alpha\rangle\langle\alpha|$ are orthogonal projectors. We have the operator equation,

$$(|\alpha\rangle\langle\alpha|)(I - |\alpha\rangle\langle\alpha|) = |\alpha\rangle\langle\alpha| - (|\alpha\rangle\langle\alpha|)^2 = |\alpha\rangle\langle\alpha| - |\alpha\rangle\langle\alpha| = 0.$$

For any given state $|\psi\rangle$, the states $|\alpha\rangle\langle\alpha|\,|\psi\rangle$ and $(I - |\alpha\rangle\langle\alpha|)\,|\psi\rangle$ are orthogonal. The corresponding scalar product reduces to

$$\{\langle\psi|\alpha\rangle\langle\alpha|\}\{|\psi\rangle - |\alpha\rangle(\langle\alpha|\psi\rangle)\} = \langle\psi|\alpha\rangle\langle\alpha|\psi\rangle - \langle\alpha|\psi\rangle\langle\psi|\alpha\rangle = 0.$$

Thus if we collect the states $(I - |\alpha\rangle\langle\alpha|)|\psi\rangle$ for all $|\psi\rangle \in \mathcal{H}$, we get a subspace of $\mathcal{H}$ containing the states orthogonal to the state $(|\alpha\rangle\langle\alpha|)|\psi\rangle$ or $|\alpha\rangle(\langle\alpha|\psi\rangle)$ or just $|\alpha\rangle$. Of course, this result is true for any pair of orthogonal projectors.

Resolution of Identity :
We can write Eq.(1), with an orthonormal basis $\{|k\rangle, k = 1, \ldots, n\}$ as

$$|\psi\rangle = \sum_k |k\rangle(\langle k|\psi\rangle) = \sum_k (|k\rangle\langle k|)|\psi\rangle = \left[\sum_k (|k\rangle\langle k|)\right]|\psi\rangle,$$

which immediately gives

$$\sum_k (|k\rangle\langle k|) = I \tag{8}$$

That is, the sum of the projection operators over all the states in an orthonormal basis equals identity operator. This result, called resolution of identity, is very useful and will be used frequently in the sequel.

2.2 Trace of a linear operator

The second piece we need is the trace of a linear operator. The trace of a linear operator A is defined by

$$tr(A) = \sum_i \langle\psi_i|A|\psi_i\rangle \tag{9}$$

where $\{|\psi_i\rangle, i = 1, \ldots, n\}$ is an arbitrary basis. The way we have defined trace implies that it is independent of the basis used to compute it and depends only on the concerned operator.

As an example we find out the trace of the operator $|\beta\rangle\langle\alpha|$ defined by

$$(|\beta\rangle\langle\alpha|)|\psi\rangle = |\beta\rangle(\langle\alpha|\psi\rangle).$$

We have,

$$tr(|\beta\rangle\langle\alpha|) = \sum_k \langle k|(|\beta\rangle\langle\alpha|)|k\rangle = \langle\alpha|\left[\sum_k |k\rangle\langle k|\right]|\beta\rangle = \langle\alpha|\beta\rangle, \tag{10}$$

where we have used Eq.(8). Obviously, the value of the trace is completely independent of the basis used to calculate it.

Basis independence of trace operation can be easily proved for orthonormal bases. We have, using orthonormal bases,

$$
\begin{aligned}
tr(A) &= \sum_k \langle \psi_k | A | \psi_k \rangle = \sum_{k,l,m} \langle \psi_k | \phi_l \rangle \langle \phi_l | A | \phi_m \rangle \langle \phi_m | \psi_k \rangle \\
&= \sum_{l,m} \langle \phi_l | A | \phi_m \rangle \langle \phi_m | \left[\sum_k | \psi_k \rangle \langle \psi_k | \right] | \phi_l \rangle = \sum_l \langle \phi_l | A | \phi_l \rangle.
\end{aligned}
\tag{11}
$$

where we have used Eq.(8) and the orthonormality condition $\langle \phi_m | \phi_l \rangle = \delta_{ml}$.

We now describe various properties of the trace operation. In the following, a, b are two complex numbers, A, B are two linear operators on the Hilbert space of states and $\{|k\rangle; k = 1, \ldots, n\}$ is an orthonormal basis.

- As can be seen from its definition, trace is a linear operation. That is,

$$
tr(aA + bB) = a\,tr(A) + b\,tr(B).
$$

-

$$
tr(AB) = tr(BA).
\tag{12}
$$

We have,

$$
\begin{aligned}
tr(AB) &= \sum_k \langle k | AB | k \rangle \\
&= \sum_k \langle k | A \left[\sum_j | j \rangle \langle j | \right] B | k \rangle \\
&= \sum_j \langle j | B \left[\sum_k | k \rangle \langle k | \right] A | j \rangle \\
&= \sum_j \langle j | BA | j \rangle = tr(BA),
\end{aligned}
$$

where we have used Eq.(8).

- Adjoint of an operator A is defined via

$$
\langle y | A | x \rangle = \langle x | A^\dagger | y \rangle^*,
$$

where $|x\rangle, |y\rangle$ are arbitrary (normalized) kets in the state space. This means, by a defining property of the scalar product, that the bra of the ket $A|x\rangle$ is $\langle x|A^\dagger$. [2] We have,

$$tr(A^\dagger) = \sum_k \langle k|A^\dagger|k\rangle = \sum_k \langle k|A|k\rangle^* = [tr(A)]^*. \tag{13}$$

Thus the trace of $A^\dagger$ is the complex conjugate of the trace of A.

2.3   Hilbert space of operators

We show that the set $L_\mathcal{H}$ of linear operators on a complex Hilbert space of states $\mathcal{H}$ forms a complex linear space.[3] Addition of two operators L_1 and L_2 is defined by

$$(L_1 + L_2)|x\rangle = L_1|x\rangle + L_2|x\rangle \quad \forall |x\rangle \in \mathcal{H},$$

scalar (complex number) multiplication by

$$(\alpha L)|x\rangle = \alpha(L|x\rangle) \quad \forall |x\rangle \in \mathcal{H}.$$

A linear combination

$$(\alpha L_1 + \beta L_2)|x\rangle = \alpha(L_1|x\rangle) + \beta(L_2|x\rangle)$$

is also a linear operator. The properties of addition defining a linear space are
   Commutativity :

$$(L_1 + L_2)|x\rangle = (L_2 + L_1)|x\rangle \quad \forall |x\rangle \in \mathcal{H}.$$

   Associativity :

$$[(L_1 + L_2) + L_3]|x\rangle = [L_1 + (L_2 + L_3)]|x\rangle \quad \forall |x\rangle \in \mathcal{H},$$

which are obviously satisfied. For a linear operator L its additive inverse is $-L = (-1)L$:

$$L - L = \hat{0},$$

where the zero operator $\hat{0}$ nullifies all states

$$\hat{0}|x\rangle = 0|x\rangle = 0.$$

---

[2]This makes the adjoint operator $A^\dagger$ act on the dual of Hilbert space, that is, the space comprising all functionals on the Hilbert space. However we can show that every functional on the Hilbert space can be realized via some bra, as described in write-up-1, and the Hilbert space is isomorphic with its dual. These facts are expressed by saying that Hilbert space is its own dual. Thus the adjoint of an operator on the Hilbert space can be treated as the operator acting on the Hilbert space. We shall not prove these statements.
[3]The phrases 'linear space' and 'vector space' have the same meaning. We use it interchangeably in this note.

Thus
$$(L + \hat{0})|x\rangle = L|x\rangle + \hat{0}|x\rangle = L|x\rangle \quad \forall |x\rangle \in \mathcal{H},$$
showing that the zero operator is the additive identity. Thus we have shown that $L_\mathcal{H}$ is a linear space, or a vector space.

An additional result is that the vector space $L_\mathcal{H}$ can be given a natural inner product, turning it into a Hilbert space. This inner or scalar product is defined as the map $L_\mathcal{H} \times L_\mathcal{H} \mapsto \mathbb{C}$ ($\mathbb{C}$: the set of complex numbers,) namely,

$$(A, B) = tr(A^\dagger B), \tag{14}$$

with $A, B \in L_\mathcal{H}$. We have to show that Eq.(14) defines a valid scalar product. We have,

$$(A, \alpha B + \beta C) = tr(A^\dagger(\alpha B + \beta C)) = \alpha tr(A^\dagger B) + \beta tr(A^\dagger C) = \alpha(A, B) + \beta(A, C),$$

$$(A, B) = tr(A^\dagger B) = tr(B^\dagger A)^\dagger = (tr(B^\dagger A))^* = (B, A)^*,$$

$$(A, A) = tr(A^\dagger A) = \sum_k \langle k|A^\dagger A|k\rangle = \sum_k ||A|k\rangle||^2 \geq 0.$$

These three equations prove that Eq.(14) defines a valid scalar product.

As we shall show later, $L_\mathcal{H}$ is isomorphic with the space of $n \times n$ complex matrices, which in turn is isomorphic with the space of n^2-tuples of complex numbers $\mathbb{C}^{n^2}$. Since the dimension of $\mathbb{C}^{n^2}$ is n^2, that of $L_\mathcal{H}$ is also n^2. A more direct way to see this isomorphism is the following. Since $L_\mathcal{H}$ is a linear space, we can choose a basis in it and express every operator in $L_\mathcal{H}$ as a unique linear combination of the basis elements. This establishes a one to one correspondence between the n^2-tuples of complex numbers which constitute the linear space $\mathbb{C}^{n^2}$ and the operators in $L_\mathcal{H}$: the n^2-tuple corresponding to an operator is the one which occurs in the corresponding linear combination of the basis. This correspondence is also onto, because an n^2-tuple gives a linear combination of the basis, defining an operator which is the pre-image of the n^2-tuple. Further, the image of a linear combination of operators under this correspondence is the same linear combination of the images of operators in it. Thus this correspondence is an isomorphism between $L_\mathcal{H}$ and $\mathbb{C}^{n^2}$.

We now inquire about the all important subset of $L_\mathcal{H}$, namely, the one comprising all the hermitian operators. If A and B are two hermitian operators, then the adjoint of their linear combination $aA + bB$ is $a^*A + b^*B$, so that such a linear combination is an hermitian operator only when the coefficients a, b are real. Thus the subset of all hermitian operators is closed under their linear combinations with real coefficients and forms a real subspace of the complex Hilbert space $L_\mathcal{H}$. The scalar product on this subspace is $(A, B) = tr(AB)$ so that this subspace is a Hilbert space in its own right. To get its dimension, we make use of the fact that this subspace is isomorphic with the space of $n \times n$ (complex) hermitian

matrices. Counting all the hermiticity conditions, it is straightforward to show that we need to fix n^2 real numbers to fix a hermitian matrix. Thus the space of hermitian matrices is isomorphic to the space of n^2 tuples of real numbers, $\mathbb{R}^{n^2}$. Since the dimension of $\mathbb{R}^{n^2}$ is n^2, that of the subspace of hermitian operators is also n^2. Thus the hermitian operators form a n^2 dimensional real subspace of $L_{\mathcal{H}}$. The isomorphism between the subspace of Hermitian operators and $\mathbb{R}^{n^2}$ can be established in exactly the same way as we established the isomorphism between $L_{\mathcal{H}}$ and $\mathbb{C}^{n^2}$ in the above para.

As an example, we show that the $4-D$ Hilbert space $L_{\mathcal{H}}$ corresponding to $2-D$ state space of a 2 state quantum system is spanned by the orthonormal basis

$$\left(\frac{1}{\sqrt{2}}I, \frac{1}{\sqrt{2}}\boldsymbol{\sigma}_1, \frac{1}{\sqrt{2}}\boldsymbol{\sigma}_2, \frac{1}{\sqrt{2}}\boldsymbol{\sigma}_3 \right) \tag{15}$$

where $\boldsymbol{\sigma}_{1,2,3}$ are Pauli operators satisfying

$$\begin{aligned}
\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j &= i\varepsilon_{ijk}\boldsymbol{\sigma}_k, \\
tr(\boldsymbol{\sigma}_i) &= 0 \quad i=1,2,3, \\
tr(I) &= 2, \\
\boldsymbol{\sigma}_i^2 &= I \quad i=1,2,3, \\
\boldsymbol{\sigma}_i^\dagger &= \boldsymbol{\sigma}_i \quad i=1,2,3.
\end{aligned} \tag{16}$$

Here ε_{ijk} are Levi-Civita symbols which equal $+1$ when ijk are a cyclic permutation of 123. $\varepsilon_{ijk} = -1$ if ijk are a permutation of 123 which is not cyclic. $\varepsilon_{ijk} = 0$ when two or more of i,j,k are equal. To show that the set (15) is an orthonormal basis, we evaluate their mutual scalar products. We have,

$$\begin{aligned}
(\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_j) &= tr(\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j) = i\varepsilon_{ijk}tr(\boldsymbol{\sigma}_k) = 0 \quad i \neq j, \;\; i,j=1,2,3 \\
tr(I, \boldsymbol{\sigma}_i) &= tr(\boldsymbol{\sigma}_i) = 0 \quad i=1,2,3
\end{aligned}$$

Thus the operators in the set (15) are mutually orthogonal and hence linearly independent. Further, they are normalized.

$$\begin{aligned}
(\frac{1}{\sqrt{2}}\boldsymbol{\sigma}_i, \frac{1}{\sqrt{2}}\boldsymbol{\sigma}_i) &= \frac{1}{2}tr(\boldsymbol{\sigma}_i^2) = \frac{1}{2}tr(I) = 1 \quad i=1,2,3 \\
(\frac{1}{\sqrt{2}}I, \frac{1}{\sqrt{2}}I) &= \frac{1}{2}tr(I) = 1
\end{aligned}$$

Thus the set (15) is an orthonormal basis of $L_{\mathcal{H}}$ for a two state system. It is also an orthonormal basis of the subspace of hermitian operators if its linear combinations are restricted to real coefficients.

Note that if we re-define the scalar product of operators by

$$(A, B) = \frac{1}{2}tr(A^\dagger B), \tag{17}$$

then the set $\{I, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3\}$ forms an orthonormal basis.

We now introduce some classes of operators and make some statements involving them, without proof, for future use.

An operator A is called *normal* if it commutes with its adjoint

$$[A, A^\dagger] = AA^\dagger - A^\dagger A = 0 \tag{18}$$

A Hermitian or the self adjoint operator $(A = A^\dagger)$ is obviously normal, as well as a unitary operator U satisfying $UU^\dagger = I = U^\dagger U$.

An operator can be diagonalized by a unitary transformation (unitary change of basis) if and only if it is normal. In other words, an operator has an orthonormal eigenbasis if and only if it is normal (see Appendix B).

An operator is Hermitian if and only if it has real eigenvalues. The only if part, that is, a Hermitian operator has real eigenvalues is proved in all the textbooks on quantum mechanics. We prove the if part, that is, an operator having real eigenvalues is Hermitian. We first note that any operator P can be written as

$$\begin{aligned} P &= X + iY \\ X &= \frac{P + P^\dagger}{2} \\ Y &= \frac{P - P^\dagger}{2i} \end{aligned} \tag{19}$$

It is straightforward to check that both X and Y are Hermitian operators. From the equation

$$P|\rangle = (X + iY)|\rangle = (\lambda_1 + i\lambda_2)|\rangle$$

and the fact that X and Y are hermitian operators, it follows that

$$X|\rangle = \lambda_1|\rangle, \quad \text{and} \quad Y|\rangle = \lambda_2|\rangle.$$

Thus an eigenstate of P with eigenvalue $\lambda_1 + i\lambda_2$ is also an eigenstate of X with eigenvalue λ_1 and an eigenstate of Y with eigenvalue λ_2. This can also be explicitly checked using their definitions. Thus all real eigenvalues of P imply that $\lambda_2 = 0$ in all of them, or $Y = 0$. Then the definition of Y and $Y = 0$ imply $P = P^\dagger$.

An operator A is called *positive,* if all its eigenvalues are non-negative real numbers. Obviously, a positive operator is also a Hermitian operator.

3 Pure state density operators

We first show that all of the pure state quantum mechanics can be done with projectors $|\psi\rangle\langle\psi|$ replacing the ket $|\psi\rangle$. In this context, we call projector the pure state density operator $\rho = |\psi\rangle\langle\psi|$.

Consider a hermitian operator A corresponding to a measurable physical quantity. Let $\{\lambda_k; k = 1,\ldots,n\}$ be its eigenvalues and $\{|\lambda_k\rangle; k = 1,\ldots,n\}$ be its orthonormal eigenbasis. If we measure A on a large number N of identical quantum systems prepared in state $|\psi\rangle$, then we know that approximately a fraction $p_1 = |\langle\lambda_1|\psi\rangle|^2$ of measurements will give value $\lambda_1,\ldots$, a fraction $p_k = |\langle\lambda_k|\psi\rangle|^2$ of measurements will give value $\lambda_k,\ldots$, a fraction $p_n = |\langle\lambda_n|\psi\rangle|^2$ of measurements will give value λ_n, the approximation getting better with increasing N. Thus the average value of quantity A in the state $|\psi\rangle$ is

$$
\begin{aligned}
\langle A \rangle &= \sum_k \lambda_k p_k = \sum_k \lambda_k |\langle\lambda_k|\psi\rangle|^2 \\
&= \sum_k \lambda_k \langle\psi|\lambda_k\rangle\langle\lambda_k|\psi\rangle \\
&= \langle\psi|A \left[\sum_k |\lambda_k\rangle\langle\lambda_k| \right] |\psi\rangle \\
&= \langle\psi|A|\psi\rangle,
\end{aligned}
$$

where we have used Eq.(8). Now consider

$$
\begin{aligned}
tr(\rho A) &= tr(|\psi\rangle\langle\psi|A) = \sum_k \langle\lambda_k|\psi\rangle\langle\psi|A|\lambda_k\rangle \\
&= \langle\psi|A \left[\sum_k |\lambda_k\rangle\langle\lambda_k| \right] |\psi\rangle \\
&= \langle\psi|A|\psi\rangle,
\end{aligned}
$$

again using Eq.(8). Thus $tr(\rho A) = tr(|\psi\rangle\langle\psi|A)$ gives the average value of any measurable quantity A in an arbitrary pure (normalized) state $|\psi\rangle$.

The probability of finding the eigenvalue λ_l of A in a measurement of A on the state $|\psi\rangle$ is $|\langle\lambda_l|\psi\rangle|^2$. Consider

$$
\begin{aligned}
tr[|\lambda_l\rangle\langle\lambda_l|\rho] &= tr[(|\lambda_l\rangle\langle\lambda_l|)(|\psi\rangle\langle\psi|)] \\
&= \sum_k \langle\lambda_k|\lambda_l\rangle\langle\lambda_l|\psi\rangle\langle\psi|\lambda_k\rangle \\
&= |\langle\lambda_l|\psi\rangle|^2.
\end{aligned}
$$

Thus the probability of the value λ_l turning up in the measurement of A on the state $|\psi\rangle$ is simply the trace of the product of the projector $|\lambda_l\rangle\langle\lambda_l|$ and the corresponding pure state density operator ρ. We denote by E_l the projector $|\lambda_l\rangle\langle\lambda_l|$ of the lth eigenstate of an observable. Therefore, the probability p_l of getting lth eigenvalue when we measure A on ρ is

$$
p_l = tr(E_l\rho). \tag{20}
$$

For a pure state $|\psi\rangle$, the density operator is a projector $\rho = |\psi\rangle\langle\psi|$. Therefore it has to satisfy the idempotency condition

$$\rho^2 = \rho \tag{21}$$

Eq.(21) can be used as a criterion to check whether a given density operator is a pure state density operator. We will further refine it in the next section.

3.1 Phases in quantum mechanics

We digress for a brief exposition of the phase factors occuring in the expressions of quantum states and their physical relevance.

For a travelling harmonic wave, its phase appears as the argument of the trigonometric functions like $\sin(\theta)$ or $\cos(\theta)$, which are the phase factors. In the complex version, the phase factors appear in the form $e^{i\theta}$. Only the relative phase between two interfering waves has an observable effect. Similarly, phase factors can occur in two different ways in an expression of a quantum state. The so called global phase factor has no observable effect, while the relative phase factor can produce observable effects.

A phase factor $e^{i\theta}$ appearing in a state like $e^{i\theta}|\psi\rangle$ is called a global phase factor. The state $e^{i\theta}|\psi\rangle$ is the same as the state $|\psi\rangle$ up to a global phase factor. The amplitude for getting an eigenvalue λ of some observable A in the state $e^{i\theta}|\psi\rangle$ is $e^{i\theta}\langle\lambda|\psi\rangle$ and the corresponding probability is

$$P(\lambda) = e^{-i\theta}\langle\psi|\lambda\rangle\langle\lambda|\psi\rangle e^{i\theta} = \langle\psi|\lambda\rangle\langle\lambda|\psi\rangle = |\langle\lambda|\psi\rangle|^2,$$

which is the same for the state $|\psi\rangle$. Thus both the states $e^{i\theta}|\psi\rangle$ and $|\psi\rangle$ produce the same measurement statistics. Therefore, these two states are identical from the observational point of view. Thus global phase factors are irrelevant to the observed properties of quantum systems. We note that the global phase factors are automatically eliminated in pure state density operators. For a state $e^{i\theta}|\psi\rangle$ we get $e^{i\theta}|\psi\rangle\langle\psi|e^{-i\theta} = |\psi\rangle\langle\psi|$.

Another kind of phase is the relative phase. Consider a state expressed as a linear combination of some basis states. Two amplitudes say a and b occuring in this expansion are said to differ by a relative phase if there is a real θ such that $b = e^{i\theta}a$. Obviously, if the same state is expanded in another basis, we may not get amplitudes differing by a relative phase. Therefore, relative phase is a basis dependent phenomenon. Generally, two states, expanded in the same basis, are said to differ by a relative phase in that basis if each of the corresponding pair of amplitudes is related as $b = e^{i\theta}a$. For example, the two states

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \text{ and } |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

are the same up to a relative phase shift because the $|0\rangle$ amplitudes are identical (a relative phase factor of 1) while the $|1\rangle$ amplitudes differ by a relative phase factor of $-1 = e^{i\pi}$. Relative phase factors change from amplitude to amplitude and hence are basis dependent. When an observable with different eigenbasis is measured on these states,

they will, in general, show different measurement statistics. For example, if we measure $\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}$ on each of the states in the example above, then they will produce pairs of beams of spin $\frac{1}{2}$ particles with different intensities.

We have already seen an example of the physical relevance of the relative phase while dealing with Ramsay interferometry in the first write-up.

Thus we see that the states which differ only by relative phases in the same basis give rise to physically observable differences in measurement statistics so that they are physically different, unlike the states differing by a global phase factor.

You can convince yourself that relative phases do not get eliminated in the pure state density operator formalism.

4 Mixed state density operators

Consider a mixed state defined by the ensemble $\{p_i, |\psi_i\rangle; i = 1, \ldots, m\}$. The corresponding pure state density operators are the projectors $\rho_i = |\psi_i\rangle\langle\psi_i|$. To get the average value of an observable A in such a mixed state, we have to find the average of all the average values of A in the pure states $|\psi_i\rangle$ using the probability distribution p_i over $|\psi_i\rangle$ given by the ensemble. Thus we have

$$
\begin{aligned}
\langle A \rangle &= \sum_i p_i tr(A\rho_i) \\
&= tr\left[\left(\sum_i p_i \rho_i\right) A\right] \\
&= tr(\rho A)
\end{aligned}
\tag{22}
$$

where the operator

$$
\rho = \sum_i p_i \rho_i = \sum_i p_i |\psi_i\rangle\langle\psi_i| \quad \sum_i p_i = 1
\tag{23}
$$

must specify the mixed state, as the observable A is arbitrary.

In general, a mixed state is specified by an ensemble $\{p_i, \rho_i; i = 1, \ldots, m\}$ that is, by the probability distribution over m states $\rho_i; i = 1, \ldots, m$. The corresponding mixed state density operator ρ is given by the convex combination

$$
\rho = \sum_{i=1}^{m} p_i \rho_i \quad \sum_i p_i = 1
\tag{24}
$$

Note that ρ_i could themselves be mixed states. However, if say ρ_l in the sum defining the mixed state operator ρ in Eq.(24) is a mixed state operator itself, we can use the ensemble defining it to replace it by the corresponding convex combination of pure state projectors. Thus every mixed state density operator can be written as a convex combination of pure

state density operators, or projectors. However, these projectors need not be orthogonal. That is, the pure states making up a mixed state density operator need not be orthogonal.

Let us now find out the probability of obtaining mth eigenvalue as the outcome of measuring an observable A on the system in the mixed state ρ with the corresponding ensemble $\{p_i, |\psi_i\rangle; i = 1, \ldots, N\}$. The mth eigenvalue turns up in such a measurement by realizing one of the following mutually exclusive alternatives.

System is in the first pure state $|\psi_1\rangle\langle\psi_1|$ and mth eigenvalue is the outcome.

$$\cdots\cdots\cdots$$

$$\cdots\cdots\cdots$$

System is in the ith pure state $|\psi_i\rangle\langle\psi_i|$ and mth eigenvalue is the outcome.

$$\cdots\cdots\cdots$$

$$\cdots\cdots\cdots$$

System is in the Nth pure state $|\psi_N\rangle\langle\psi_N|$ and mth eigenvalue is the outcome.

What is the probability of each of these alternatives? say the ith alternative? The corresponding event is : the system is in the ith pure state *and* mth outcome turns up, given that the system is in the ith pure state $|\psi_i\rangle$. This is a compound event comprising two events and compound event occurs when both the events occur. Therefore the probability of the compound event is the product of the probabilities of the individual events. The probability of the system being in the ith pure state $\rho_i = |\psi_i\rangle\langle\psi_i|$ is p_i as specified by the ensemble. The probability that mth outcome turns up assuming the system to be in the ith pure state is $p(m|i) = tr(E_m\rho_i)$ (see Eq.(20)). Therefore the probability of ith alternative is $p_i tr(E_m\rho_i)$. The event that mth eigenvalue turns up can be realized by realizing any one of the above alternatives. The corresponding compound event is $x_1 \vee x_2 \vee \cdots x_i \vee \cdots \vee x_n$ where x_i stands for ith alternative and $\vee$ stands for the clause 'or'. The probability of this compound event is the addition of the probabilities of individual alternatives. Thus the probability that the mth outcome turns up after measuring the observable A on the system in the mixed state ρ is

$$
\begin{aligned}
p(m) &= \sum_{i=1}^{n} p_i p(m|i) \\
&= \sum_{i=1}^{n} p_i tr(E_m\rho_i) \\
&= tr\left(E_m \sum_{i=1}^{n} p_i\rho_i\right) \\
&= tr(E_m\rho).
\end{aligned}
\tag{25}
$$

4.1 Comparison with superposition of pure states

We emphasize that a mixed state must be clearly distinguished from a pure state superposition of the form [4]

$$|\psi\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle. \tag{26}$$

As we have seen in the first write-up, all component states are simultaneously present in the superposed state $|\psi\rangle$ which can always be experimentally verified, at least in principle. On the other hand, a mixed state is a classical mixture of pure states, where the system is exactly in one of the pure states $\{|\psi_i\rangle\}$ and not simultaneously in all of them. The mixed state is then specified by a distribution of probabilities p_i over $\{|\psi_i\rangle\}$ where p_i is the probability of the system being in the state $|\psi_i\rangle$. There is no a priori probabilistic element in the superposition like that in Eq.(26) : $|\psi\rangle$ is a pure state and therefore contains maximum possible knowledge about the system. The pure state density operator corresponding to the superposed state $|\psi\rangle$ in Eq.(26) is

$$\rho = |\psi\rangle\langle\psi| = \sum_{i,j} \sqrt{p_i p_j}|\psi_i\rangle\langle\psi_j| = \sum_i p_i|\psi_i\rangle\langle\psi_i| + \sum_{i\neq j} \sqrt{p_i p_j}|\psi_i\rangle\langle\psi_j| \tag{27}$$

The presence of the terms with $i \neq j$ (the so called off diagonal terms, see section 11) which represent interference between different states $|\psi_i\rangle$ clearly distinguishes this pure state density operator from the mixed state density operator in Eq.(23,24).

4.2 State after measurement

Let an observable A with eigenbasis $\{|k\rangle; k = 1, \ldots, n\}$ and eigenvalues $\{\lambda_k; k = 1, \ldots, n\}$ be measured on a pure state $|\psi\rangle$ to get eigenvalue λ_m as a result, so that the eigenstate $|m\rangle$ gets prepared. The probability that eigenvalue λ_m turns up is

$$p_m = \langle\psi|m\rangle\langle m|\psi\rangle = \langle\psi|E_m|\psi\rangle = tr(E_m\rho_\psi), \tag{28}$$

where $\rho_\psi = |\psi\rangle\langle\psi|$ is the projector for the input pure state $|\psi\rangle$. So the state after the measurement, $|m\rangle$, can be written as [5]

$$|m\rangle = \frac{|E_m|\psi\rangle|}{\sqrt{\langle\psi|E_m|\psi\rangle}}$$

or, in terms of projectors

$$\rho_m = |m\rangle\langle m| = \frac{E_m\rho_\psi E_m}{tr(E_m\rho_\psi)}. \tag{29}$$

---

[4]The form of superposition in Eq.(26) is chosen to make the distinction between the superposition and the mixed state evident.

[5]Substitute $E_m = |m\rangle\langle m|$ on RHS and remember that $\langle m|\psi\rangle$ is, in general, a complex number.

Now suppose the observable A defined above be measured on the mixed state given by the density operator

$$\rho = \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i| = \sum_{i=1}^{N} p_i \rho_i \qquad \sum_{i=1}^{N} p_i = 1.$$

If the initial state was the pure state $\rho_{in} = |\psi\rangle\langle\psi|$, we know that the state prepared after measurement with outcome λ_m is

$$\rho_{in}^m = |m\rangle\langle m| = \frac{E_m \rho_i E_m}{tr(E_m \rho_i)}.$$

When the initial state is the mixed state ρ defined above, the mixed state ρ' prepared after the outcome λ_m is defined over the ensemble made up of eigenstates $\{|m\rangle; m = 1, \ldots, n\}$. The probability of each $|m\rangle$ in ρ' is the conditional probability $p(i|m)$, that is, the probability that the system was in the pure state $|\psi\rangle$ given the outcome m. Using elementary probability theory, we have,

$$p(i|m) = \frac{p(i,m)}{p(m)} = \frac{p(m|i)p_i}{p(m)}.$$

We know that (see Eq.(20), Eq.(25))

$$\begin{aligned} p(m|i) &= tr(E_m \rho_i) \\ p(m) &= tr(E_m \rho) \end{aligned} \qquad (30)$$

so that

$$\begin{aligned} \rho' &= \sum_i p(i|m)|m\rangle\langle m| \\ &= \sum_i p_i \frac{tr(E_m \rho_i)}{tr(E_m \rho)} |m\rangle\langle m| \\ &= \sum_i p_i \frac{tr(E_m \rho_i)}{tr(E_m \rho)} \frac{E_m \rho_i E_m}{tr(E_m \rho_i)} \\ &= \frac{E_m \left(\sum_i p_i \rho_i\right) E_m}{tr(E_m \rho)} \\ &= \frac{E_m \rho E_m}{tr(E_m \rho)} \qquad (31) \end{aligned}$$

Thus the state of the system after the measurement of A as defined above on the mixed state ρ is

$$\rho' = \frac{E_m \rho E_m}{tr(E_m \rho)}. \qquad (32)$$

4.3   Some properties of the density operator

(i) Hermiticity : Consider the operator $|\beta\rangle\langle\alpha|$. We have

$$
\begin{aligned}
\langle\psi|[|\beta\rangle\langle\alpha|]^\dagger|\phi\rangle &= \langle\phi|[|\beta\rangle\langle\alpha|]|\psi\rangle^* \\
&= \langle\psi|\alpha\rangle\langle\beta|\phi\rangle \\
&= \langle\psi|[|\alpha\rangle\langle\beta|]|\phi\rangle
\end{aligned}
$$

Since $|\psi\rangle, |\phi\rangle$ are arbitrary, we get

$$
(|\beta\rangle\langle\alpha|)^\dagger = |\alpha\rangle\langle\beta|.
$$

We immediately get

$$
(|\alpha\rangle\langle\alpha|)^\dagger = |\alpha\rangle\langle\alpha|,
$$

that is, a projector is a hermitian operator. Therefore, for a mixed state density operator ρ we have

$$
\rho^\dagger = \sum_i p_i \rho_i^\dagger = \sum_i p_i \rho_i = \rho \tag{33}
$$

since $\rho_i^\dagger = \rho_i$ as ρ_i is a projector or a pure state. Thus a density operator is hermitian.

(ii) Unit trace : Consider the operator $|\beta\rangle\langle\alpha|$. We have already proved that

$$
tr(|\beta\rangle\langle\alpha|) = \langle\alpha|\beta\rangle.
$$

For a projector $|\alpha\rangle\langle\alpha|$ we get

$$
tr(|\alpha\rangle\langle\alpha|) = \langle\alpha|\alpha\rangle = 1
$$

because $|\alpha\rangle$ is normalized. Now consider

$$
tr(\rho) = tr\left(\sum_i p_i \rho_i\right) = \sum_i p_i tr(\rho_i) = \sum_i p_i = 1,
$$

as $tr(\rho_i) = 1$ since ρ_i are projectors and $\sum_i p_i = 1$ as p_is are probabilities and total probability must be unity. Thus the trace of a density operator must be unity :

$$
tr(\rho) = 1. \tag{34}
$$

(iii) Positivity : Consider a hermitian operator (observable) A with its orthonormal eigenbasis [6] $\{|k\rangle; k = 1, \ldots, n\}$ and eigenvalues $\{\lambda_k; k = 1, \ldots, n\}$. Using Eq.(8) we can

---

[6]In section 2 and Appendix B we have seen that a normal operator is diagonalized by an orthonormal basis of the state space and that a Hermitian operator is normal.

write

$$
\begin{aligned}
A &= \sum_{i,j} |i\rangle\langle i|A|j\rangle\langle j| \\
&= \sum_{i,j} \lambda_j |i\rangle\langle j|\langle i|j\rangle \\
&= \sum_i \lambda_i |i\rangle\langle i| \tag{35}
\end{aligned}
$$

The second equality follows because $|j\rangle$ is an eigenstate of A with eigenvalue λ_j. The third equality follows because the eigenbasis of an hermitian operator is orthonormal, so that $\langle i|j\rangle = \delta_{ij}$. If an eigenvalue is m fold degenerate, it will repeat m times in the sum in Eq.(35). Thus any hermitian operator can be written as the linear combination of the projectors on its eigenvectors ($|i\rangle\langle i|$) with coefficients as its eigenvalues (λ_i). This is called spectral decomposition of a Hermitian, (in general normal), operator. Eq.(35) is often called spectral representation of the operator.

Since any density operator is a hermitian operator, it is always diagonalizable, that is, we can always find all its eigenvalues and its orthonormal eigenbasis (see Appendix B). Applying spectral theorem to ρ we get,

$$
\rho = \sum_\alpha p_\alpha |\alpha\rangle\langle\alpha|
$$

with $\{p_\alpha : \alpha = 1, \ldots, n\}$ as its eigenvalues and $\{|\alpha\rangle; \alpha = 1, \ldots, n\}$ as its orthonormal eigenbasis. This is the mixed state corresponding to the classical mixture of eigenstates $\{|\alpha\rangle; \alpha = 1, \ldots, n\}$ with probabilities $\{p_\alpha : \alpha = 1, \ldots, n\}$. To see that p_α are the eigenvalues of ρ, consider its action on some eigenstate $|\beta\rangle$.

$$
\rho|\beta\rangle = \sum_\alpha p_\alpha |\alpha\rangle\langle\alpha|\beta\rangle = \sum_\alpha p_\alpha |\alpha\rangle\delta_{\alpha\beta} = p_\beta|\beta\rangle,
$$

where we have used the orthonormality of the eigenbasis of ρ. Further, the unit trace condition on ρ implies that the sum of all its eigenvalues must be unity, namely, $\sum_\alpha p_\alpha = 1$. We can thus unambiguously identify the eigenvalues of a density operator as the probabilities occuring in the ensemble comprising its eigenstates which generates this density operator. We can immediately see that the same density operator can be prepared using different ensembles, because the density operator ρ was given (so it must have been prepared by the preparation procedure corresponding to some ensemble) and the same density operator is shown to be prepared by the ensemble comprising its eigenstates with its eigenvalues as the probabilities. Thus these two ensembles and the corresponding preparation procedures produce the same density operator and hence the same mixed state. This observation and its consequences will be discussed in detail in section 7.

Since the eigenvalues of a density operator ρ are probabilities we must have $\{p_\alpha \geq 0; \alpha = 1, \ldots, n\}$ Thus all the eigenvalues of a density operator have to be non-negative. If

the eigenvalue corresponding to some eigenstates is zero, these eigenstates are absent from the ensemble defining ρ so that ρ is defined over the subspace of the state space which is orthogonal to these eigenstates. Restricted to this subspace, the density operator ρ has all positive eigenvalues. If ρ is non-negative, then for any normalized state $|\psi\rangle$ we have

$$\langle\psi|\rho|\psi\rangle = \sum_\alpha p_\alpha \langle\psi|\alpha\rangle\langle\alpha|\psi\rangle = \sum_\alpha p_\alpha |\langle\alpha|\psi\rangle|^2 \geq 0$$

since $p_\alpha \geq 0\ \forall\alpha$. Thus the general condition for non-negativity of an operator is

$$\langle\psi|\rho|\psi\rangle \geq 0\ \forall|\psi\rangle \in \mathcal{H}. \tag{36}$$

(iv) Criterion for ρ to be a pure state. : A density operator ρ corresponding to a pure state is a projector $\rho = |\psi\rangle\langle\psi|$ satisfying $\rho^2 = \rho$. This gives

$$tr(\rho^2) = tr(\rho) = 1 \tag{37}$$

as a criterion for the system to be in the pure state. In fact this is both necessary and sufficient condition, because for a mixed state with at least two non-zero probabilities p_i, $tr(\rho^2) < 1$. We have,

$$tr(\rho^2) = tr\left(\sum_{ij} p_i p_j |\psi_i\rangle\langle\psi_j|\langle\psi_i|\psi_j\rangle\right) = \sum_{ij} p_i p_j |\langle\psi_i|\psi_j\rangle|^2 < 1,$$

assuming that $\langle\psi_i|\psi_j\rangle \neq 0$ for $i \neq j$. Even if $\langle\psi_i|\psi_j\rangle = 0$ for $i \neq j$, we have

$$tr(\rho^2) = tr\left(\sum_i p_i^2 |\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i^2 < 1.$$

Thus for a mixed state $tr(\rho^2) < 1$, which means that Eq.(37) gives a necessary and sufficient condition for ρ to be a pure state.

4.4   An important theorem

It turns out that any operator ρ satisfying the trace condition, Eq(34), hermiticity and non-negativity condition Eq.(36), is a density operator associated with some ensemble. In fact we have the following theorem.

Theorem : A diagonalizable linear operator ρ is a density operator associated with some ensemble $\{p_i, |\psi_i\rangle\}$ if and only if it satisfies the conditions [7]

(i) (Trace condition) $tr(\rho) = 1$.

(ii) (Hermiticity) $\rho = \rho^\dagger$

---

[7]We have seen in section that a positive operator is necessarily Hermitian although we prefer to state the Hermiticity condition explicitly here.

(iii) (Positivity condition) ρ satisfies Eq.(36).

Proof : Suppose $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ is a density operator associated with an ensemble $\{p_i, \psi_i\}$. Then we have already proved that it satisfies all the three conditions stated in the theorem. Now suppose that ρ is an operator satisfying the three conditions stated in the theorem. Since ρ is hermitian, and positive, it must have the spectral decomposition

$$\rho = \sum_j \lambda_j |j\rangle\langle j|; \quad \lambda_j \geq 0, \tag{38}$$

where $\{|j\rangle\}$ forms the orthonormal eigenbasis of ρ and λ_j are the eigenvalues of ρ. From the trace condition we see that $\sum_j \lambda_j = 1$. Therefore, the ensemble $\{\lambda_j, |j\rangle\}$ defining a mixed state is the ensemble giving rise to the density operator ρ in Eq.(38). [QED]

Sometimes we get an operator which is positive and has a finite trace which does not equal unity. In such a case, we can just divide this operator by its trace, to make it a density operator.

5 Geometry of mixed states : convex sets

We wish to understand a very important characteristic of mixed states. We start with defining a convex set. Let E^n be a space same as $\mathbb{R}^n$ with Euclidean metric defined on it. A set of points in E^n is convex if all points on the line joining every pair of points in it are the elements of that set . If $\mathbf{x}_1$ and $\mathbf{x}_2$ are two points in a convex set in E^n, then all the points $\{\mathbf{x}\}$ on the line joining $\mathbf{x}_1$ and $\mathbf{x}_2$, that is, on the line (see Fig. 1),

$$\mathbf{x}(\lambda) = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \quad 0 \leq \lambda \leq 1 \tag{39}$$

are elements of that set. The RHS of Eq.(39) is called a convex combination of two terms. Another way to characterize a convex set is that, sitting at any point inside the set, you can see every point in it, because every other point in the set can be joined by a straight line to that point, which lies totally inside the set. Fig.2 gives some examples of convex sets. We have already established that the real linear space comprising Hermitian operators is isomorphic with $\mathbb{R}^{n^2}$. We know that the set of all density operators sits in the linear space of Hermitian operators, which is a real space of dimension n^2, given that the state space has dimension n. However, in addition to the Hermiticity conditions a density operator has to satisfy the unit trace condition, Eq(34), so that this isomorphism is restricted to that between the set of density operators and an $n^2 - 1$ dimensional hypersurface in $\mathbb{R}^{n^2}$. In order to get the $n^2 - 1$ dimensional image set of the set of all density operators under this isomorphism, we have to find the set of real linear combinations of a basis in the Hermitian operator space which give valid density operators (positive operators with unit trace). This task is not easy and explicitly constructed image sets of density operators are available only in the $n = 2$ case. The corresponding 3 dimensional image set is called Bloch ball with a spherical boundary called Bloch sphere. We will go through the detailed
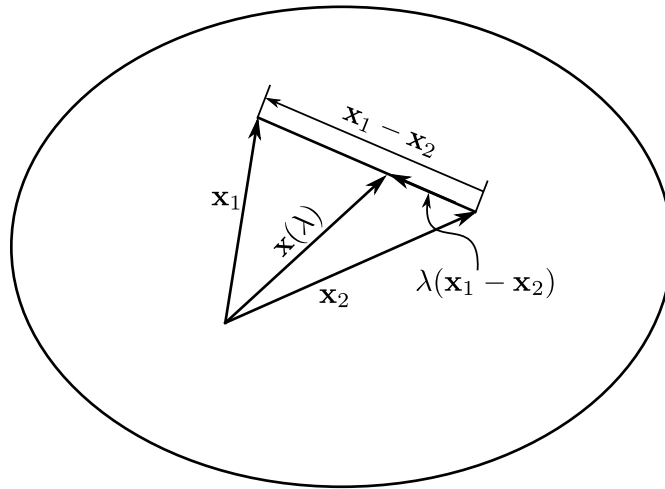
Figure 1: $\mathbf{x}(\lambda) = \mathbf{x}_2 + \lambda(\mathbf{x}_1 - \mathbf{x}_2) = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$.

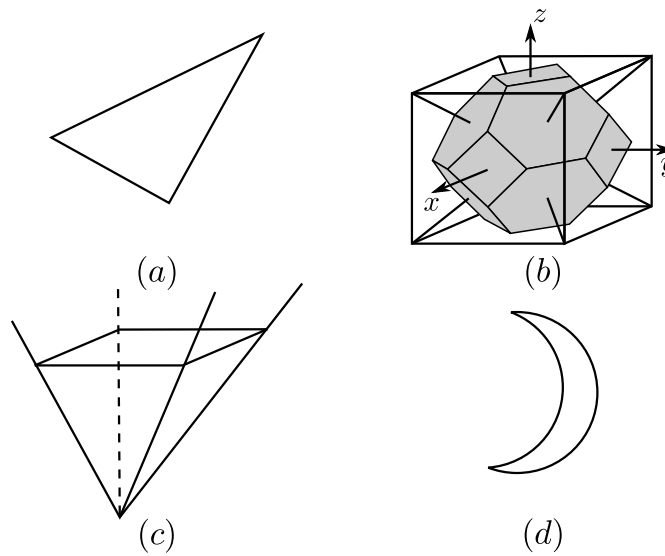

Figure 2: Examples of convex sets (a) A triangle in $2-D$. (b)The unit cell of Body Center Cubic lattice as well as its Wigner seitz cell are both convex sets. (c) A cone with square base. (d) The crescent moon is NOT a convex set.

construction of the Bloch ball and sphere in section 9. However, we can show that in all dimensions, this image set is a convex set.

To get to the convexity of the image set, we show that the operator given by a convex combination of two density operators, namely,

$$\rho = \lambda\rho_1 + (1 - \lambda)\rho_2 \quad 0 \leq \lambda \leq 1 \tag{40}$$

is also a density operator. In fact ρ defined by Eq.(40) has all the characteristics of a density operator. Using the fact that ρ_1 and ρ_2 are density operators, we get,

(i) Hermiticity : $\rho^\dagger = \lambda\rho_1^\dagger + (1 - \lambda)\rho_2^\dagger = \lambda\rho_1 + (1 - \lambda)\rho_2 = \rho$.

(ii) Unit trace : $tr(\rho) = \lambda tr(\rho_1) + (1 - \lambda)tr(\rho_2) = \lambda + (1 - \lambda) = 1$.

(iii) Positivity : For an arbitrary state $|\psi\rangle$ we get,

$$\langle\psi|\rho|\psi\rangle = \lambda\langle\psi|\rho_1|\psi\rangle + (1 - \lambda)\langle\psi|\rho_2|\psi\rangle \geq 0,$$

since both the terms on the right are non-negative.

Thus we see that a convex combination of two density operators is also a density operator. The isomorphic images of ρ_1, ρ_2 and ρ, say $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}$ would then satisfy Eq.(40), or, $\mathbf{x}$ lies on the line joining $\mathbf{x}_1$ and $\mathbf{x}_2$. This shows that the image set is convex.

Let us find the conditions on λ, ρ_1 and ρ_2 for ρ to be a pure state. For a pure state we need $tr(\rho^2) = 1$. We have, using Eq.(40),

$$tr(\rho^2) = \lambda^2 tr(\rho_1^2) + 2\lambda(1 - \lambda)tr(\rho_1\rho_2) + (1 - \lambda)^2 tr(\rho_2^2).$$

Thus $tr(\rho^2) = 1$ provided either (a) $\lambda = 0$ and ρ_2 is a pure state, $(tr(\rho_2^2) = 1)$, or (b) $\lambda = 1$ and ρ_1 is a pure state $(tr(\rho_1^2) = 1)$. In fact, for $\lambda = 1$ or $\lambda = 0$, $\rho = \rho_1$ or $\rho = \rho_2$. Thus a pure state density operator cannot be written as a convex combination of two *different* density operators with $0 < \lambda < 1$. Geometrically, such points in a convex set are called extreme points. They lie on the boundary of the convex set and can only be the end points of a line wholly lying in the set. Thus pure state density operators form a subset of extreme points in the convex set corresponding to all density operators.

6 Temporal evolution of a mixed state

To find the time evolution of the mixed states we work in the Schroedinger picture, where states evolve in time. At time $t = 0$ let the mixed state be given by the density operator

$$\rho(0) = \sum_i p_i\rho_i(0) = \sum_i p_i|\psi_i(0)\rangle\langle\psi_i(0)|.$$

The probabilities p_i for the pure states $|\psi_i(0)\rangle$ are independent of time, while $|\psi_i(0)\rangle$ unitarily evolve in time, that is, the states evolve in time such that their norms and scalar products remain invariant in time :

$$|\psi_i(0)\rangle \rightarrow U(t)|\psi_i(0)\rangle = |\psi_i(t)\rangle.$$

Therefore,

$$
\begin{aligned}
\rho(t) &= \sum_i p_i |\psi_i(t)\rangle\langle\psi_i(t)| \\
&= \sum_i p_i U(t)|\psi_i(t)\rangle\langle\psi_i(t)|U^\dagger(t) \\
&= U(t)\rho(0)U^\dagger(t).
\end{aligned}
\tag{41}
$$

In order to get $\rho(t)$ from $\rho(0)$ we differentiate Eq.(41) with time to get

$$
\dot{\rho} \equiv \frac{d\rho}{dt} = \frac{dU}{dt}\rho(0)U^\dagger + U\rho(0)\frac{dU^\dagger}{dt}.
\tag{42}
$$

To obtain the time evolution equation for $U(t)$ we assume that the infinitesimal time evolution operator has the form

$$
U(dt) = I - \frac{iHdt}{\hbar}
$$

where I is the identity operator and H is the Hamiltonian of the system. This form guarantees the continuity of time translation by satisfying $U(t) \to I$ as $t \to 0$. Hamiltonian is taken to be the generator of evolution. Using the composition property of $U(t)$ we can write

$$
U(t+dt,0) = U(t+dt,t)U(t,0) = \left(I - \frac{iHdt}{\hbar}\right)U(t,0),
$$

from which we get, dropping 0,

$$
U(t+dt) - U(t) = -\frac{i}{\hbar}HU(t)dt.
$$

Dividing by dt and taking the limit $dt \to 0$ we get,

$$
i\hbar\frac{dU(t)}{dt} = HU(t).
\tag{43}
$$

For the adjoint operator $U^\dagger$ we get

$$
-i\hbar\frac{dU^\dagger(t)}{dt} = U^\dagger(t)H.
\tag{44}
$$

Substituting from Eq.(43) and Eq.(44) the expressions for $\frac{dU(t)}{dt}$ and $\frac{dU^\dagger(t)}{dt}$ in Eq.(42) we get

$$
\begin{aligned}
i\hbar\frac{d\rho}{dt} &= HU\rho(0)U^\dagger - U\rho(0)U^\dagger H \\
&= [H, \rho(t)].
\end{aligned}
\tag{45}
$$

Eq.(45) looks like the Heisenberg equation of motion, but remember that ρ is not a dynamical variable in the Heisenberg picture. On the contrary, ρ is built up by Schroedinger picture state kets and bras which evolve in time according to Schroedinger equation.

We can get explicit form of $U(t)$ by expressing $U(t)$ as the composition of infinitesimal time evolution operators as

$$U(t) = \lim_{N \to \infty} \left[I - \frac{\frac{iH}{\hbar}t}{N} \right]^N = \exp\left[-\frac{iHt}{\hbar} \right], \tag{46}$$

in which case we can obtain $\rho(t)$ from $\rho(0)$ by directly using Eq.(41). However, this prescription works only when the Hamiltonian is independent of time. You can easily check that Eq.(46) solves Eq.(43) provided Hamiltonian is independent of time.

In general, to get the time evolution $\rho(t)$ from $\rho(0)$ we have to solve Eq.(45) with the initial condition $\rho(t) = \rho(0)$ at $t = 0$. Various methods are developed to get approximate solutions of Eq.(45) in different situations.

7   The preparation ambiguity of mixed state density operators

We have seen how a density operator representing a mixed state can be constructed in terms of a classical probability distribution $\{p_i, i = 1, \ldots, m\}$ over a set of pure quantum states $\{|\psi_i\rangle, i = 1, \ldots, m\}$. Such a density operator is given by Eq.(23). The probability distribution $\{p_i\}$ over $\{|\psi_i\rangle\}$ is the result of the nebulous preparation procedure of the state : we know that the system is actually prepared in one of the states $\{|\psi_i\rangle\}$ but we do not possess specific information about which of the states $\{|\psi_i\rangle\}$ has been prepared. Therefore the best we can do is to assign probabilities $\{p_1 \mapsto |\psi_1\rangle, p_2 \mapsto |\psi_2\rangle, \ldots, p_i \mapsto |\psi_i\rangle, \ldots, p_m \mapsto |\psi_m\rangle, \}$ for the system to be actually in the corresponding state, using whatever partial information on the preparation that we have. Thus every set of pairs, (also called ensemble), $\{(p_i, |\psi_i\rangle); i = 1, \ldots, m\}$ corresponds to one nebulous preparation procedure which gives rise to a mixed state with density operator as in Eq.(23). The probability p_i is approximated by the fraction of systems, in a large set of non-interacting systems called their ensemble, get prepared in the pure state $|\psi_i\rangle$ when the same preparation is carried out on each one of them.

The interesting question is : Given a density operator ρ can we find out the ensemble $\{(p_i, |\psi_i\rangle); i = 1, \ldots, m\}$, that is, the preparation which gives rise to it? The answer is NO. It turns out that, given a density operator ρ, we can find infinity of ensembles $\{p_i, |\psi_i\rangle\}$ which give rise to the same density operator. This is called the preparation ambiguity of mixed state density operators, or, if the set $\{|\psi_i\rangle\}$ forms a basis of the state space $\mathcal{H}$, the basis ambiguity of the mixed state density operator ρ. (see subsection 7.1)

In section 5 we saw that a pure state density operator cannot be decomposed into a convex combination of different density operators. However, a mixed state density

operator can be decomposed into pure state density operators (projectors) in infinitely many ways. For example, a density operator over a $2 - D$ Hilbert space with $\{|0\rangle, |1\rangle\}$ as an orthonormal basis, namely,

$$\rho = 0.7|0\rangle\langle 0| + 0.3|1\rangle\langle 1|$$

can be decomposed as

$$\rho = 0.4|0\rangle\langle 0| + 0.3|\alpha\rangle\langle\alpha| + 0.3|\beta\rangle\langle\beta|, \tag{47}$$

where

$$|\alpha\rangle = 0.8|0\rangle + 0.6|1\rangle \ \text{ and } \ |\beta\rangle = 0.6|0\rangle - 0.8|1\rangle$$

as can be verified by a straightforward calculation. What is important is to realize that we can change the probabilities $(0.4, 0.3, 0.3)$ in Eq.(47) and redefine $|\alpha\rangle$ and $|\beta\rangle$ in infinitely many ways to get the same density operator ρ.

We can generically show that there is a continuum of ensembles $\{p_i, |\psi_i\rangle\}$ giving rise to the same density operator ρ by using Eq.(40) which expresses ρ as a convex combination of density operators ρ_1 and ρ_2. We can generically define the preparation procedure for ρ in Eq.(40) as follows. Let a random process have probability λ to succeed and probability $(1 - \lambda)$ to fail. ('Succeed' and 'fail' correspond to two mutually exclusive outcomes of the process). In the case of success we prepare the state ρ_1 while in the case of failure we prepare the state ρ_2. That this preparation procedure prepares ρ in Eq.(40) can be seen by evaluating the average value of an arbitrary observable A using this procedure. We have,

$$
\begin{aligned}
\langle A \rangle &= \lambda tr(\rho_1 A) + (1 - \lambda)tr(\rho_2 A) \\
&= tr[(\lambda\rho_1 + (1 - \lambda)\rho_2)A] \\
&= tr(\rho A).
\end{aligned}
\tag{48}
$$

Since A is an arbitrary observable, Eq.(48) implies Eq.(40), thus showing that the above preparation procedure is consistent with Eq.(40). The fact that the generic preparation procedure given above for ρ in Eq.(40) is actually a continuum of preparation procedures can be seen from Fig.3 which shows that ρ is a convex combination of two pure states at the end points of a line passing through ρ. Each such line gives a new set $\{\lambda, \rho_1, \rho_2\}$ giving the same density operator ρ. This proves that every mixed state density operator can be prepared in infinitely many ways corresponding to infinitely many preparations $(p_i, |\psi_i\rangle)$. This infinite ambiguity regarding the preparation of a density operator ρ for a mixed state implies that the density operator contains all the information, that we can extract, about the (mixed) state of the system and its evolution, independent of its preparation. Thus in Eq.(48) we see that the average value $\langle A \rangle$ of an arbitrary observable A depends exclusively on the density operator ρ and is independent of which triplet $\{\lambda, \rho_1, \rho_2\}$ was used to prepare ρ.
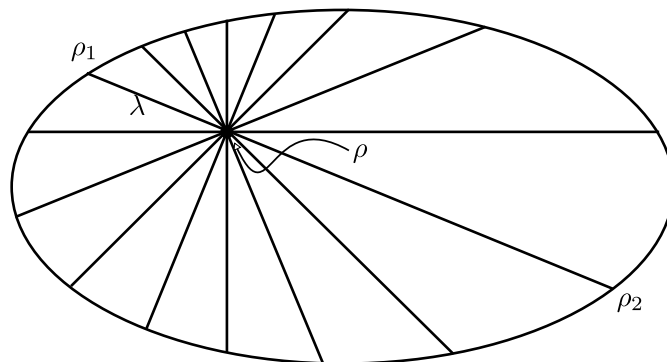
Figure 3: Each line passing through point ρ, gives a set $(\rho_1, \rho_2, \lambda)$ satisfying Eq.(40) for the same ρ.

As an example, we may have an experimental set up preparing the linear polarization states $|H\rangle$ and $|V\rangle$ each with probability $\frac{1}{2}$. Thus each photon is in the mixed polarization state specified by the ensemble $\{(\frac{1}{2}, |H\rangle), (\frac{1}{2}, |V\rangle)\}$ whose density operator is

$$\rho = \frac{1}{2}\{|H\rangle\langle H| + |V\rangle\langle V|\} = \frac{1}{2}I. \tag{49}$$

The last equality follows because $\{|H\rangle, |V\rangle\}$ is an orthonormal basis in the polarization state space so that Eq.(8) applies.

We may now use another experimental set up to prepare the ensemble of two circularly polarized states $\{|L\rangle, |R\rangle\}$ each with probability $\frac{1}{2}$, with the corresponding ensemble $\{(\frac{1}{2}, |L\rangle), (\frac{1}{2}, |R\rangle)\}$. The density operator of this state is

$$\rho = \frac{1}{2}\{|L\rangle\langle L| + |R\rangle\langle R|\} = \frac{1}{2}I, \tag{50}$$

because $\{|L\rangle, |R\rangle\}$ is also an orthonormal basis in the polarization state space so that Eq.(8) applies.

An observer receiving megajules of these photons (in the mixed polarization state $\rho = \frac{1}{2}I$) will never be able to discover which one of these two methods of preparation were used to produce ρ, despite the fact that these are two different preparation procedures. Thus all possible information about the system that can be obtained is totally contained in ρ alone.

As another example imagine an experimental set up preparing a mixed state corresponding to the ensemble $\{(\frac{1}{2}, |\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}, 0\rangle), (\frac{1}{2}, |\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}, 1\rangle)\}$ where $\{|\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}, 0\rangle, |\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}, 1\rangle\}$ is the orthonormal eigenbasis of the observable $\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}$, that is, these are the spin up and spin down states along the axis specified by the unit vector $\hat{\mathbf{n}}$. The resulting mixed state has the density operator

$$\rho = \frac{1}{2}\{|\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}, 0\rangle\langle\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}, 0| + |\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}, 1\rangle\langle\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{n}}, 1|\} = \frac{1}{2}I, \tag{51}$$

27

where we have used Eq.(8). As we continuously rotate $\hat{\mathbf{n}}$, we scan through infinity of preparation procedures corresponding to infinity of ensembles given above, each specified by different unit vector $\hat{\mathbf{n}}$. All these infinitely many different preparation procedures produce the same density operator $\rho = \frac{1}{2}I$. An observer receiving a flux of these spin $\frac{1}{2}$ particles (each in the state $\rho = \frac{1}{2}I$) will not be able to discover exactly which of these infinitely many preparation procedures was used to produce ρ. Thus ρ contains all the information that can be obtained about the system.

We express this fact as a fundamental postulate of quantum mechanics.

The density operator completely specifies all the properties of the system in any state.
Note that this postulate applies to pure states as well.

7.1 Ensembles generating the same density operator

If we consider two arbitrary ensembles of pure states, then, in general, they will generate different density operators. Therefore the question arises as to the condition under which two ensembles, say $\{p_i, |\psi_i\rangle\}$ and $\{q_j, |\phi_j\rangle\}$ generate the same density operator[8]. If we look at the examples we have considered in the previous section, the states in the two ensembles generating the same density operator form orthonormal bases of the corresponding state space. Two orthonormal bases are basically connected by a unitary transformation. Thus we expect that two ensembles generating the same density operator must be unitarily connected. This is indeed the case and in fact the condition of unitary connection between the sets of states forming the ensembles is both necessary and sufficient.

Consider two ensembles $\{p_i, |\psi_i\rangle\}$ $i = 1, 2, \ldots, N$ and $\{q_j, |\phi_j\rangle\}$ $j = 1, 2, \ldots, M$. In order to express unitary connection between them, it is convenient to define kets

$$
\begin{aligned}
|\tilde{\psi}_i\rangle &= \sqrt{p_i}|\psi_i\rangle \ \ i = 1, 2, \ldots, N \\
|\tilde{\phi}_j\rangle &= \sqrt{q_j}|\phi_j\rangle \ \ j = 1, 2, \ldots, M
\end{aligned}
\tag{52}
$$

Note that the states $\{|\tilde{\psi}_i\rangle\}$ and $\{|\tilde{\phi}_j\rangle\}$ are not normalized although $\{|\psi_i\rangle\}$ and $\{|\phi_j\rangle\}$ are.

In general, $M \neq N$, that is, the cardinality of the sets $\{|\psi_i\rangle\}$ and $\{|\phi_j\rangle\}$ defining two ensembles are different. If $N > M$, we pad up the set $\{|\tilde{\phi}_j\rangle\}$ by $N - M$ zero (null) kets. If $N < M$ we pad up the set $\{|\tilde{\psi}_i\rangle\}$ by $M - N$ zero (null) kets. We assume that this is done and that both these sets are of the same size say N.

We now wish to prove the following result.

Theorem : The sets $\{|\tilde{\psi}_i\rangle\}$ $i = 1, 2, \ldots, N$ and $\{|\tilde{\phi}_j\rangle\}$ $j = 1, 2, \ldots, N$ generate the same density operator ρ if and only if

$$
|\tilde{\psi}_i\rangle = \sum_j u_{ij}|\tilde{\phi}_j\rangle \ \ i = 1, 2, \ldots, N,
\tag{53}
$$

where $[u_{ij}]$ is a $N \times N$ complex unitary matrix.

---

[8]This subsection may be dropped at the first reading.

Note that, in terms of normalized states $\{|\psi_i\rangle\}$ and $\{|\phi_j\rangle\}$ and the corresponding probability distributions $\{p_i\}, \{q_j\}$ Eq.(53) becomes

$$\sqrt{p_i}|\psi_i\rangle = \sum_j u_{ij}\sqrt{q_j}|\phi_j\rangle \tag{54}$$

for some $N \times N$ complex unitary matrix $[u_{ij}]$.

Proof [1] : Suppose Eq.(53) holds for some unitary matrix $[u_{ij}]$. Then

$$
\begin{aligned}
\sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| &= \sum_{ijk} u_{ij}u_{ik}^*|\tilde{\phi}_j\rangle\langle\tilde{\phi}_k| \\
&= \sum_{jk}\left(\sum_i u_{ki}^\dagger u_{ij}\right)|\tilde{\phi}_j\rangle\langle\tilde{\phi}_k| \\
&= \sum_{jk}\delta_{jk}|\tilde{\phi}_j\rangle\langle\tilde{\phi}_k| \\
&= \sum_j |\tilde{\phi}_j\rangle\langle\tilde{\phi}_j| \tag{55}
\end{aligned}
$$

which shows that the sets $\{|\tilde{\psi}_i\rangle\}$ $i = 1, 2, \ldots, N$ and $\{|\tilde{\phi}_j\rangle\}$ $j = 1, 2, \ldots, N$ generate the same density operator ρ

Conversely, suppose

$$\sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \rho = \sum_j |\tilde{\phi}_j\rangle\langle\tilde{\phi}_j|. \tag{56}$$

where ρ is a density operator. Let ρ have the spectral decomposition

$$\rho = \sum_{k=1}^m \lambda_k|k\rangle\langle k| \quad m \leq n \tag{57}$$

Here n is the dimension of the state space, $\{|k\rangle\}$ $k = 1, \ldots, m$ are the orthonormal eigenstates and $\{\lambda_k\}$ are strictly positive eigenvalues of ρ.

If $m = n$ in Eq.(57), then the eigenbasis $\{|k\rangle\}$ $k = 1, \ldots, n$ of the density operator ρ spans the whole state space so that every $\{|\tilde{\psi}_i\rangle\}$ $i = 1, 2, \ldots, N$ in the first ensemble can be expressed as a linear combination of the orthogonal basis set $\{|\tilde{k}\rangle = \sqrt{\lambda_k}|k\rangle\}$ $k = 1, \ldots, n$. That is,

$$|\tilde{\psi}_i\rangle = \sum_{k=1}^n v_{ik}|\tilde{k}\rangle \quad i = 1, \ldots, N. \tag{58}$$

If $m < n$ in Eq.(57), let $|\psi\rangle$ be a state in the orthogonal complement of the subspace spanned by $|k\rangle$ $k = 1, \ldots, m$. Therefore, $\langle\psi|\tilde{k}\rangle\langle\tilde{k}|\psi\rangle = 0$ for all $|\tilde{k}\rangle$, so that $\langle\psi|\rho|\psi\rangle = 0$ via Eq.(57). We thus get

$$0 = \langle\psi|\rho|\psi\rangle = \sum_i \langle\psi|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|\psi\rangle = \sum_i |\langle\psi|\tilde{\psi}_i\rangle|^2.$$

Since the vanishing of the sum of non-negative terms implies vanishing of each term separately, we get $\langle \psi | \tilde{\psi}_i \rangle = 0$ for all $i = 1, \ldots, N$ and all $|\psi\rangle$ in the orthogonal complement of the subspace spanned by $\{|\tilde{k}\rangle\}$. Thus all $\{|\tilde{\psi}_i\rangle\}$ $i = 1, 2, \ldots, N$ are in the subspace spanned by $\{\tilde{k}\}$ $k = 1, \ldots, m$ and hence can be expressed as a linear combination of $\{|\tilde{k}\rangle\}$ $k = 1, 2, \ldots, m$:

$$|\tilde{\psi}_i\rangle = \sum_{k=1}^{m} v_{ik}|\tilde{k}\rangle \quad i = 1, \ldots, N. \tag{59}$$

By Eq.(56) coupled with Eq.(58) or Eq.(59) we get

$$\sum_k |\tilde{k}\rangle\langle \tilde{k}| = \sum_{kl} \left(\sum_i v_{ik}^* v_{il} \right) |\tilde{l}\rangle\langle \tilde{k}|.$$

Taking scalar product on both sides with eigenket $|\tilde{s}\rangle$ of ρ and using their orthogonality we get

$$|\tilde{s}\rangle = \sum_l \left(\sum_i v_{is}^* v_{il} \right) |\tilde{l}\rangle.$$

This equality is possible only when

$$\sum_i v_{is}^* v_{il} = \delta_{sl}. \tag{60}$$

This just means that every pair of columns in the $V = [v_{ij}]_{N \times m}$ matrix connecting $|\tilde{\psi}_i\rangle$ and $\{|k\rangle\}$ is orthogonal. Written in the matrix form Eq.(59) reads

$$[|\tilde{\psi}\rangle]_{N \times 1} = V_{N \times m}[|\tilde{k}\rangle]_{m \times 1}. \tag{61}$$

In exactly similar way we can show that

$$|\tilde{\phi}_i\rangle = \sum_{k=1}^{m} w_{ik}|\tilde{k}\rangle \quad i = 1, \ldots, N. \tag{62}$$

where every pair of columns in the $W = [w_{ik}]_{N \times m}$ matrix are orthogonal. In the matrix form we have,

$$[|\tilde{\phi}\rangle]_{N \times 1} = W_{N \times m}[|\tilde{k}\rangle]_{m \times 1}. \tag{63}$$

Thus we get the matrix equations,

$$[|\tilde{\psi}\rangle] = V[|\tilde{k}\rangle] \tag{64}$$

$$[|\tilde{\phi}\rangle] = W[|\tilde{k}\rangle]. \tag{65}$$

30

By virtue of the orthonormality of columns of V (see Eq.(60)) and the corresponding relation for W we get

$$
\begin{aligned}
V^\dagger_{m\times N} V_{N\times m} &= I_{m\times m} \\
W_{N\times m} W^\dagger_{m\times N} &= I_{N\times N}
\end{aligned}
\tag{66}
$$

Therefore, if we multiply Eq.(65) on the left by $W^\dagger$ and substitute the resulting expression for $[|\tilde{k}\rangle]$ in Eq.(64)we get

$$
[|\tilde{\psi}\rangle]_{N\times 1} = [VW^\dagger]_{N\times N} [|\tilde{\phi}\rangle]_{N\times 1}.
\tag{67}
$$

Unitarity of $[VW^\dagger]$ is easily established using Eqs.(66)

$$
[VW^\dagger]^\dagger [VW^\dagger] = WV^\dagger VW^\dagger = I_{N\times N}.
$$

Thus we finally get

$$
[|\tilde{\psi}\rangle]_{N\times 1} = [U]_{N\times N} [|\tilde{\phi}\rangle]_{N\times 1},
$$

where

$$
U = VW^\dagger
$$

is unitary. This completes the proof.

8 Random mixtures

The density operators like $\frac{1}{2}I$ are quite special. The corresponding mixed states are called random mixtures. This is the equal weight classical mixture of orthonormal pure states which form a basis of the state space of the system. The density operator $\frac{1}{N}I$ (for a quantum system with N states, that is, $dim(\mathcal{H}) = N$) is invariant under any change of orthonormal basis : $U^\dagger \left(\frac{1}{N}I\right) U = \frac{1}{N}I$. Since every orthonormal basis defines (at least in principle,) a possible measurement, (of the observable whose eigenbasis is the given orthonormal basis), we see that, in the state $\rho = \frac{1}{N}I$, all outcomes of every measurement are equally probable, with probability $\frac{1}{N}$. We have, using Eq.(25),

$$
p(m) = tr(E_m \rho) = \frac{1}{N} tr(E_m I) = \frac{1}{N}.
$$

Here we have assumed, without loss of generality, that there is no degeneracy.

Examples :

(i) If we pass unpolarized light through an analyzer, we get two equal intensity beams at the output irrespective of the orientation of the axes of the analyzer. This implies $\rho = \frac{1}{2}I$ for the ensemble of photons making unpolarized light.

(ii) If we prepare the ensemble of electron states

$$
\rho = \frac{1}{2}\{|\hat{\boldsymbol{\sigma}}\cdot\hat{\mathbf{n}}, 0\rangle\langle\hat{\boldsymbol{\sigma}}\cdot\hat{\mathbf{n}}, 0| + |\hat{\boldsymbol{\sigma}}\cdot\hat{\mathbf{n}}, 1\rangle\langle\hat{\boldsymbol{\sigma}}\cdot\hat{\mathbf{n}}, 1|\} = \frac{1}{2}I
\tag{68}
$$

and pass such a beam of electrons through a Stern-Gerlach magnet, we get two equal intensity beams at the output irrespective of the direction of the field gradient.

We now show that no two or more states correspond to random mixture, that is, a state which is random mixture is unique. This is because the corresponding operator is identity which is unique. The density operator for random mixture has to be proportional to identity because no other observable A satisfies $U^{\dagger}AU = A$ for *all* unitary operators U. Thus there cannot be several distinct types of random mixtures.

Further, this unique random mixture is dynamically invariant. Suppose $\rho(0) = \frac{1}{N}I$ is a random mixture. Then,

$$\rho(t) = U\rho(0)U^{\dagger} = U\left(\frac{1}{N}I\right)U^{\dagger} = \frac{1}{N}I = \rho(0),$$

which proves the point. Thus if the system is in the random mixture state at $t = 0$, then after its evolution according to its internal interactions for a finite time t, the system must again be in the same random mixture state.

8.1 Degree of mixedness/ignorance

When the density operator of a mixed state made up of N orthonormal basis states is $\frac{1}{N}I$, then the probability that the system is prepared in each of the N orthonormal basis states making up the density operator is the same and equals $\frac{1}{N}$. Thus the system is equally likely to be prepared in each of the N basis states making up the density operator. This makes us completely ignorant about exactly which of the N basis states the system is prepared. This justifies the name 'random mixture' for such a mixed state. This state is called maximally mixed and expresses a maximum degree of ignorance.

Now consider an arbitrary density operator expressed as a classical mixture of N orthonormal basis states $\{|\psi_i\rangle\}$[9] defined by the ensemble $\{|\psi_i\rangle, p_i\}$. We know that, if only one of these probabilities, say p_1 is unity and all other probabilities are zero, the density operator becomes a projector $|\psi_1\rangle\langle\psi_1|$ and the corresponding state a pure state. In this case, we know with certainty which pure state the system is in and there is no ignorance about the state of the system. If we now reduce p_1 from unity, since the normalization condition $\sum_i p_i = 1$ is to be preserved, some $p_i \neq p_1$ must become greater than zero, say $p_2 > 0$. Some degree of ignorance regarding the state of the system then prevails, as we do not know with certainty in which state out of $|\psi_1\rangle$ and $|\psi_2\rangle$ the system has been prepared. However, if, say, $p_1 = 0.9$ and $p_2 = 0.1$ then we are reasonably confident that the state $|\psi_1\rangle$ has been prepared. Accordingly, the 'degree of ignorance' or the 'degree of mixedness' is low. As we steadily increase $p_2, p_3, \ldots, p_N$ and decrease p_1, the degree of ignorance keeps increasing until it finally gets maximum at $p_1 = p_2 = p_3 \cdots = p_N = \frac{1}{N}$, with the density operator $\rho = \frac{1}{N}I$, that is, a random mixture.

---

[9]Every density operator can be expressed as a classical mixture of orthonormal pure states forming a basis in the state space, which could be its own eigenbasis.

Can we introduce a measure for the degree of mixedness of a mixed state? The answer is yes. One such measure is the so called purity of the state defined by

$$\zeta = tr(\rho^2) \tag{69}$$

If ρ is a pure state, then we know that $\zeta = 1$. Assuming that $\{|\psi_i\rangle\}$ making up ρ forms an orthonormal set, we know that

$$\zeta = \sum_{i=1}^{N} p_i^2. \tag{70}$$

This sum is bounded from below by $\frac{1}{N}$ which implies $p_i = \frac{1}{N}$, $i = 1, 2, \ldots, N$ and $\rho = \frac{1}{N}I$.

Another measure of mixedness is the so called von Neumann entropy, which is a generalization of the notion of entropy in classical statistical mechanics to the case of quantum mechanical density operators. It is given by

$$S(\rho) = -tr(\rho \log \rho) = -\sum_i \lambda_i \log_2 \lambda_i \tag{71}$$

where λ_i are the eigenvalues of the density operator ρ. Conventionally, the case $\lambda_i = 0$ is handled by defining $0 \log 0 = 0$, so that the states absent from the mixture do not contribute to entropy.

If ρ is pure, then all $\lambda_i = 0$ except for one which has to be unity. This gives $S(\rho) = 0$ which means that von Neumann entropy vanishes for a pure state. On the other hand, for a random mixture, $\lambda_i = \frac{1}{N}$ which gives

$$S(\rho) = \log_2(N) \tag{72}$$

which is the maximum value $S(\rho)$ can take. For the intermediate case of non-maximally mixed states, the value of $S(\rho)$ is correspondingly lower. For example, if the preparation procedure improves and assigns non-zero probabilities to $M < N$ states and zero probabilities to $N - M$ states we get $S(\rho) = \log_2(M)$. Thus we see that von Neumann entropy is intuitively similar to the entropy in classical setting, where it is the measure of the amount of information - or of ignorance - about the state of the system, quantified by the number of states available to the system.

9   Bloch ball and Bloch sphere

We start by making a quick revision.

- The set of all linear operators on the state space of dimension n forms a complex Hilbert space $L_{\mathcal{H}}$ of dimension n^2 with scalar product defined by Eq.(14) or Eq.(17).

- The set of all Hermitian operators is a subspace of $L_{\mathcal{H}}$ which is a real Hilbert space of dimension n^2 with the same scalar product.

- The set of all density operators (corresponding to both, mixed and pure states) is an $n^2 - 1$ dimensional subset of the space of Hermitian operators and has an isomorphic image in $\mathbb{R}^{n^2-1}$ which is a convex set.

As an important realization of the convex set representing all density operators, as described in the last item above, we consider the set of density operators acting on the two dimensional state space of a spin $\frac{1}{2}$ particle with orthonormal basis $|0\rangle$: spin up and $|1\rangle$: spin down, in the z direction, that is, $\{|0\rangle, |1\rangle\}$ is the eigenbasis of the Pauli operator $\boldsymbol{\sigma}_z \equiv \boldsymbol{\sigma}_3$.

We know that the set $\{I, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3\}$ forms an orthonormal basis of the $4 - D$ real Hilbert space comprising Hermitian operators. Therefore, all possible real linear combinations of the operators in the above basis span all Hermitian operators acting on the $2 - D$ state space. Thus we can write a general Hermitian operator acting on the state space as

$$\rho(\mathbf{P}) = \frac{1}{2}(I + \mathbf{P} \cdot \boldsymbol{\sigma}) \equiv \frac{1}{2}(I + P_1\boldsymbol{\sigma}_1 + P_2\boldsymbol{\sigma}_2 + P_3\boldsymbol{\sigma}_3) \tag{73}$$

where (P_1, P_2, P_3) is a triplet of real numbers defining a vector

$$\mathbf{P} \equiv (P_1, P_2, P_3) \tag{74}$$

in $\mathbb{R}^3$. All the Hermitian operators are spanned as this vector $\mathbf{P}$, that is, the triplet of real numbers in Eq.(74), is varied continuously in $\mathbb{R}^3$.

However, we are not interested in the whole $4 - D$ space of Hermitian operators, but only in the $3 - D$ subset of this space comprising density operators, which are positive operators having unit trace. Thus we have to find conditions on vector $\mathbf{P}$ or the triplet (P_1, P_2, P_3) occuring in the linear combination in Eq.(73), which will make the corresponding operator $\rho(\mathbf{P})$ a positive operator with unit trace. We know that the determinant of an operator is the product of its eigenvalues so that if we restrict to triplets (P_1, P_2, P_3) for which $det(\rho) \geq 0$, then the product of the eigenvalues of ρ must be non-negative. We can find the determinant of ρ by

$$det(\rho) = \langle 0|\rho|0\rangle\langle 1|\rho|1\rangle - \langle 0|\rho|1\rangle\langle 1|\rho|0\rangle \tag{75}$$

These scalar products can be evaluated by using the expression for ρ in Eq.(73) and the following definitions of the Pauli operators.

$$\begin{aligned}
\boldsymbol{\sigma}_1 &= |0\rangle\langle 1| + |1\rangle\langle 0| \\
\boldsymbol{\sigma}_2 &= -i|0\rangle\langle 1| + i|1\rangle\langle 0| \\
\boldsymbol{\sigma}_3 &= |0\rangle\langle 0| - |1\rangle\langle 1|.
\end{aligned} \tag{76}$$

The result is

$$det(\rho) = \frac{1}{4}(1 - \mathbf{P}^2). \tag{77}$$

Therefore a necessary condition for the operator to have non-negative eigenvalues is that $det(\rho) \geq 0$, or

$$\mathbf{P}^2 \leq 1. \tag{78}$$

This condition is also sufficient, because although $\mathbf{P}^2 \leq 1$ allows both the eigenvalues to be negative, that cannot be, as $tr(\rho) = 1$ and two negative numbers cannot add upto 1.

Thus we see that there is a one to one correspondence between the set of all density operators acting on the $2 - D$ state space and the set of points in $\mathbb{R}^3$ satisfying $|\mathbf{P}| \leq 1$ via Eq.(73), which we write as

$$\mathbf{P} \mapsto \rho(\mathbf{P}) \tag{79}$$

This set of points constitutes the interior and the surface of a unit sphere in $\mathbb{R}^3$. This is a closed unit ball whose boundary surface is a unit sphere. These are called Bloch ball and Bloch sphere respectively, after their inventor Felix Bloch. Note that this is a convex set.

Where do the pure states live? The pure states must satisfy $\rho^2 = \rho$. Evaluating ρ^2 we get

$$\rho^2 = \frac{1}{4}(I + \mathbf{P} \cdot \boldsymbol{\sigma})^2 = \frac{1}{4}(I + 2\mathbf{P} \cdot \boldsymbol{\sigma} + (\mathbf{P} \cdot \boldsymbol{\sigma})^2). \tag{80}$$

Therefore, $\rho^2 = \rho$ is satisfied if

$$(\mathbf{P} \cdot \boldsymbol{\sigma})^2 = I$$

or,

$$\mathbf{P}^2 I + \sum_{i \neq j} P_i P_j (\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j + \boldsymbol{\sigma}_j \boldsymbol{\sigma}_i) = I, \tag{81}$$

or,

$$(\mathbf{P} \cdot \boldsymbol{\sigma})^2 = \mathbf{P}^2 I = I, \tag{82}$$

as the last three terms in Eq.(81) vanish because Pauli operators anticommute :

$$\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j + \boldsymbol{\sigma}_j \boldsymbol{\sigma}_i = 0, \quad i \neq j. \tag{83}$$

Eq.(82) is satisfied if

$$\mathbf{P}^2 = 1. \tag{84}$$

Thus the points corresponding to pure states lie on the Bloch sphere, which is the sphere forming the boundary of the Bloch ball. This confirms our general conclusion that the image points of pure states, in the convex set imaging the set of density operators, form a subset of the set of extreme points. For the two state system we are dealing with, we have shown that the condition (84) is both, necessary and sufficient, giving a one to one correspondence between the pure states of a two level system and the Bloch sphere. The points interior to the sphere obviously correspond to mixed states.

Which density operator the center of the Bloch ball corresponds to? At the center $\mathbf{P} = \mathbf{0}$ so Eq.(73) reduces to

$$\rho(\mathbf{0}) = \frac{1}{2}I \tag{85}$$

Thus the centre of the Bloch ball corresponds to the random mixture.

From Eq.(80) and Eq.(82) we see that

$$tr(\rho^2) = \frac{1}{2} + \frac{\mathbf{P}^2}{2} \tag{86}$$

Thus we see that, as $\mathbf{P}^2$ reduces continuously from 1 to 0, $tr(\rho^2)$ reduces continuously from 1 to $\frac{1}{2}$ which corresponds to random mixture. Thus the degree of mixedness or classical ignorance (see subsection 8.1) increases continuously as we travel from boundary to the centre.

We have shown that the set of pure states of a two level system is in one to one correspondence with the points on the Bloch sphere, but it is useful to get exactly what this correspondence is. Let us denote the general pure state density operator by

$$\rho = \frac{1}{2}(I + \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}) \tag{87}$$

where $\hat{\mathbf{n}}$ is the unit vector making polar angle θ and azimuthal angle ϕ. It turns out that the states corresponding to the spin pointing up or down along $\hat{\mathbf{n}}$, denoted $|0, \theta, \phi\rangle$ and $|1, \theta, \phi\rangle$ respectively, are the eigenstates of the operator $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}$. When expanded in the $\{|0\rangle, |1\rangle\}$ basis, these states are

$$
\begin{aligned}
|0, \theta, \phi\rangle &= \cos\frac{\theta}{2} e^{-i\frac{\phi}{2}} |0\rangle + \sin\frac{\theta}{2} e^{i\frac{\phi}{2}} |1\rangle \\
|1, \theta, \phi\rangle &= \sin\frac{\theta}{2} e^{-i\frac{\phi}{2}} |0\rangle - \cos\frac{\theta}{2} e^{i\frac{\phi}{2}} |1\rangle
\end{aligned}
\tag{88}
$$

Evaluating the projector $|0, \theta, \phi\rangle\langle\phi, \theta, 0|$ we get,

$$
\begin{aligned}
|0, \theta, \phi\rangle\langle\phi, \theta, 0| &= \cos^2\frac{\theta}{2}|0\rangle\langle0| + \cos\frac{\theta}{2}\sin\frac{\theta}{2}e^{-i\phi}|0\rangle\langle1| + \cos\frac{\theta}{2}\sin\frac{\theta}{2}e^{i\phi}|1\rangle\langle0| + \sin^2\frac{\theta}{2}|1\rangle\langle1| \\
&= \frac{1}{2}(1 + \cos\theta)|0\rangle\langle0| + \frac{1}{2}\sin\theta e^{-i\phi}|0\rangle\langle1| + \frac{1}{2}\sin\theta e^{i\phi}|1\rangle\langle0| + \frac{1}{2}(1 - \cos\theta)|1\rangle\langle1| \\
&= \frac{1}{2}(|0\rangle\langle0| + |1\rangle\langle1|) + \frac{1}{2}\cos\theta(|0\rangle\langle0| - |1\rangle\langle1|) + \frac{1}{2}\sin\theta\cos\phi(|0\rangle\langle1| + |1\rangle\langle0|) \\
&\quad+ \frac{1}{2}\sin\theta\sin\phi(-i|0\rangle\langle1| + i|1\rangle\langle0|) \\
&= \frac{1}{2}(I + \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}) \tag{89}
\end{aligned}
$$

where $\hat{\mathbf{n}} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ and where we have used Eq.(8) and Eq.(76).

Thus the pure state $|0, \theta, \phi\rangle$ corresponding to the spin up along $\hat{\mathbf{n}}$ direction is the same as the density operator $\rho = \frac{1}{2}(I + \hat{\mathbf{n}} \cdot \boldsymbol{\sigma})$. By the map (79) defined in Eq.(73), the corresponding point on the Bloch sphere is the one whose position vector is $\hat{\mathbf{n}}$. Similarly, we can show that

$$|1, \theta, \phi\rangle\langle\phi, \theta, 1| = \frac{1}{2}(I - \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}), \tag{90}$$

which means that the point on the Bloch sphere corresponding to the state $|1, \theta, \phi\rangle$ is the one whose position vector is $-\hat{\mathbf{n}}$. Thus the points representing the orthonormal pair $|0, \theta, \phi\rangle$ and $|1, \theta, \phi\rangle$ are diametrically opposite on the Bloch sphere, the diameter coinciding with $\hat{\mathbf{n}}$. In particular, the points representing the eigenstates $|0\rangle, |1\rangle$ of $\boldsymbol{\sigma}_3$ are given by the north and south poles respectively on the Bloch sphere, that is, the points of intersection of the sphere and the z axis. The Bloch ball and the Bloch sphere are depicted in Fig.4.
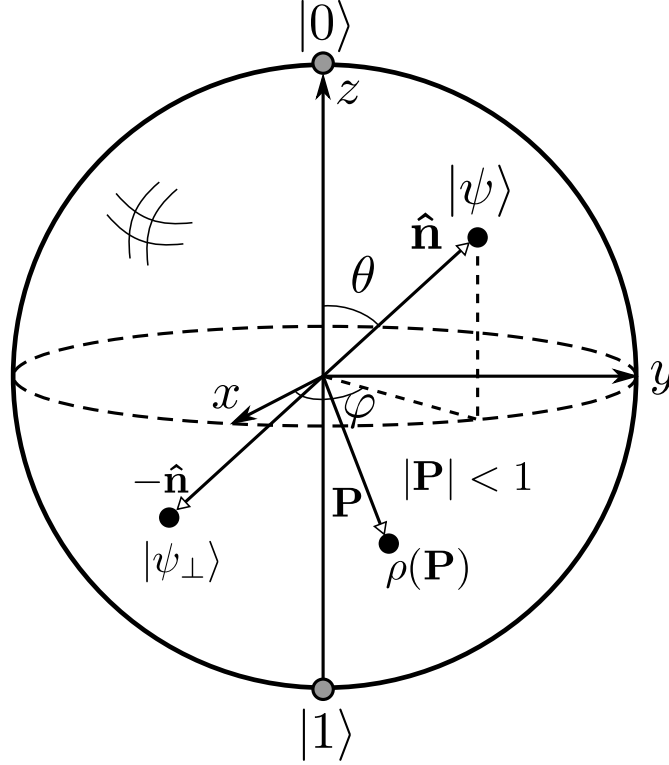

Figure 4: The Bloch ball and the Bloch sphere.

We have seen that the pure states corresponding to the spin up or spin down along $\hat{\mathbf{n}}$ direction are the eigenstates of $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}$. These states correspond to the diametrically opposite points $\pm\hat{\mathbf{n}}$ on the Bloch sphere. The values of the spin along $\pm\hat{\mathbf{n}}$ are definitely ± 1. However, when the system is in the mixed state $\rho(\mathbf{P})$, corresponding to some interior point $\mathbf{P}$ in the Bloch ball, we can speak of the average value of the spin along some direction $\hat{\mathbf{n}}$, that is, the average value of $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}$ in $\rho(\mathbf{P})$. Using the property

$$\frac{1}{2} tr \boldsymbol{\sigma}_i \boldsymbol{\sigma}_j = \delta_{ij}$$

which can easily be derived from Eq.(16), we can compute this average using Eq.(73) :

$$\langle \hat{\mathbf{n}} \cdot \boldsymbol{\sigma} \rangle_{\mathbf{P}} = tr((\hat{\mathbf{n}} \cdot \boldsymbol{\sigma})\rho(\mathbf{P})) = \hat{\mathbf{n}} \cdot \mathbf{P} \tag{91}$$

37

Thus the average value of spin along $\hat{\mathbf{n}}$ vanishes if $\hat{\mathbf{n}}$ is orthogonal to $\mathbf{P}$. If $\hat{\mathbf{n}}$ lies on a fixed cone making an angle θ with $\mathbf{P}$, then the average spin along $\hat{\mathbf{n}}$ increases monotonically with $|\mathbf{P}|$. This fact is sometimes expressed by saying that the vector $\mathbf{P}$ defining $\rho(\mathbf{P})$ polarizes the spin along $\hat{\mathbf{n}}$. If we choose $\hat{\mathbf{n}}$ as the triad of unit vectors defining the x, y, z axes, we get

$$
\begin{aligned}
\langle \hat{\mathbf{x}} \cdot \boldsymbol{\sigma} \rangle_{\mathbf{P}} &= \hat{\mathbf{x}} \cdot \mathbf{P} = P_x \\
\langle \hat{\mathbf{y}} \cdot \boldsymbol{\sigma} \rangle_{\mathbf{P}} &= \hat{\mathbf{y}} \cdot \mathbf{P} = P_y \\
\langle \hat{\mathbf{z}} \cdot \boldsymbol{\sigma} \rangle_{\mathbf{P}} &= \hat{\mathbf{z}} \cdot \mathbf{P} = P_z
\end{aligned}
\tag{92}
$$

Thus if we measure $\hat{\mathbf{x}} \cdot \boldsymbol{\sigma}$, $\hat{\mathbf{y}} \cdot \boldsymbol{\sigma}$ and $\hat{\mathbf{z}} \cdot \boldsymbol{\sigma}$ on a large number of systems identically prepared in the state $\rho(\mathbf{P})$ and compute $\langle \hat{\mathbf{x}} \cdot \boldsymbol{\sigma} \rangle$, $\langle \hat{\mathbf{y}} \cdot \boldsymbol{\sigma} \rangle$, $\langle \hat{\mathbf{z}} \cdot \boldsymbol{\sigma} \rangle$ using the measured values, then by Eq.(92) we get P_x, P_y, P_z or the vector $\mathbf{P}$ corresponding to $\rho(\mathbf{P})$. We can use this value of $\mathbf{P}$ in Eq.(73) to determine the density operator $\rho(\mathbf{P})$. Thus we get an experimental method to find out the state of the system.

10 Matrices for the linear operators

Consider a linear operator A on an n dimensional state space $\mathcal{H}$ and a basis $\{|\phi_k\rangle\} \in \mathcal{H}$. By $A|\phi_k\rangle$ we mean the ket obtained by acting A on $|\phi_k\rangle$. This ket can be expanded in the basis $\{|\phi_k\rangle\}$. Expanding each of the n kets $A|\phi_k\rangle$; $k = 1, \ldots, n$ in the basis $\{|\phi_k\rangle\}$, we get the following system of n equations

$$
A|\phi_k\rangle = \sum_j a_{jk}|\phi_j\rangle \quad ; k = 1, \ldots, n.
\tag{93}
$$

There are n coefficients a_{jk}, $k = 1, \ldots, n$ in kth equation and $k = 1, \ldots, n$ is the index spanning n equations, so the coefficients a_{jk} can be arranged in a $n \times n$ square matrix

$$
A \leftrightarrow \begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} \\
\vdots & \vdots & \vdots & \vdots \\
a_{k1} & a_{k2} & \ldots & a_{kn} \\
\vdots & \vdots & \vdots & \vdots \\
a_{n1} & a_{n2} & \ldots & a_{nn}
\end{bmatrix}.
$$

The matrix $[a_{jk}]$ defined above is called the matrix representing the operator A in the $\{|\phi_k\rangle\}$ basis. We can use it to evaluate the action of A on an arbitrary state $|\psi\rangle$ in the following way. Let the expansion of $|\psi\rangle$ in the basis $\{|\phi_k\rangle\}$ be

$$
|\psi\rangle = \sum_{k=1}^{n} x_k |\phi_k\rangle.
\tag{94}
$$

and let the expansion of the ket $A|\psi\rangle$ in the $\{|\phi_k\rangle\}$ basis be

$$A|\psi\rangle = \sum_{j=1}^{n} y_j |\phi_j\rangle. \qquad (95)$$

Acting by A on both sides of Eq.(94) and using linearity of A we get

$$A|\psi\rangle = \sum_{k=1}^{n} x_k A|\phi_k\rangle. \qquad (96)$$

Expanding the kets $A|\phi_k\rangle$ in terms of the basis $\{|\phi_k\rangle\}$, as in Eq.(93) and substituting in Eq.(96), we get,

$$A|\psi\rangle = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_{jk} x_k \right) |\phi_j\rangle. \qquad (97)$$

Comparison with Eq.(95) immediately gives the system of n equations

$$y_j = \sum_{k=1}^{n} a_{jk} x_k \quad j = 1, \ldots, n. \qquad (98)$$

Written in matrix form, this system of equations becomes

$$
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}
=
\begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} \\
\vdots & \vdots & \vdots & \vdots \\
a_{j1} & a_{j2} & \ldots & a_{jn} \\
\vdots & \vdots & \vdots & \vdots \\
a_{n1} & a_{n2} & \ldots & a_{nn}
\end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},
$$

where $[a_{jk}]$ is the matrix representing the operator A in the basis $\{|\phi_k\rangle\}$, $\{x_k, \ k = 1, \ldots, n\}$ are the coefficients of expansion of $|\psi\rangle$ in the basis $\{|\phi_k\rangle\}$ and $\{y_j, \ j = 1, \ldots, n\}$ are the coefficients of the expansion of the ket $A|\psi\rangle$ in the same basis. Since this matrix equation determines the coefficients $\{y_j\}$ which determine $A|\psi\rangle$, the action of A on $|\psi\rangle$ is completely determined by the action of matrix $[a_{jk}]$ on the column vector comprising $\{x_k, \ k = 1, \ldots, n\}$, which are the coefficients of expansion of $|\psi\rangle$ in the $\{|\phi_k\rangle\}$ basis. Note that all quantities refer to the same basis.

Thus if we fix a basis in the state space $\mathcal{H}$, we can establish a one to one and onto correspondence between the set of linear operators on $\mathcal{H}$ and the set of $n \times n$ (complex) matrices $[a_{ij}]$ in the following way. Given an operator A we know how to set up its matrix using the fixed basis. This matrix representing the operator is unique as long as the basis is fixed. Given an $n \times n$ complex matrix, we can obtain its action on each of the basis vectors $|\phi\rangle_k$ via Eq.(93), which, when read from right to left, defines an operator on $\mathcal{H}$. Thus every $n \times n$ complex matrix has a unique operator as its pre-image as long as all

expansions are referred to a fixed basis. Thus the correspondence between the operators and matrices is onto. It is straightforward to show that, given two linear operators A and B, and their respective matrices $[a_{ij}]$ and $[b_{ij}]$ with respect to a fixed basis, then the matrix representing the operator $\alpha A + \beta B$ (α, β complex numbers) with respect to the same basis, is $\alpha[a_{ij}] + \beta[b_{ij}]$. This shows that, for a fixed basis, there is an isomorphism between the set of linear operators on $\mathcal{H}$ and the set of complex $n \times n$ matrices. Setting up the matrices for linear operators using a different basis will generate a new isomorphism, because different matrix will represent the same operator through Eq.(93). The only exception is the identity operator : $a_{jk} = \delta_{jk}$ in all bases for the identity, as can be seen from Eq.(93). Thus identity operator is always represented by a unit matrix.

If the basis $\{|\phi_k\rangle, k = 1, \ldots, n\}$ is an orthonormal basis, then

$$\langle \phi_j | \{A | \phi_k \rangle\} = \langle \phi_j | \{\sum_{i=1}^{n} a_{ik} | \phi_i \rangle\} = \sum_{i=1}^{n} a_{ik} \langle \phi_j | \phi_i \rangle = \sum_{i=1}^{n} a_{ik} \delta_{ji} = a_{jk}. \tag{99}$$

Thus the jkth matrix element a_{jk} is the scalar product of jth eigenvector with $A|\phi_k\rangle$, provided the basis $\{|\phi_k\rangle\}$ is orthonormal. In the context of matrices representing linear operators, by a basis we mean an orthonormal basis.

Written explicitly, the matrix of an operator A in the basis $\{|\phi_k\rangle\}$ looks like

$$[A] = \begin{bmatrix} \langle \phi_1 | A | \phi_1 \rangle & \langle \phi_1 | A | \phi_2 \rangle & \ldots & \langle \phi_1 | A | \phi_n \rangle \\ \langle \phi_2 | A | \phi_1 \rangle & \langle \phi_2 | A | \phi_2 \rangle & \ldots & \langle \phi_2 | A | \phi_n \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle \phi_n | A | \phi_1 \rangle & \langle \phi_n | A | \phi_2 \rangle & \ldots & \langle \phi_n | A | \phi_n \rangle \end{bmatrix}.$$

Mnemonic : Elements of lth column are the components of $A|\phi_l\rangle$ in the basis $\{|\phi_k\rangle\}$.

As an example, consider the operator $R\left(\frac{\pi}{2}\hat{\mathbf{i}}\right)$ in $3 - D$ Euclidean space for the counterclockwise rotation about the basis vector $\hat{\mathbf{i}}$ in the orthonormal $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ basis. Identify $|1\rangle \equiv \hat{\mathbf{i}}, |2\rangle \equiv \hat{\mathbf{j}}, |3\rangle \equiv \hat{\mathbf{k}}$. Then

$$R\left(\frac{\pi}{2}\hat{\mathbf{i}}\right)|1\rangle = |1\rangle \equiv 1\ 0\ 0$$

$$R\left(\frac{\pi}{2}\hat{\mathbf{i}}\right)|2\rangle = |3\rangle \equiv 0\ 0\ 1$$

$$R\left(\frac{\pi}{2}\hat{\mathbf{i}}\right)|3\rangle = -|2\rangle \equiv 0\ -1\ 0$$

where the triplets on the right are the coefficients in the expansion of the transformed vectors in the $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ basis. Using the mnemonic, the matrix for $R\left(\frac{\pi}{2}\hat{\mathbf{i}}\right)$ is obtained by arranging these coefficients columnwise.

$$R\left(\frac{\pi}{2}\hat{\mathbf{i}}\right) \leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Let A, B be two linear operators on $\mathcal{H}$. We have,

$$AB|\phi_k\rangle = A\sum_{i=1}^n b_{ik}|\phi_i\rangle = \sum_{i=1}^n b_{ik}A|\phi_i\rangle = \sum_{i=1}^n \beta_{ik}\sum_{j=1}^n a_{ji}|\phi_j\rangle = \sum_{j=1}^n \left(\sum_{i=1}^n a_{ji}b_{ik}\right)|\phi_j\rangle$$

In other words, the matrix elements of $C = AB$ are

$$c_{jk} = \sum_{i=1}^n a_{ji}b_{ik} \tag{100}$$

or, the matrix for the product $C = AB$ is given by the product of the matrices for A and B. In general, of course, $AB \neq BA$. If A and B do not commute, their matrices also do not commute.

Using the definition of the adjoint of an operator, we see that the elements of the matrix representing the adjoint operator must satisfy

$$a_{jk}^\dagger = \langle\phi_j|A^\dagger|\phi_k\rangle = \langle\phi_k|A|\phi_j\rangle^* = a_{kj}^*. \tag{101}$$

So the jkth element of the $A^\dagger$ matrix equals complex conjugate of the kjth element of the A matrix. If the operator is Hermitian or self adjoint, then $A^\dagger = A$ so that, $a_{jk}^\dagger = a_{jk} = a_{kj}^*$. This also implies that the diagonal elements are real : $a_{jj} = a_{jj}^*$. The corresponding matrix is also called Hermitian. The matrix representing a Hermitian operator A with respect to its orthonormal eigenbasis $\{|u_k\rangle\}$ is diagonal with its eigenvalues $\{\lambda_k\}$ on the diagonal. This is because

$$\langle u_l|A|u_k\rangle = \lambda_k\langle u_l|u_k\rangle = \lambda_k\delta_{kl}.$$

The matrix representing a ket comprises the coefficients ocurring in its expansion in the given orthonormal basis which are arranged as a column vector, that is, a $n \times 1$ matrix. The column matrix representing a ket $|a\rangle$ in the orthonormal basis $\{|\phi_k\rangle\}$ is

$$|a\rangle \leftrightarrow \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

where $a_k = \langle\phi_k|a\rangle$, $k = 1, \ldots, n$ are defined via

$$|a\rangle = \sum_k a_k|\phi_k\rangle = \sum_k |\phi_k\rangle(\langle\phi_k|a\rangle).$$

The matrix representing a bra $\langle a|$ is obtained via

$$1 = \langle a|a\rangle = \sum_k a_k^* a_k = \begin{bmatrix} a_1^* & a_2^* & \ldots & a_n^* \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Thus the matrix representing $\langle a |$ is the row matrix

$$\langle a | \leftrightarrow \begin{bmatrix} a_1^* & a_2^* & \dots & a_n^* \end{bmatrix}.$$

11   Density matrix

The matrix representing a density operator ρ in a given orthonormal basis is called density matrix. Since ρ is hermitian, its density matrix must also be Hermitian.

Consider the density operator

$$\rho = \sum_{i=1}^{m} p_i |\psi_i\rangle \langle \psi_i| \tag{102}$$

corresponding to the ensemble $\{p_i, |\psi_i\rangle \ i = 1, \dots, m\}$. We require $\{|\psi_i\rangle\}$ to be normalized but they need not be orthogonal. Consider the density matrix of ρ in an orthonormal basis $\{|u_j\rangle, \ j = 1, \dots, n\}$. Its diagonal elements ρ_{kk} have the form

$$\rho_{kk} = \langle u_k|\rho|u_k\rangle = \sum_{i=1}^{m} p_i \langle u_k|\psi_i\rangle \langle \psi_i|u_k\rangle = \sum_{i=1}^{m} p_i |\langle u_k|\psi_i\rangle|^2, \tag{103}$$

which shows that ρ_{kk} is non-negative. We know that the last expression is the probability of the state $|u_k\rangle$ getting prepared as a result of the measurement to the basis $\{|u_j\rangle, \ j = 1, \dots, n\}$ carried out on the state ρ. Another way of looking at this is

$$tr(E_k\rho) = \sum_{l=1}^{n} \langle u_l|\rho E_k|u_l\rangle = \sum_{l=1}^{n} \langle u_l|\rho|u_k\rangle \langle u_k|u_l\rangle = \sum_{l=1}^{n} \langle u_l|\rho|u_k\rangle \delta_{kl} = \langle u_k|\rho|u_k\rangle = \rho_{kk}, \tag{104}$$

where we have used the orthonormality of the basis : $\langle u_k|u_l\rangle = \delta_{kl}$. We know that $tr(E_k\rho)$ is the probability of preparing the state $|u_k\rangle$ as a result of a measurement to the orthonormal basis $\{|u_j\rangle, \ j = 1, \dots, n\}$ on the state ρ. Thus the diagonal elements ρ_{kk} give this probability. If we carry out the same measurement on a large number, say N, of systems in state ρ, then $N\rho_{kk}$ systems will be found in the state $|u_k\rangle$ just after the measurement. For this reason, ρ_{kk} is often called the population of the state $|u_k\rangle$.

The off diagonal elements are

$$\rho_{kl} = \langle u_k|\rho|u_l\rangle = \sum_{i=1}^{m} p_i \langle u_k|\psi_i\rangle \langle \psi_i|u_l\rangle. \tag{105}$$

We know that these amplitudes appear in the coherent superposition giving $|\psi_i\rangle$ in the $\{|u_j\rangle\}$ basis

$$|\psi_i\rangle = \sum_{j=1}^{n} |u_j\rangle \langle u_j|\psi_i\rangle$$

and produce quantum interference effects which show up via relative phases between various terms in the coherent superposition. As per Eq.(105), ρ_{kl} is the average of the products like $\langle u_k|\psi_i\rangle\langle u_l|\psi_i\rangle^*$ taken over the states $|\psi_i\rangle$ which make ρ. Note that ρ_{kl} can be zero even if none of the products $\langle u_k|\psi_i\rangle\langle u_l|\psi_i\rangle^*$ is zero. In contrast ρ_{kk} can vanish only when each term in Eq.(103) vanishes separately. $\rho_{kl} = 0$ implies that the averaging in Eq.(105) has canceled out any interference effects between $|u_k\rangle$ and $|u_l\rangle$. This is why off diagonal elements of ρ are called coherences.

As is true for any linear operator acting on the state space, the matrix representing a density operator depends on the orthonormal basis used to set it up. This means that the populations and the coherences depend on the basis $\{|u_j\rangle\}$ used to set up the density matrix. In particular, if we choose the orthonormal eigenbasis of ρ to set up the density matrix, all the off diagonal elements vanish, that is, all coherences vanish. The eigenvalues p_k on the diagonal are the populations.

If the basis $\{|u_j\rangle\}$ used to set up the density matrix is the eigenbasis of the Hamiltonian H of the system, so that $\{|u_j\rangle\}$ are the energy eigenkets, we can find the evolution of the diagonal elements ρ_{kk} by sandwiching the equation of motion within $|u_k\rangle$. We get

$$i\hbar\frac{d}{dt}\rho_{kk} = \langle u_k|[H,\rho]|u_k\rangle = E_k(\rho_{kk} - \rho_{kk}) = 0, \tag{106}$$

or ρ_{kk} is constant in time. To get the coherences we sandwich the equation of motion between $|u_k\rangle$ and $|u_l\rangle$, $l \neq k$, so that

$$i\hbar\frac{d}{dt}\rho_{kl} = (E_k - E_l)\rho_{kl}. \tag{107}$$

Integrating this equation with initial condition $\rho_{kl}(t = 0) = \rho_{kl}(0)$, we get

$$\rho_{kl}(t) = \exp\left[\frac{-i}{\hbar}(E_k - E_l)t\right]\rho_{kl}(0). \tag{108}$$

Thus populations are constant in time and the coherences oscillate at the Bohr frequencies of the system.

Following example illustrates coherences. Consider the pure state

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \; ; \; |\alpha|^2 + |\beta|^2 = 1, \tag{109}$$

where $(|0\rangle, |1\rangle)$ is an orthonormal basis in the $2-D$ state space. The corresponding projector is

$$\rho = |\psi\rangle\langle\psi| = |\alpha|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1| + \alpha\beta^*|0\rangle\langle 1| + \alpha^*\beta|1\rangle\langle 0|. \tag{110}$$

We set up the density matrix for ρ in Eq.(110) with respect to the $(|0\rangle, |1\rangle)$ basis. We get

$$\rho_{00} = |\alpha|^2, \; \rho_{11} = |\beta|^2, \; \rho_{01} = \alpha\beta^*, \; \rho_{10} = \alpha^*\beta. \tag{111}$$

That is,

$$[\rho] = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{bmatrix},$$

with $tr[\rho] = 1$ as it should be. Since this density matrix represents a projector, it must satisfy $\rho^2 = \rho$ or $tr(\rho^2) = 1$, which we leave for you to check.

The diagonal elements of ρ, $|\alpha|^2$ and $|\beta|^2$, are the probabilities of the states $|0\rangle$ or $|1\rangle$ respectively getting prepared, as a result of measuring σ_z whose eigenbasis is $(|0\rangle, |1\rangle)$,[10] when the system is in the state $|\psi\rangle$ as in Eq.(109), or, equivalently in the state given by the projector ρ in Eq.(110). Thus the diagonal elements give the probabilities for the two mutually exclusive (that is, orthogonal) alternatives $|0\rangle$ and $|1\rangle$.

To understand the meaning of the off-diagonal elements, we consider the ensemble

$$\{(|0\rangle, |\alpha|^2), (|1\rangle, |\beta|^2)\}, \tag{112}$$

as against the pure state $|\psi\rangle$ in Eq.(109). Since the pure states $|0\rangle$ and $|1\rangle$ defining this ensemble are orthogonal and hence correspond to two mutually exclusive alternatives (the system could be either in the $|0\rangle$ state or in the $|1\rangle$ state but it cannot be in both these states. In contrast, the superposition defining the pure state $|\psi\rangle$ means that the system can be simultaneously in both the states $|0\rangle$ and $|1\rangle$.) Therefore, the mixed state generated by the ensemble (112) is expected to behave classically, that is, we do not expect it to show any quantum effects. The density operator corresponding to the ensemble (112) is

$$\rho_c = |\alpha|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1|, \tag{113}$$

with the corresponding density matrix

$$[\rho_c] = \begin{bmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{bmatrix}.$$

You can check that $\rho_c^2 \neq \rho_c$, so that this is a mixed state. ρ_c differs from ρ by the absence of the off-diagonal elements. Therefore we expect that the quantum effects in ρ could be embeded in its off-diagonal elements.

To relate the off-diagonal elements of ρ to quantum coherence, we first note that we cannot differentiate between ρ and ρ_c if we restrict ourselves to the measurement of σ_z with its $(|0\rangle, |1\rangle)$ eigenbasis, as the probabilities for detecting spin up and spin down alternatives are equal in both cases and this is all the information that can be extracted from this measurement. However, the differences between ρ and ρ_c come about when we measure σ_x with its rotated eigenbasis

$$\begin{aligned} |0, x\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ |1, x\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \end{aligned} \tag{114}$$

---

[10]We say that we are measuring to the $(|0\rangle, |1\rangle)$ basis. We will explain this phrase in detail in write-up 3.

Inverting Eq.(114) we get

$$
\begin{aligned}
|0\rangle &= \frac{1}{\sqrt{2}}(|0,x\rangle + |1,x\rangle) \\
|1\rangle &= \frac{1}{\sqrt{2}}(|0,x\rangle - |1,x\rangle)
\end{aligned}
\tag{115}
$$

Substituting Eq.(115) in the expression of $|\psi\rangle$ in Eq.(109) we get

$$
|\psi\rangle = \frac{\alpha + \beta}{\sqrt{2}}|0,x\rangle + \frac{\alpha - \beta}{\sqrt{2}}|1,x\rangle,
\tag{116}
$$

with the corresponding density matrix

$$
[\rho_x] = \frac{1}{2}\begin{bmatrix} |\alpha+\beta|^2 & (\alpha+\beta)(\alpha^*-\beta^*) \\ (\alpha^*+\beta^*)(\alpha-\beta) & |\alpha-\beta|^2 \end{bmatrix}.
$$

On the other hand, the density operator ρ_c, as in Eq.(113), gets transformed, after substituting Eq.(115) into it, to

$$
\rho_{c,x} = \frac{|\alpha|^2 + |\beta|^2}{2}|0,x\rangle\langle 0,x| + \frac{|\alpha|^2 + |\beta|^2}{2}|1,x\rangle\langle 1,x| + \frac{|\alpha|^2 - |\beta|^2}{2}(|0,x\rangle\langle 1,x| + |1,x\rangle\langle 0,x|),
\tag{117}
$$

giving the probabilities for detecting spin up or spin down along x-axis to be

$$
P(0,x) = \frac{|\alpha|^2 + |\beta|^2}{2} = P(1,x)
\tag{118}
$$

while the corresponding probabilities given by ρ_x corresponding to Eq.(116) are

$$
\begin{aligned}
P(0,x) &= \frac{|\alpha + \beta|^2}{2}, \\
P(1,x) &= \frac{|\alpha - \beta|^2}{2}.
\end{aligned}
\tag{119}
$$

The expressions for these probabilities clearly show quantum interference, or quantum coherence, as these are obtained by first combining the amplitudes and then getting the squares of their absolute values. The probabilities obtained from ρ_c (Eq.(118)) do not show such quantum interference effects. Since the quantum effects go to zero as the off-diagonal terms in ρ go to zero, (making ρ go over to the classical case ρ_c,) the off-diagonal terms in a density matrix are called quantum coherences.

Finally we note that the ensembles generating density operators we considered consisted of orthonormal pure states and the density matrices were set up using the orthonormal bases containing the states in these ensambles. In such cases the operators of the form $|\psi_1\rangle\langle\psi_2|$ occurring in the expression of the density operator do not contribute to

45

the diagonal elements of the density matrix and will not contribute to its trace. These terms will then exclusively generate only the off-diagonal elements. On the other hand, If the ensemble generating the density operator has non-orthogonal states, (eg $\rho = |\psi\rangle\langle\psi|$ with $|\psi\rangle = a|\psi_1\rangle + b|\psi_2\rangle, \langle\psi_1|\psi_2\rangle \neq 0, \{|\psi_1\rangle, |\psi_2\rangle\}$ linearly independent,) the operators of the form $|\psi_1\rangle\langle\psi_2|$ occurring in the expression of the density operator will, in general, contribute to both, the off-diagonal as well as the diagonal elements of the density matrix and will contribute $\langle\psi_2|\psi_1\rangle$ to its trace.[11]

Exercise : Set up the density operator for the ensemble $\{(|0\rangle, a)(|1, x\rangle, b)\}$ and its density matrix using the orthonormal basis $\{|0\rangle, |1\rangle\}$, Find the trace.

11.1   Density matrix in position representation

Sometimes it is useful to set up the density matrix for a density operator ρ (as in Eq.(102)) in position representation, that is, using the eigenbasis of the position operator with basis kets labeled by $x : |x\rangle$. The position operator has continuous eigenvalues x and hence x labeling the basis kets has to be treated as a continuous variable. As a result, the matrix elements have continuous indices and can be viewed as functions of these indices. We have, for the xx'th element,

$$
\begin{aligned}
\rho_{xx'} = \rho(x, x') &\equiv \langle x'|\rho|x\rangle \\
&= \sum_{i=1}^{m} p_i \langle x'|\psi_i\rangle\langle\psi_i|x\rangle \\
&= \sum_{i=1}^{m} p_i \psi_i(x')\psi_i^*(x) \quad (120)
\end{aligned}
$$

where $\psi_i(x)$ is the wave function for the state $|\psi_i\rangle$. For a pure state $|\psi\rangle$ we get

$$
\rho(x, x') = \psi(x')\psi^*(x) \quad (121)
$$

For the average value of an observable A we get

$$
\begin{aligned}
\langle A \rangle = tr(\rho A) &= \int dx \langle x|\rho A|x\rangle \\
&= \int dx \langle x|\rho \left(\int dx' |x'\rangle\langle x'| \right) A|x\rangle \\
&= \int\int dx dx' \langle x|\rho|x'\rangle\langle x'|A|x\rangle \\
&= \int\int dx dx' \rho(x, x')A(x', x), \quad (122)
\end{aligned}
$$

---

[11]Since a density operator is Hermitian, if a term like $|\psi_1\rangle\langle\psi_2|$ occurs in it, its Hermitian conjugate $|\psi_2\rangle\langle\psi_1|$ must also occur. Sum of their contribution to trace is $\langle\psi_2|\psi_1\rangle + \langle\psi_1|\psi_2\rangle$ which is a real number. Same consideration applies to the contribution of such terms to diagonal elements.

where we have used Eq.(8).

We work out the average values of the position and momentum operators. Replacing A by the position operator x we get

$$A(x', x) = \langle x'|A|x \rangle \mapsto \langle x'|x|x \rangle = x \langle x'|x \rangle = x\delta(x' - x).$$

Thus,

$$\langle x \rangle = \int \int dx dx' x \rho(x, x') \delta(x' - x) = \int dx \, x \, \rho(x, x). \tag{123}$$

To get the average momentum in the state ρ we replace A by momentum operator p. Since

$$\langle x|p|\psi \rangle = \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x|\psi \rangle,$$

we have,

$$\langle x'|p|x \rangle = \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x'|x \rangle = \frac{\hbar}{i} \frac{\partial}{\partial x} \delta(x' - x).$$

Then we get, for the average momentum,

$$\langle p \rangle = \int \int dx dx' \rho(x, x') \frac{\hbar}{i} \frac{\partial}{\partial x} \delta(x' - x).$$

Using the well known identity

$$\int \phi(x) \frac{\partial}{\partial x} \delta(x - a) dx = -\frac{\partial \phi}{\partial x}|_{x=a},$$

we get,

$$\langle p \rangle = -\frac{\hbar}{i} \int dx \left[\frac{\partial}{\partial x} \rho(x, x') \right]_{x'=x}. \tag{124}$$

12 Density operator in statistical mechanics

Here we deal with systems comprising large number of particles, of the order of Avogadro number $\approx 10^{23}$. Generally such large assemblies of particles constitute macroscopic bodies which obey laws of classical physics. However, their quantum behavior is fundamental to a large cross-section of phenomena. Thus the sea of conduction plus valence electrons in crystals stay in Bloch states which are quantum states and give rise to their quantized energy structure, that is, band structure. Energy bands in solids are basically responsible for the occurrence of metals, insulators and semiconductors. Many phenomena like heat capacities of solids, superconductivity, super fluidity, ferro and antiferromagnetism in solids, quantum phase transitions, Bose-Einstein condensation, are the consequences of quantum behavior of large assemblies of particles. The pure quantum states of such large assemblies of particles, capable of superposition, are called cat states. Cat states are extremely fragile and decohere very fast to mixed states, that is, classical mixtures of

pure states. [12] Thus the naturally occuring as well as laboratory prepared states of large quantum systems turn out to be mixed states. Thus physics of such systems is contained in their density operators.

In order to set up the density operator of such large systems, we need the corresponding ensemble : the relevant set of pure states and a probability distribution over this set of states.

We first note that our system is an open system, that is, it is interacting with its environment. We deal with this situation in two stages : we first consider the system interacting with its local environment, which we call heat bath or reservoir and then the system and the heat bath together interacting very weakly with the rest of the world. We do not precisely know about the interaction between the heat reservoir and the system, except the fact that these two can exchange energy and particles between them. When such a system and reservoir are left coupled for long enough time they together evolve to a state of thermal equilibrium, where all of its macroscopic characteristics (eg energy, volume, concentrations of atomic and molecular species, total number of particles etc) are time independent and these characteristics completely specify its state. A state of equilibrium is characterized by a constant and uniform temperature throughout the system plus reservoir. At the atomic/molecular level, the macroscopic equilibrium state is a consequence of incessant and rapid transitions of the system in equilibrium through all the atomic states consistent with all the imposed boundary conditions. If the transitions between atomic states are sufficiently fast, the system rapidly passes through all the relevant atomic states equally frequently, (that is, with equal probability), in the course of a macroscopic observation, (of macroscopic parameters like energy, volume etc), which gives rise to the equilibrium state. The equilibrium state has no record of system's previous history, that is, how the equilibrium state was reached. Thus, for our system plus heat bath configuration in equilibrium, the behavior of the system in equilibrium is independent of the details of its interaction with the heat bath. Therefore, we can choose the heat bath to have its mass and energy content far larger than that of the system. Thus if E_h and E_s are the energies of the heat bath and the system respectively, we assume that $E_h >> E_s$. Also, the smallest change in the energy of the heat bath is obtained by exciting/de-exciting a single molecule, which is very tiny compared to its total energy. Therefore, the energy levels of the heat bath are semi-continuous.

Let us take the energy states of the system, (assumed to be in thermal equilibrium with a heat bath), to be the set of states defining the density operator. As the energy of the system keeps changing, (through its energy exchange with the heat bath), states with different energies become accessible to the system. Thus we expect that the probability of the system to be in a state with energy E is a function of E. If two or more states, say $|r\rangle$ and $|s\rangle$ have equal energy, $E_r = E_s$, then these are equally probable, because a tiniest

---

[12]This does not mean that cat states cannot be produced and used in a laboratory. An example is the Superconducting QUantum Interference Device (SQUID), which can be used to superpose two quantum states associated with a macroscopic current flowing simultaneously clockwise and counterclockwise.

fluctuation can make the system go back and forth between $|r\rangle$ and $|s\rangle$. This makes the probability of a state to have energy E to be a function only of E, say $p(E)$.

The total system comprising the bath and the system has a very weak interaction with the external world. Thus its energy E_0 can fluctuate by a tiny amount Δ due to this interaction. Its energy can be anywhere in the range $E_0 \pm \Delta$ and $\Delta << E_0$ as well as $\Delta << E_h$. By energy conservation we must have

$$E_0 - \Delta \leq E_h + E_s \leq E_0 + \Delta. \tag{125}$$

Since Δ is too tiny as compared to E_0, we can treat energies in the range $E_0 \pm \Delta$ to be equal, that is, we treat the difference between these energies negligible. As a consequence, we can take the probability of the heat bath states within $E_h \pm \Delta$ to be equal, that is, we take these states to be equally likely. Inequality (125) tells us that if the reservoir is in the state with energy in the range $E_h \pm \Delta$, then the system is in a state with energy E_s. Therefore the probability $p(E_s)$ is the same as the probability of finding the reservoir in one of the states with energy in the range $E_h \pm \Delta$. Since the probabilities of all such states are equal, $p(E_s)$ is proportional to the number of ways the state of the system with energy E_s can be realized, which in turn is equal to the number of states with energy in the range $E_h \pm \Delta$. Let $\eta(E_h) = \eta(E_0 - E_s)$ be the number of states per unit energy at E_h, then $p(E_s)$ is proportional to $\eta(E_0 - E_s)(2\Delta)$. Therefore the ratio of the probabilities for states with energies E_s and $E_{s'}$ is given by

$$\begin{aligned}
\frac{p(E_s)}{p(E_{s'})} &= \frac{\eta(E_0 - E_s)}{\eta(E_0 - E_{s'})} = \exp\{\ln \eta(E_0 - E_s) - \ln \eta(E_0 - E_{s'})\} \\
&= \exp\left\{\left[\frac{\ln \eta(E_0 - E_s) - \ln \eta(E_0 - E_{s'})}{E_s - E_{s'}}\right](E_s - E_{s'})\right\} \tag{126}
\end{aligned}$$

Since both E_s and $E_{s'}$ are tiny as compared to E_0, the ratio in the square bracket can be approximated by $\frac{d}{dE}\ln \eta(E) = \beta(E)$. Assuming that $\beta(E)$ is almost constant over the range under consideration, we get

$$\frac{p(E_s)}{p(E_{s'})} = e^{-\beta(E_s - E_{s'})} \tag{127}$$

or,

$$p(E_s) \propto e^{-\beta E_s} \tag{128}$$

Since the total probability has to be unity, we must have, for the state with energy E_k,

$$p(E_k) = \frac{1}{Z}e^{-\beta E_k} \tag{129}$$

where

$$Z = \sum_k e^{-\beta E_k} \tag{130}$$

49

is called the partition function of the system.

Thus our ensemble giving rise to the density operator is $(|E_k\rangle, p_k = \frac{1}{Z}e^{-\beta E_k})$ so that the density operator is given by

$$
\begin{aligned}
\rho &= \frac{1}{Z}\sum_k e^{-\beta E_k}|E_k\rangle\langle E_k| \\
&= \frac{1}{Z}e^{-\beta H} \tag{131}
\end{aligned}
$$

where H is the Hamiltonian of the system. The last equality follows because $\sum_k e^{-\beta E_k}|E_k\rangle\langle E_k|$ is the spectral representation of the operator $e^{-\beta H}$.

The average value of an observable A pertaining to the system is given by, (see Eq.(22)),

$$
\langle A \rangle = tr(\rho A) = \frac{1}{Z}\sum_k e^{-\beta E_k}\langle E_k|A|E_k\rangle. \tag{132}
$$

This is the key equation which has to be used to get the physics out of any large system in thermal equilibrium. Since we deal only with the systems in thermal equilibrium, this equation applies to all the cases we know of.
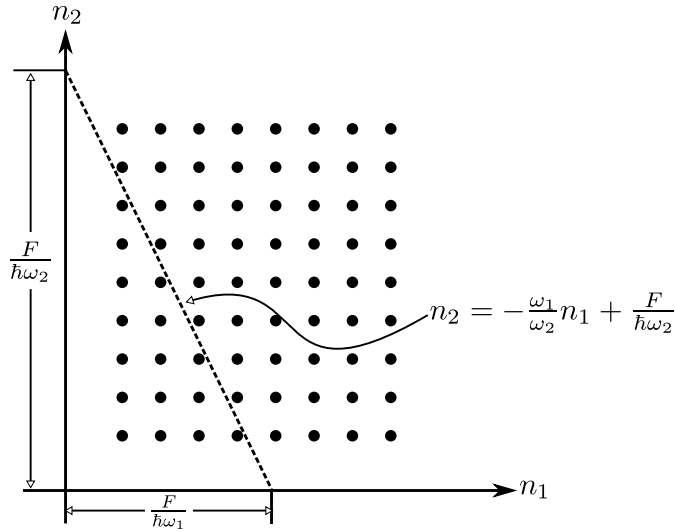

Figure 5: The number of states with energy less than F is proportional to area of the triangle below the line intersecting the axes.

We close by obtaining Eq.(126) for the case where the heat bath comprises N independent harmonic oscillators. The energy of the bath is

$$
F = \sum_{k=1}^{N} n_k \hbar\omega_k \tag{133}
$$

where we assume that n_k are very large and we neglect the zero point energy. Our aim is to get the function $\eta(F)$, that is, the number of states per unit energy at F. To get it, we first find the number of states with energy less than F. For $N = 2$, Eq.(133) is an equation to a straight line in the (n_1, n_2) plane with slope $-\frac{\omega_1}{\omega_2}$ and with intercept on the n_2 axis equal to $\frac{F}{\hbar\omega_2}$ (see Fig.5). The number of states with energy less than F is proportional to the area of the triangle formed by this line cutting the n_1 and n_2 axes. This area is proportional to F^2. Similarly, for any N the number of states with energy less than F is proportional to F^N. Therefore, number of states per unit energy is given by

$$\eta(F) \propto (d/dF)F^N \propto F^{N-1}.$$

This gives

$$\ln \eta(F) = constant + (N-1)\ln F.$$

Therefore, we get for $\beta(F)$,

$$\beta(F) = (d/dF)\ln \eta(F) \propto \frac{N-1}{F}. \tag{134}$$

Since $F = E_0 - E_s$ and $E_s << E_0$, we can replace F in Eq.(134) by E_0. Further, since the system plus reservoir are in thermal equilibrium, the energy is equally distributed over oscillators, so that we can define the energy per oscillator $W = E_0/N$ giving

$$\frac{1}{W} = \frac{1}{E_0/N} = \frac{N}{E_0} \approx \frac{N-1}{E_0}.$$

Substituting in Eq.(134) we get

$$\beta = \frac{1}{W} = \frac{1}{\text{Average energy per oscillator}}$$

which is constant independent of F. Thus we finally get, with $E_0 = NW$,

$$
\begin{aligned}
\frac{p(E_s)}{p(E_{s'})} &= \frac{\eta(E_0 - E_s)}{\eta(E_0 - E_{s'})} = \frac{(E_0 - E_s)^{N-1}}{(E_0 - E_{s'})^{N-1}} \\
&= \frac{(1 - E_s/(NW))^{N-1}}{(1 - E_{s'}/(NW))^{N-1}} \to \frac{e^{-E_s/W}}{e^{-E_{s'}/W}} = \frac{e^{-\beta E_s}}{e^{-\beta E_{s'}}}
\end{aligned} \tag{135}
$$

and we are done.

Statistical mechanics is an all-important branch of physics and deserves to be developed as a separate course.

Acknowledgments

13 Appendix A : Homomorphism and Isomorphism

We need two algebraic concepts, namely, homomorphism and its special case isomorphism between two sets, which we now define. Consider two pairs $(S_1, \circ)$ and $(S_2, \times)$, where S_1, S_2 are sets and $\circ, \times$ are binary operations on S_1 and S_2 respectively. We assume that S_1, S_2 are closed under the corresponding binary operations. Let $\varphi : S_1 \mapsto S_2$ be a map from S_1 to S_2 such that for every $a \in S_1$ there is a $\varphi(a) \in S_2$. We say that φ is a *homomorphism* if, for every $a, b \in S_1$

$$\varphi(a \circ b) = \varphi(a) \times \varphi(b). \tag{136}$$

In other words, the image of the product of a and b in S_1 is the product of their images $\varphi(a)$ and $\varphi(b)$ in S_2.

Example

Consider $(\mathbb{Z}, +)$ and $((1, -1), \cdot)$ where $\mathbb{Z}$ is the set of integers and $+$ is the usual addition on it, while $\cdot$ is the usual multiplication on the two element set $(1, -1)$. Define a map φ by

$$\varphi(n) = (-1)^n,$$

that is,

$$(-1)^{a+b} = (-1)^a \cdot (-1)^b$$

which is clearly true.

If the map φ defining a homomorphism is also one to one and onto, then it is called *isomorphism*.

Example : The set $\mathbb{Z}_2$ of integers modulo 2 and $((1, -1), \cdot)$ defined above, are isomorphic.

14 Appendix B : A Theorem About Normal Operators

We prove the following statement.

A normal operator acting on a linear space $\mathcal{H}$ has an eigenbasis in $\mathcal{H}$. That is, every normal operator is diagonal with respect to some orthonormal basis of $\mathcal{H}$. Conversely, every operator on $\mathcal{H}$ having an orthonormal eigenbasis is normal.

We follow the proof given in [1] but make it much more explicit.

We prove the forward implication (normality implies orthonormal eigenbasis) by induction on the dimension m of $\mathcal{H}$.

For $m = 1$ the statement is trivially true because any normalized ket can be taken to be a basis and the action of a linear operator just multiplies it by a scalar, that is, a complex number. Further, orthogonality condition drops out as there is only one basis vector.

Now assume that for some $m > 1$, every normal operator acting on the space with dimension $\leq m$ has an orthonormal eigenbasis. Under this assumption, we want to show that a normal operator acting on a space of dimension $m+1$ has an orthonormal eigenbasis.

In what follows, $|\psi\rangle$ is an arbitrary state in $\mathcal{H}$.

Let A be a normal operator on a $m + 1$ dimensional space. Let λ be an eigenvalue of A and let P be the projector on the eigenspace of λ, or the $\lambda-$eigenspace. Let $Q = I - P$ be the projector on the orthogonal complement of the $\lambda-$eigenspace. Now consider

$$A = (P + Q)A(P + Q) = PAP + QAP + PAQ + QAQ. \tag{137}$$

For an arbitrary ket $|\psi\rangle$, $P|\psi\rangle$ is in $\lambda-$eigenspace so that

$$PAP|\psi\rangle = \lambda P^2|\psi\rangle = \lambda P|\psi\rangle.$$

Therefore, we get the operator equality

$$PAP = \lambda P.$$

Similarly,

$$QAP|\psi\rangle = \lambda QP|\psi\rangle = 0,$$

as Q and P are orthogonal projectors. This gives the operator equality

$$QAP = 0.$$

Since $P|\psi\rangle$ is in $\lambda-$eigenspace, we get

$$AA^\dagger P|\psi\rangle = A^\dagger AP|\psi\rangle = \lambda A^\dagger P|\psi\rangle.$$

so that $A^\dagger P|\psi\rangle$ is in $\lambda-$eigenspace. This leads to $QA^\dagger P|\psi\rangle = 0$, or

$$QA^\dagger P = 0.$$

Taking adjoints we get

$$PAQ = 0.$$

Thus two terms in Eq.(137) vanish and it becomes

$$A = PAP + QAQ. \tag{138}$$

Next we prove that QAQ is a normal operator. First note that

$$QA = QA(P + Q) = QAQ$$

and

$$QA^\dagger = QA^\dagger(P + Q) = QA^\dagger Q$$

because $QA^\dagger P = 0$. Now by normality of A and the observation that $Q^2 = Q$ we get, using two previous equations,

$$(QAQ)(QA^\dagger Q) = QAQA^\dagger Q = QAA^\dagger Q = QA^\dagger AQ = QA^\dagger QAQ = (QA^\dagger Q)(QAQ)$$

which just means that QAQ commutes with its adjoint so that it is a normal operator.

Let us call the subspace orthogonal to $\lambda-$eigenspace projected out by Q to be S_Q. The dimension of S_Q must be $\leq m$ because the dimension of the parent space $\mathcal{H}$ is $m+1$ and the dimension of the $\lambda-$eigenspace (which is orthogonal to S_Q) is at least unity. Thus QAQ is a normal operator acting on S_Q whose dimension is $\leq m$. Hence by induction hypothesis, the normal operator QAQ must have an orthonormal eigenbasis in S_Q. Since all kets in the $\lambda-$eigenspace are eigenkets of PAP, an orthonormal basis in $\lambda-$eigenspace is an orthonormal eigenbasis of PAP in $\lambda-$eigenspace. Thus the first term in Eq.(138) has an orthonormal eigenbasis in the $\lambda-$eigenspace while the second term has an orthonormal eigenbasis in its orthogonal complement S_Q in the $m+1$ dimensional space $\mathcal{H}$. This means that the normal operator A has an orthonormal eigenbasis in the $m+1$ dimensional space $\mathcal{H}$ and we are done.

To prove the converse, solve the following problem.

Let A be an operator on a $m-$dimensional space $\mathcal{H}$ having orthonormal eigenbasis $\{|k\rangle\}$. Then for an arbitrary $|\psi\rangle = \sum_k c_k |k\rangle \in \mathcal{H}$, show that

$$A^\dagger A |\psi\rangle = \sum_k |\lambda_k|^2 c_k |k\rangle = AA^\dagger |\psi\rangle,$$

where λ_k is the eigenvalue of A in the eigenstate $|k\rangle$.

Hint : Use Eq.(8).

References

[1] Nielsen M.A. and I.L.Chuang, Quantum Computation and Quantum Information, Cambridge University Press (2000).