

## Mixed states and pure states

(Dated: April 9, 2009)

These are brief notes on the abstract formalism of quantum mechanics. They will introduce the concepts of *pure* and *mixed* quantum states. Some statements are indicated by a **P**. You should try and prove these statements. If you understand the formalism, then these statements should not be hard to prove; they are good tests for your understanding. The homework will contain more difficult questions.

1. A pure state of a quantum system is denoted by a vector (ket)  $|\psi\rangle$  with unit length, i.e.  $\langle\psi|\psi\rangle = 1$ , in a complex Hilbert space  $H$ . Previously, we (and the textbook) just called this a ‘state’, but now we call it a ‘pure’ state to distinguish it from a more general type of quantum states (‘mixed’ states, see step 21).
2. We already know from the textbook that we can define dual vectors (bra)  $\langle\phi|$  as linear maps from the Hilbert space  $H$  to the field  $C$  of complex numbers. Formally, we write

$$\langle\phi|(|\psi\rangle) = \langle\phi|\psi\rangle.$$

The object on the right-hand side denotes the inner product in  $H$  for two vectors  $|\phi\rangle$  and  $|\psi\rangle$ . That notation for the inner product used to be just that, notation. Now that we have defined  $\langle\phi|$  as a dual vector it has acquired a second meaning.

3. Given vectors and dual vectors we can define operators (i.e., maps from  $H$  to  $H$ ) of the form

$$\hat{O} = |\psi\rangle\langle\phi|.$$

$\hat{O}$  acts on vectors in  $H$  and produces as result vectors in  $H$  (**P**).

4. The hermitian conjugate of this operator is

$$\hat{O}^\dagger \equiv |\phi\rangle\langle\psi|.$$

This follows (**P**) straight from the definition of the hermitian conjugate:

$$(\langle m|\hat{O}|n\rangle)^* = \langle n|\hat{O}^\dagger|m\rangle,$$

for all states  $|n\rangle$  and  $|m\rangle$  in  $H$ .

5. A special case of such an operator is

$$\hat{P}_\psi = |\psi\rangle\langle\psi|.$$

It is hermitian and it satisfies **(P)**

$$\hat{P}_\psi^2 = \hat{P}_\psi.$$

That is, it is a projection operator, or projector.

6. Completeness of a basis  $\{|n\rangle\}$  of  $H$  can be expressed as

$$\hat{I} = \sum_n |n\rangle\langle n|,$$

where  $\hat{I}$  is the identity operator on  $H$ . ‘Inserting the identity’ is a useful trick. We’ll use it several times in the following, and it may be useful for proving some of the **P** statements. [If  $n$  is a continuous variable, we should integrate over  $n$  rather than sum. In the following we’ll keep using sums to keep notation simple, except when we’re discussing position or momentum.]

7. It is useful to define the ‘Trace’ operation:

$$\text{Tr}(\hat{K}) = \sum_n \langle n|\hat{K}|n\rangle,$$

where  $\hat{K}$  is an arbitrary operator, and the sum is over a set of basis vectors  $\{|n\rangle\}$ . If we write down a matrix representation for  $\hat{O}$ , i.e., a matrix with elements  $\langle n|\hat{O}|m\rangle$ , then the Trace is the sum over all diagonal elements (i.e., with  $m = n$ ).

8. A nice property of the Trace operation is that a basis change leaves it invariant (**P**): that is, it does not matter which basis we choose in the definition (step 7) of Trace. Indeed, the Trace would be far less useful if it did depend on the basis chosen.

9. Another property of the Trace is the ‘cyclical’ property. For example, for the Trace of three arbitrary operators (which do not necessarily commute!) we have **(P)**

$$\text{Tr}(\hat{A}\hat{B}\hat{C}) = \text{Tr}(\hat{B}\hat{C}\hat{A}) = \text{Tr}(\hat{C}\hat{A}\hat{B}).$$

Generalizations to any number of operators are obvious. In an infinite-dimensional Hilbert space some care has to be taken . . . see homework!

10. On the other hand, in general we have

$$\text{Tr}(\hat{A}\hat{B}\hat{C}) \neq \text{Tr}(\hat{C}\hat{B}\hat{A}).$$

11. Expectation values can be expressed in terms of projectors  $\hat{P}_\psi$  rather than in terms of state vectors  $|\psi\rangle$ . Namely, by inserting the identity  $\hat{I} = \sum_n |n\rangle\langle n|$  we find that for any operator  $\hat{O}$  we can write (**P**)

$$\langle \hat{O} \rangle_\psi = \langle \psi | \hat{O} | \psi \rangle = \text{Tr}(\hat{O}\hat{P}_\psi) = \text{Tr}(\hat{P}_\psi\hat{O}).$$

12. Similarly, we can express probabilities and overlaps in terms of projectors (**P**):

$$|\langle \phi | \psi \rangle|^2 = \text{Tr}(\hat{P}_\psi |\phi\rangle\langle \phi|) = \text{Tr}(|\phi\rangle\langle \phi| \hat{P}_\psi).$$

13. And so we might as well use the projector  $\hat{P}_\psi$  to describe all physical quantities we can derive from the state  $|\psi\rangle$ . We use the symbol  $\hat{\rho}$  to indicate (or emphasize) we're talking about a physical state rather than an arbitrary operator

$$\hat{\rho} = |\psi\rangle\langle \psi|,$$

and we call  $\hat{\rho}$  the density matrix, or the density operator describing the state  $|\psi\rangle$ .

14. We always have the normalization condition (**P**)

$$\text{Tr}(\hat{\rho}) = 1.$$

15. For example, if we take the two pure states

$$|\psi_\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}},$$

then the corresponding  $\hat{\rho}$ s are 2x2 matrices. Written in the basis  $\{|0\rangle, |1\rangle\}$ , they have the form

$$\hat{\rho}_\pm = \begin{pmatrix} 1/2 & \pm 1/2 \\ \pm 1/2 & 1/2 \end{pmatrix}.$$

16. The notation in step 15 is borrowed from quantum information theory: there, instead of considering classical bits, which take values 0 and 1, we consider qubits, which can be in any superposition of  $|0\rangle$  and  $|1\rangle$ . You could think of these states as spin up and spin down,  $|\uparrow\rangle$  and  $|\downarrow\rangle$ , of a spin-1/2 particle, as an example.

17. For another example, consider a particle moving in 1-D. Its state  $|\psi\rangle$  lives in an infinite-dimensional Hilbert space. Suppose we, as usual, define the wavefunction  $\psi(x)$  by

$$\langle x|\psi\rangle = \psi(x).$$

Then the corresponding density matrix in the basis  $\{|x\rangle\}, x \in (-\infty, +\infty)$ , is

$$\hat{\rho} = |\psi\rangle\langle\psi| = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' |x'\rangle\langle x'| |\psi\rangle\langle\psi| |x\rangle\langle x|$$

where we inserted two identities. Rearranging terms gives

$$\hat{\rho} = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \psi^*(x) \psi(x') |x'\rangle\langle x|$$

We can denote the matrix elements of  $\hat{\rho}$  as

$$\langle x'|\hat{\rho}|x\rangle \equiv \rho(x', x) \equiv \rho_{x'x} = \psi(x')\psi^*(x).$$

18. On the ‘diagonal’ of the matrix  $\rho(x', x)$ , where  $x = x'$ , we get the usual probability density  $|\psi(x)|^2$ . We also have

$$\text{Tr}(\hat{\rho}) = \int dx |\psi(x)|^2 = 1,$$

which expresses normalization in old and new ways.

19. Similarly, we can use momentum eigenstates and expand the *same* matrix in the form **(P)**

$$\hat{\rho} = \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dp' \tilde{\psi}^*(p) \tilde{\psi}(p') |p'\rangle\langle p|$$

where  $\tilde{\psi}(p) = \langle p|\psi\rangle$ .

20. Here is one advantage a density operator has compared to a ket: a given physical state can be described by any ket of the form  $\exp(i\theta)|\psi\rangle$  with  $\theta$  an arbitrary phase, but by only *one* density matrix  $\hat{\rho}$ . This is more economical, to say the least.
21. Now let us define a more general type of states, still described by density operators and keeping the advantage of step 20, by introducing ‘mixtures’ of pure states:

$$\hat{\rho} = \sum_{k=1}^N p_k |\psi_k\rangle\langle\psi_k|,$$

where  $\{|\psi_k\rangle\}$  is some set of pure states, not necessarily orthogonal. The number  $N$  could be anything, and is not limited by the dimension of the Hilbert space. The  $N$  numbers (or ‘weights’)  $p_k$  are nonzero and satisfy the relations

$$0 < p_k \leq 1; \quad \sum_{k=1}^N p_k = 1.$$

The normalization of the weights  $p_k$  expresses the condition  $\text{Tr}(\hat{\rho}) = 1$  (**P**).

22. We could interpret the weights  $p_k$  as probabilities, but we have to be careful: we should not think of  $p_k$  as the probability to find the particle in the state  $|\psi_k\rangle$ ! You can give one reason for that right now (**P**), and soon we will see another reason.

23. The quantum state described by  $\hat{\rho}$  is called a mixed state whenever  $\hat{\rho}$  cannot be written as a density matrix for a pure state (for which  $N = 1$  and  $p_1 = 1$ ).

24. An example: from statistical physics you may know the following statistical mixture of energy eigenstates  $|\psi_n\rangle$  in thermal equilibrium:

$$\hat{\rho} = \sum_n p_n |\psi_n\rangle \langle \psi_n|,$$

where  $p_n = \exp(-E_n/kT)/Z$  with  $Z = \sum_n \exp(-E_n/kT)$  the partition function. When the Hamiltonian does not depend on time, this mixture is time-independent (**P**).

25. Expectation values in mixed states are probabilistic weighted averages of the expectation values of the pure states:

$$\text{Tr}(\hat{\rho}\hat{O}) = \sum_k p_k \text{Tr}(|\psi_k\rangle \langle \psi_k| \hat{O}).$$

That is, in the pure state  $|\psi_k\rangle$  we would expect an average value of  $O_k \equiv \text{Tr}(|\psi_k\rangle \langle \psi_k| \hat{O})$  if we measured  $\hat{O}$ , and for the mixed state we simply get  $\sum_k p_k O_k$ .

26. Since  $\hat{\rho}$  is hermitian, we can diagonalize it, such that

$$\hat{\rho} = \sum_{k=1}^M \lambda_k |\phi_k\rangle \langle \phi_k|,$$

where the states  $|\phi_k\rangle$  are orthogonal (unlike those used in step 21). The numbers  $\lambda_k$  satisfy

$$0 \leq \lambda_k \leq 1; \quad \sum_{k=1}^M \lambda_k = 1.$$

The numbers  $\lambda_k$  are, in fact, nothing but the eigenvalues of  $\hat{\rho}$ . They sum to one because of normalization. There are exactly  $M = d$  of these numbers, where  $d$  is the dimension of the Hilbert space. In contrast, in step 21,  $N$  could be any number.

27. Comparing steps 21 and 26 shows that a given mixed (not pure) density matrix can be written in multiple [infinitely many ways, in fact] ways as probabilistic mixtures of pure states. And that is another reason why we have to be careful interpreting coefficients  $\lambda_k$  or  $p_k$  as probabilities.

28. On the other hand, we can certainly *prepare* a mixed state in a probabilistic way. If we prepare with probability  $p_k$  a pure state  $|\psi_k\rangle$ , and then forget which pure state we prepared, the resulting mixed state is  $\hat{\rho} = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ . In this case,  $p_k$  certainly has the meaning of probability.

29. For example, consider the following mixture: with probability 1/2 we prepare  $|0\rangle$  and with probability 1/2 we prepare  $(|0\rangle + |1\rangle)/\sqrt{2}$ . The mixture can be represented as a 2x2 matrix in the basis  $\{|0\rangle, |1\rangle\}$ :

$$\hat{\rho} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}.$$

The eigenvalues and eigenvectors of this matrix are

$$\lambda_{\pm} = 1/2 \pm \sqrt{2}/4,$$

and

$$|\phi_{\pm}\rangle = \sqrt{\lambda_{\pm}}|0\rangle \mp \sqrt{1 - \lambda_{\pm}}|1\rangle.$$

And so we can also view the mixed state as a mixture of the two eigenstates  $|\phi_{\pm}\rangle$  with weights equal to  $\lambda_{\pm}$ .

30. There are two simple tests to determine whether  $\hat{\rho}$  describes a pure state or not:

$$\text{pure state : } \hat{\rho}^2 = \hat{\rho}; \quad \text{mixed state : } \hat{\rho}^2 \neq \hat{\rho}.$$

Or (**P**):

$$\text{pure state : } \text{Tr}(\hat{\rho}^2) = 1; \quad \text{mixed state : } \text{Tr}(\hat{\rho}^2) < 1.$$

In fact,  $P \equiv \text{Tr}(\hat{\rho}^2)$  is called the purity of a state. A state is pure when its purity equals 1, and mixed otherwise.

31. Another important quantity is the entropy

$$S(\rho) = -\text{Tr}(\hat{\rho} \log[\hat{\rho}]).$$

You might wonder how to calculate the log of a matrix: just diagonalize it, and take the log of the diagonal elements. That is,

$$S(\hat{\rho}) = -\sum_{k=1}^d \lambda_k \log \lambda_k,$$

where  $\lambda_k$  are the eigenvalues of  $\hat{\rho}$ . Note that these eigenvalues are nonnegative, so the entropy can always be defined. Indeed, a zero eigenvalue contributes zero to the entropy, as

$$\lim_{x \downarrow 0} x \log x = 0.$$

32. We use the log base 2, so that the unit of entropy is the *bit*. We can then interpret the entropy as the missing information (in bits) about the state.

33. Using the entropy we get another criterion for pure states vs mixed states (**P**):

$$\text{pure state} : S(\hat{\rho}) = 0$$

$$\text{mixed state} : S(\hat{\rho}) > 0.$$

Thus, there is no missing information for a pure state. A pure quantum state corresponds to maximum information. It doesn't tell us all we could know classically (for example, momentum *and* position), but it is the maximum knowledge quantum mechanics allows us to have.

34. Example: both 2x2 matrices displayed in step 15 have two eigenvalues, 1 and 0. Thus their entropies are zero, and they both describe pure states; we already knew that, of course.

35. In a Hilbert space of dimension  $d$  the entropy can be at most  $\log d$  bits, namely when all eigenvalues  $\lambda_k$  are equal; they must equal  $1/d$  then (**P**). That state is called the maximally mixed (mm) state in that Hilbert space, and the density matrix is proportional to the identity matrix  $\hat{I}$ :

$$\hat{\rho}_{\text{mm}} = \frac{\hat{I}}{d}.$$

36. A unitary operation  $\hat{U}$ —which satisfies by definition  $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}$ — acts on kets as

$$|\psi\rangle \mapsto \hat{U}|\psi\rangle,$$

and hence **(P)** it acts on density matrices as

$$\hat{\rho} \mapsto \hat{U}\hat{\rho}\hat{U}^\dagger.$$

37. The maximally mixed state  $\hat{\rho}_{\text{mm}}$  is easily seen to be a scalar, invariant under rotations, and in fact invariant under *any* unitary transformation **(P)**. Don't confuse  $\hat{\rho}_{\text{mm}}$  with a zero-angular momentum *pure* state, which is also invariant under rotations, but for a different reason **(P)**!

38. Now let us finally see how mixed states arise from pure states of multiple quantum systems, if we consider only one of the systems by itself. Consider a pure state for two quantum systems  $A$  and  $B$ . In general, we can write such a state as a superposition

$$|\Psi\rangle_{AB} = \sum_{nm} a_{nm} |n\rangle_A |m\rangle_B,$$

in terms of bases  $\{|n\rangle_A\}$  for  $A$  and  $\{|m\rangle_B\}$  for  $B$ . Because of normalization we have

$$\sum_{nm} |a_{nm}|^2 = 1.$$

39. Now consider a measurement just on system  $A$ , say an observable  $\hat{O}_A$ . The expectation value of that observable in terms of  $\hat{\rho}_{AB} = |\Psi\rangle_{AB}\langle\Psi|$  is

$$\langle\hat{O}_A\rangle = \text{Tr}(\hat{\rho}_{AB}\hat{O}_A) = \sum_n \sum_m \langle n|\langle m|\hat{\rho}_{AB}|m\rangle_B \hat{O}_A |n\rangle_A,$$

where the ket  $|m\rangle_B$  has been moved to the left, as  $\hat{O}_A$  doesn't do anything to states of system  $B$  (that is, the observable is really  $\hat{O}'_{AB} = \hat{O}_A \hat{I}_B$ ). Thus we can define an operator

$$\hat{\rho}_A = \sum_m \langle m|\hat{\rho}_{AB}|m\rangle_B \equiv \text{Tr}_B \hat{\rho}_{AB},$$

in terms of which

$$\langle\hat{O}_A\rangle = \text{Tr}(\hat{\rho}_A \hat{O}_A).$$

That is,  $\hat{\rho}_A$  as defined above, describes the state of system  $A$  by itself. We also call  $\hat{\rho}_A$  the reduced density matrix of system  $A$ .



40. The operation indicated by  $\text{Tr}_B$  in the previous step is called a ‘partial trace,’ and more in particular, a ‘trace over  $B$ ’. We also say, we ‘traced out’ system  $B$ . So, if you want to describe a quantum system by itself, you have to trace out the rest of the universe!

41. For example, take a pure state of the form

$$|\Psi\rangle_{AB} = \frac{|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B}{\sqrt{2}}.$$

What is  $\hat{\rho}_A$ ? Answer: First calculate

$$|\Psi\rangle_{AB}\langle\Psi| = \frac{1}{2}[|0\rangle_A\langle 0| \otimes |1\rangle_B\langle 1| + |0\rangle_A\langle 1| \otimes |1\rangle_B\langle 0| + |1\rangle_A\langle 0| \otimes |0\rangle_B\langle 1| + |1\rangle_A\langle 1| \otimes |0\rangle_B\langle 0|],$$

where for convenience I inserted a  $\otimes$  sign, to indicate the different projectors before and after the sign act on different Hilbert spaces (namely, those of  $A$  and  $B$ , respectively).

Then take the trace over  $B$ . Only the first and fourth term survive this:

$$\hat{\rho}_A = \text{Tr}_B|\Psi\rangle_{AB}\langle\Psi| = \frac{1}{2}[|0\rangle_A\langle 0| + |1\rangle_A\langle 1|].$$

This is a mixed state:  $\text{Tr}(\hat{\rho}_A^2) = 1/2 < 1$ .

42. In matrix form we have (w.r.t to the basis  $\{|0\rangle, |1\rangle\}$ )

$$\hat{\rho}_A = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

Compare this to step 15: this is really a different matrix than an equal superposition.  $\hat{\rho}_A$  could be interpreted, while exercising lots of caution, as ‘either  $|0\rangle$  or  $|1\rangle$ ’, each with probability  $1/2$ .

43. Another example: Take a pure state of the form

$$|\Psi\rangle_{AB} = |\psi\rangle_A|\phi\rangle_B.$$

Then we have

$$\hat{\rho}_A = \text{Tr}_B|\psi\rangle_A\langle\psi| \otimes |\phi\rangle_B\langle\phi| = |\psi\rangle_A\langle\psi|.$$

This is so because

$$\sum_n \langle n|\phi\rangle\langle\phi|n\rangle = \sum_n \langle\phi|n\rangle\langle n|\phi\rangle = \langle\phi|\phi\rangle = 1.$$

Thus,  $\hat{\rho}_A$  is pure in this case.

44. The pure state

$$|\Psi\rangle_{AB} = \frac{|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B}{\sqrt{2}}$$

is *entangled*, i.e., it cannot be written as a product of states of  $A$  and  $B$ :

$$|\Psi\rangle_{AB} \neq |\psi\rangle_A|\phi\rangle_B.$$

How can we be sure that a pure state *cannot* be written as a product state? Answer: just trace out system  $B$ , and check if  $\hat{\rho}_A$  is pure or not

$$|\Psi\rangle_{AB} \text{ entangled} \leftrightarrow \hat{\rho}_A \text{ mixed}$$

$$|\Psi\rangle_{AB} \text{ not entangled} \leftrightarrow \hat{\rho}_A \text{ pure}$$

45. Example: take

$$|\Psi\rangle_{AB} = (|0\rangle_A|1\rangle_B + |1\rangle_A|1\rangle_B + |0\rangle_A|0\rangle_B + |1\rangle_A|0\rangle_B)/2.$$

This state is not entangled, as tracing out  $B$  gives a pure state for  $A$  ( $\mathbf{P}$ ). Similarly, tracing out  $A$  yields a pure state for  $B$ .

46. In fact, for pure states of two systems,  $A$  and  $B$ , a measure of entanglement is the entropy of the reduced density matrix of either  $A$  or  $B$  (both give the same number!).

$$E(|\Psi\rangle_{AB}) = S(\hat{\rho}_A).$$

Entanglement has units of bits, too. We actually call them ebits. (This simple expression for the amount of entanglement does not hold for mixed states of two systems, unfortunately. Quantifying entanglement for mixed states is a hard problem.)

47. We can always consider a mixed state of a system  $A$  as entangled with some fictitious auxiliary system,  $F$ . Namely, first write it in its diagonal form:  $\hat{\rho}_A = \sum_k \lambda_k |\phi_k\rangle\langle\phi_k|$ , then write

$$|\Psi\rangle_{AF} = \sum_k \sqrt{\lambda_k} |\phi_k\rangle_A |k\rangle_F,$$

where  $\{|k\rangle\}$  is some set of orthogonal states of system  $F$ . Tracing out  $F$  gives  $\hat{\rho}_A$ . Note the square root here! Square roots must appear as we will construct  $\hat{\rho}_A$  from the operator  $|\Psi\rangle\langle\Psi|$ , which is *bilinear* in  $|\Psi\rangle$ .

48. One more (long) application: consider the 1-D harmonic oscillator with frequency  $\omega$ . We know one nice basis to describe this system,  $\{|n\rangle\}$ , the set of number states, eigenstates of the Hamiltonian

$$\hat{H}|n\rangle = \hbar\omega(n + 1/2)|n\rangle.$$

49. But there is another set of states (not a basis!), which are very useful in practice, the so-called coherent states. For reasons soon to become clear, these states are sometimes called “quasi-classical” states. They can be defined as eigenstates of the lowering (or annihilation) operator  $\hat{a}$ . Since that operator is not hermitian, its eigenvalues do not have to be real. So let’s solve the equation

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle,$$

where  $\alpha$  is a complex number, the eigenvalue of  $\hat{a}$ .

50. Expanding  $|\alpha\rangle$  in the number states basis as

$$|\alpha\rangle = \sum_{n=0}^{\infty} a_n |n\rangle,$$

and substituting this in the eigenvalue equation gives

$$a_n \sqrt{n} = \alpha a_{n-1}.$$

for all  $n > 0$ . We can use this recursive relation to express all coefficients in terms of  $a_0$ :

$$a_1 = \alpha a_0; \quad a_2 = \alpha^2 a_0 / \sqrt{2}; \quad a_3 = \alpha^3 a_0 / \sqrt{3!} \dots$$

That is, we find

$$a_n = \frac{\alpha^n}{\sqrt{n!}} a_0.$$

Now the state  $|\alpha\rangle$  should be normalized:

$$1 = \sum_n |a_n|^2 = \sum_n \frac{(|\alpha|^2)^n}{n!} |a_0|^2 = \exp(|\alpha|^2) |a_0|^2,$$

and so we can choose

$$a_0 = \exp(-|\alpha|^2/2).$$

And thus we finally arrive at

$$|\alpha\rangle = \exp(-|\alpha|^2/2) \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

51. The coherent state evolves in time as

$$|\alpha\rangle(t) = \exp(-|\alpha|^2/2) \sum_n \exp(-i(n + 1/2)\omega t) \frac{\alpha^n}{\sqrt{n!}} |n\rangle = \exp(-i\omega t/2) |\alpha \exp(-i\omega t)\rangle.$$

That is, the coherent state stays a coherent state, but its eigenvalue evolves in time! Note the different meaning of the two phase factors appearing here (**P**)!

52. We like coherent states because the expectation values of position and momentum evolve in time in a nice, classical way (both oscillate at frequency  $\omega$ ):

$$\langle x \rangle(t) = \langle \alpha(t) | \hat{x} | \alpha(t) \rangle = A \cos(\omega t - \phi)$$

and

$$\langle p \rangle(t) = \langle \alpha(t) | \hat{p} | \alpha(t) \rangle = -m\omega A \sin(\omega t - \phi) = md\langle x \rangle/dt,$$

where the amplitude of the oscillation is

$$A = \sqrt{\frac{2\hbar}{m\omega}} |\alpha|,$$

and  $\alpha \equiv |\alpha| \exp(i\phi)$ . All this follows from the expectation values of  $a$  and  $a^+$ :

$$\langle \alpha(t) | \hat{a} | \alpha(t) \rangle = \alpha(t),$$

and

$$\langle \alpha(t) | \hat{a}^+ | \alpha(t) \rangle = (\alpha(t))^*.$$

53. Now consider the following mixed state:

$$\hat{\rho} = \frac{1}{2\pi} \int_0^{2\pi} d\phi |\alpha| \exp(i\phi) \langle |\alpha| \exp(i\phi) |.$$

We can view this as a mixture of coherent states with random phase and fixed amplitude  $|\alpha|$ .

We can perform the integration over  $\phi$ , if we first expand the coherent states in number states. The result is that we can rewrite the same mixed state as

$$\hat{\rho} = \exp(-|\alpha|^2) \sum_n \frac{|\alpha|^{2n}}{n!} |n\rangle \langle n|.$$

That is, the mixture of coherent states with random phase but fixed amplitude is equivalent to a mixture of number states with a Poissonian distribution, with  $\bar{n} = |\alpha|^2$ .

54. And one more application: Let us see why a mixed state can be written in infinitely many different ways as mixtures of pure states. This will take a few steps that will teach us two things along the way: (i) that “tracing out a system” can be interpreted in a new way, namely as resulting from a particular measurement on that system; and (ii) that entanglement cannot be used for signaling superluminally (some researchers in the past mistakenly thought one could: by the time we get to the “no-cloning theorem” in Chapter 12 of the textbook, we’ll return to this subject).

55. Given a mixed state for system  $A$  written in its diagonal form as  $\hat{\rho}_A = \sum_k \lambda_k |\phi_k\rangle\langle\phi_k|$ , we write as in item 47

$$|\Psi\rangle_{AF} = \sum_k \sqrt{\lambda_k} |\phi_k\rangle_A |k\rangle_F,$$

where  $\{|k\rangle_F\}$  are orthogonal states of (a fictitious) system  $F$ . When we calculate the Trace over system  $F$  by sandwiching between the orthogonal bras and kets  $\langle k|$  and  $|k\rangle$ , we get back the diagonal form of  $\hat{\rho}_A$ . We can interpret this as follows: Suppose we measure some observable  $\hat{O}$  on system  $F$  that has the set  $\{|k\rangle\}$  as eigenstates. Then, if we get outcome  $k$ , the state of  $A$  collapses to  $|\phi_k\rangle$ , and this occurs with probability  $\lambda_k$ . Thus, the diagonal representation of  $\hat{\rho}_A$  can be interpreted as arising from this particular measurement on  $F$ .

56. We can write the same Trace in (infinitely many) other ways. Namely, we can pretend we measure a different observable on  $F$ , with different eigenstates, say,

$$|\psi_m\rangle = \sum_k U_{mk} |k\rangle.$$

For the set of states  $\{|\psi_m\rangle\}$  to be orthonormal we require that the complex numbers  $U_{mk}$  are matrix elements of a unitary matrix  $U$  ( $\mathbf{P}$ ). In terms of these states we can rewrite

$$|\Psi\rangle_{AF} = \sum_k \sqrt{\lambda_k} |\phi_k\rangle \sum_m U_{mk}^* |\psi_m\rangle = \sum_{m,k} U_{mk}^* \sqrt{\lambda_k} |\phi_k\rangle |\psi_m\rangle.$$

since  $U^{-1} = U^\dagger = U^{(*T)}$ . So, if we introduce the normalized states

$$|\tilde{\phi}_m\rangle = \frac{\sum_k U_{mk}^* \sqrt{\lambda_k} |\phi_k\rangle}{\sqrt{\sum_k \lambda_k U_{mk} U_{mk}^*}} \equiv \frac{\sum_k U_{mk}^* \sqrt{\lambda_k} |\phi_k\rangle}{\sqrt{p_m}},$$

then we can write

$$|\Psi\rangle_{AF} = \sum_m \sqrt{p_m} |\tilde{\phi}_m\rangle |\psi_m\rangle.$$

That is, we can now also write the reduced density matrix as

$$\hat{\rho}_A = \sum_m p_m |\tilde{\phi}_m\rangle\langle\tilde{\phi}_m|.$$

Note the states  $|\tilde{\phi}_m\rangle$  are not orthogonal in general.

57. Exactly because the density matrix stays the same no matter what is measured on system  $F$ , as long as the outcome is not known, no information about what observable has been measured is transmitted.

58. If the dimension of system  $A$ 's Hilbert space is  $d$ , then there are at most  $d$  terms in the decomposition of  $\rho_A$  from step 56. However, we can actually also write it as a mixture of a *larger* number of different terms. Namely, just make the system  $F$  larger by adding more orthonormal states ( $F$  is a fictitious system after all: we can choose its Hilbert space to be larger than that of  $A$ .) Going through the same procedure but with a different (larger) matrix  $U'$  instead of  $U$  leads to a similar expression

$$\hat{\rho}_A = \sum_m p'_m |\tilde{\phi}'_m\rangle\langle\tilde{\phi}'_m|,$$

except that now the sum is over more terms.

59. Example: Take  $\hat{\rho}_A = \frac{1}{3}|0\rangle\langle 0| + \frac{2}{3}|1\rangle\langle 1|$ . Write

$$|\Psi\rangle_{AF} = \sqrt{\frac{1}{3}}|0\rangle|0\rangle + \sqrt{\frac{2}{3}}|1\rangle|1\rangle.$$

Let's say we measure  $F$  in the basis  $(|0\rangle \pm |1\rangle)/\sqrt{2}$ . Then we collapse  $A$  to the normalized states (**P**: insert the identity!)

$$\sqrt{\frac{1}{3}}|0\rangle \pm \sqrt{\frac{2}{3}}|1\rangle$$

with equal probability (of  $1/2$ ). It is easy to check that an equal mixture of these two states indeed gives back the same density operator.

But we could also pretend  $F$  has a third dimension and measure  $F$  in the 3-d basis

$$(|0\rangle + |2\rangle)/\sqrt{2}, \quad (|0\rangle + |1\rangle - |2\rangle)/\sqrt{3}, \quad (|0\rangle - 2|1\rangle - |2\rangle)/\sqrt{6}$$

This collapses  $A$  to the normalized (but nonorthogonal!) states

$$|0\rangle, \quad \sqrt{\frac{1}{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle, \quad \frac{1}{3}|0\rangle - \frac{2\sqrt{2}}{3}|1\rangle$$

with probabilities  $1/6$ ,  $1/3$ ,  $1/2$ , respectively. This gives yet another way of decomposing the mixture  $\hat{\rho}_A$  in terms of pure states. Check this explicitly!