

Density Operators and Ensembles

Robert B. Griffiths

Version of 30 January 2014

Contents

1	Density Operators	1
1.1	Introduction	1
1.2	Partial trace	1
1.3	Interpretation and use of density operators	2
2	Classical and quantum correlations	3
3	Ensembles	4
3.1	Definition	4
3.2	Density operators of ensembles	5
4	Purification of a density operator	5
5	Time development of density operators	5

Reference: CQT = *Consistent Quantum Theory* by Griffiths (Cambridge, 2002)

1 Density Operators

1.1 Introduction

• Notation. What will here be called *density operators* are often (as in CQT) referred to as *density matrices*. Alas, a “density matrix” is not always a matrix, and has nothing whatsoever to do with “density” in the usual meaning of that term. Thus the term “density operator,” which only suffers from the second of these defects, is to be preferred.

★ Mathematical definition: any operator ρ on the Hilbert space \mathcal{H} which is positive (semidefinite) and has trace 1 is a density operator.

◦ Positive: $\rho = \rho^\dagger$ is Hermitian, and none of its eigenvalues is negative.

• If one eigenvalue is 1 and the rest are zero, $\rho = |\psi\rangle\langle\psi|$ for some normalized $|\psi\rangle$, and is called a *pure* state. Otherwise it represents a *mixed* state.

★ The density operator for a single qubit can always be written as a matrix in the standard basis in the following form

$$\rho = \frac{1}{2} \left(I + \mathbf{r} \cdot \boldsymbol{\sigma} \right) = \frac{1}{2} \begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}, \quad (1)$$

where $\mathbf{r} = (x, y, z)$ is a real three-dimensional vector of length ≤ 1 , $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, the σ 's are Pauli matrices, and $\mathbf{r} \cdot \boldsymbol{\sigma} = x\sigma_x + y\sigma_y + z\sigma_z$.

• Pure states correspond to $|\mathbf{r}| = 1$, the surface of the Bloch sphere; mixed states to $|\mathbf{r}| < 1$, the interior of the Bloch sphere (or Bloch ball, if one is a mathematician).

1.2 Partial trace

• Compare notes “Measurements”, Sec. 3.3

★ Let $|\psi\rangle$ be a pre-probability on $\mathcal{H}_a \otimes \mathcal{H}_b$, and suppose we are only interested in calculating probabilities for the subsystem b , i.e., those associated with some decomposition of the identity

$$I = I_a \otimes \sum_k Q^k \quad (2)$$

where the Q^k are projectors on \mathcal{H}_b and sum to I_b . Or perhaps we are interested in the average $\langle B \rangle$ of an operator acting on \mathcal{H}_b .

- These can be calculated using the *reduced density operator*

$$\rho_b := \text{Tr}_a(|\psi\rangle\langle\psi|) \quad (3)$$

- Tr_a is the *partial trace* over \mathcal{H}_a . Applied to any operator on $\mathcal{H}_a \otimes \mathcal{H}_b$ it produces some operator on \mathcal{H}_b . In particular

$$\text{Tr}_a(A \otimes B) = \text{Tr}_a(A)B, \quad (4)$$

where on the right side “ Tr_a ” is the ordinary trace of the operator A on the Hilbert space \mathcal{H}_a .

- This formula, together with the requirement that Tr_a be linear, defines what Tr_a does to a general operator on $\mathcal{H}_a \otimes \mathcal{H}_b$, since any operator can be written as a sum of products of operators.

- Tr_b , the partial trace over \mathcal{H}_b is defined in an analogous fashion.
- Exercise. Show that for a general state

$$|\psi\rangle = \sum_{jk} c_{jk} |a^j b^k\rangle, \quad (5)$$

where $\{|a^j\rangle\}$ and $\{|b^k\rangle\}$ are orthonormal bases of \mathcal{H}_a and \mathcal{H}_b , (3) yields

$$\rho_b = \sum_{jkk'} c_{jk} c_{jk'}^* |b^k\rangle\langle b^{k'}|. \quad (6)$$

- ★ Let R be any operator on $\mathcal{H}_a \otimes \mathcal{H}_b$ and $\langle a^j b^k | R | a^{j'} b^{k'} \rangle$ its matrix elements using orthonormal bases $\{|a^j\rangle\}$ and $\{|b^k\rangle\}$. Then $\text{Tr}_a(R)$ is defined to be the operator whose matrix elements are given by

$$\langle b^k | \text{Tr}_a(R) | b^{k'} \rangle = \sum_j \langle a^j b^k | R | a^j b^{k'} \rangle \quad (7)$$

- Exercise. What is the corresponding expression for matrix elements of $\text{Tr}_b(R)$?

- ★ Let $\{Q^k\}$ be a decomposition of the identity I_b . For a given pre-probability $|\psi\rangle$ on $\mathcal{H}_a \otimes \mathcal{H}_b$, one can calculate the corresponding probability distribution using either $|\psi\rangle$ or the reduced density operator ρ_b :

$$\text{Pr}(Q^k) = \langle \psi | I_a \otimes Q^k | \psi \rangle = \text{Tr}_b(Q^k \rho_b) \quad (8)$$

- Exercise. Check that the second equality is correct when ρ_b is given by (6)

- ★ Let $C = C^\dagger$ be an observable on system b with spectral decomposition

$$C = \sum_k c_k Q^k. \quad (9)$$

Its average can be conveniently written using ρ_b :

$$\langle C \rangle = \langle \psi | I_a \otimes C | \psi \rangle = \sum_k c_k \text{Pr}(Q_k) = \text{Tr}_b(C \rho_b) \quad (10)$$

1.3 Interpretation and use of density operators

- ★ Equations (8) and (10) illustrates the principal role of a density operator in quantum mechanics: it serves as a *pre-probability* from which a probability distribution can be generated by employing a quantum sample space. Thus it is in some sense a quantum analog of a probability distribution in classical statistical mechanics.

- One should not take the attitude that a density operator represents a specific property of a quantum system; instead, it is used to assign a probability to a set of mutually-exclusive properties.

- Sometimes, as discussed above, a density operator is obtained by partial trace over a pure quantum state. But sometimes it is simply a guess, or is obtained by fitting parameters to experimental data.

- The density operator

$$\rho = e^{-\beta H} / \text{Tr}(e^{-\beta H}) \quad (11)$$

used in quantum statistical mechanics for a system in thermal equilibrium, where H is its Hamiltonian and $\beta = 1/k_B T$ the inverse temperature, belongs to this category.

- The polarization of a beam of spin-half particles used in a scattering experiment can be conveniently expressed using the \mathbf{r} in (1), with the parameters x , y , and z determined by experiment.

2 Classical and quantum correlations

★ Classical correlations. Charlie prepares red R and green G slips of paper, inserts them in two opaque envelopes. One goes to Alice in Atlanta, the other to Bob in Boston.

- Probability that Alice has R or G is $1/2$, but if Bob opens his envelope and sees R , he can conclude that Alice's envelope contains G .

• This is an example of statistical inference based upon the use of a conditional probability: $\Pr(A=G)$ and $\Pr(A=G|B=R)$ are two very different things

- No mysterious long-range influences, wave function collapse or the like in the classical world. Or the quantum world.

★ Probabilistic model based on $\Pr(A, B)$, where A and B can be R or G . We have

$$\Pr(R, G) = \Pr(A=R, B=G) = 1/2 = \Pr(G, R); \quad \Pr(R, R) = 0 = \Pr(G, G). \quad (12)$$

- Marginal probabilities

$$\Pr(A=R) = \sum_B \Pr(A=R, B) = 1/2 = \Pr(A=G), \quad (13)$$

and same thing for $\Pr(B=\dots)$. These do *not* reveal the correlation.

- The *conditional* probability $\Pr(B|A) = \Pr(A, B) / \Pr(A)$ reveals the correlations, e.g.,

$$\Pr(B=R|A=R) = 0, \quad \Pr(B=G|A=R) = 1. \quad (14)$$

• Statistical correlations indicate the presence of information: by examining A (and knowing the joint probability distribution) one finds out something about B ; in this sense A contains information about B .

★ In the case of a quantum bipartite system $\mathcal{H}_a \otimes \mathcal{H}_b$ the relation between an entangled $|\psi\rangle$ used as a pre-probability and the two reduced density operators ρ_a and ρ_b obtained from it by taking a partial trace is analogous to the relationship of a classical joint probability distribution $\Pr(A, B)$ and the marginal distributions $\Pr(A)$ and $\Pr(B)$.

• Thus the reduced density operator ρ_b can be used to find the probability distribution for a decomposition of the identity $\{Q^k\}$ for system b , and ρ_a for a decomposition $\{P^j\}$ for system a . However, neither of these density operators contain information about the *correlations* between systems a and b , whereas the state $|\psi\rangle$ from which they are obtained will in general contain additional information.

★ As an example, in the case of two qubits in the famous EPR-Bohm singlet state

$$|\psi\rangle = (|01\rangle - |10\rangle) / \sqrt{2} \quad (15)$$

taking partial traces yields the two density operators

$$\rho_a = I_a/2, \quad \rho_b = I_b/2. \quad (16)$$

These tell us that the given any orthonormal basis for a the probability of each of the two states is $1/2$, and the same for any orthonormal basis for b .

- On the other hand, from (15) one can show that

$$\Pr([0]_a, [0]_b) = \Pr([1]_a, [1]_b) = 0, \quad \Pr([0]_a, [1]_b) = \Pr([1]_a, [0]_b) = 1/2. \quad (17)$$

Thus, for example if $S_z = +1/2$ for particle a , then $S_z = -1/2$ for particle b . This information is *not* available in either of the reduced density operators ρ_a or ρ_b .

□ Exercise. Find a $|\psi\rangle$ for two qubits such that the reduced density operators are the same as in (16), but $S_{az} = S_{bz}$, i.e., the z component of spin is the same for the two particles.

3 Ensembles

3.1 Definition

★ See notes “Probabilities”, Sec. 1.4, for the general idea of an ensemble in probability theory. The term is employed in a similar way in quantum mechanics, but often in a way that causes a lot of confusion because a sample space has not been properly defined. Hence it is best, when possible, to avoid using this term for discussions of quantum probabilities. There is, however, a usage in quantum theory which is precisely defined and of some use.

• Let $\{p_k\}$ be a set of probabilities summing to 1, and $\{|\psi_k\rangle\}$ a set of normalized states, generally not orthogonal to each other. The ensemble \mathcal{E} is then the set of pairs

$$\mathcal{E} = \{p_k, |\psi_k\rangle\}, \quad (18)$$

and the intuitive idea is that the state of the quantum system is $|\psi_k\rangle$ with probability p_k .

◦ Classical counterpart. Think of a slip of colored paper in an opaque envelope, or a die that has been rolled but is concealed under a cup. One may be able to assign probabilities to the different possibilities even though one does not know the actual state of affairs.

• In place of $\{p_k, |\psi_k\rangle\}$ it is sometimes more convenient to employ a collection of un-normalized kets:

$$\mathcal{E} = \{|\check{\psi}_k\rangle\}, \quad |\check{\psi}_k\rangle = \sqrt{p_k} |\psi_k\rangle, \quad p_k = \langle \check{\psi}_k | \check{\psi}_k \rangle. \quad (19)$$

• It is best to think of such an ensemble not as a set of isolated possibilities existing “by themselves,” since the $\{|\psi_k\rangle\}$ need not be orthogonal to each other, and in that case do not represent physically distinct states. Instead, think of them as *labeled* or *tagged* states, meaning that with probability p_k there exists some distinctive state-of-affairs *external* to the system described by one of the $|\psi_k\rangle$, and with which it is correlated.

★ One way to construct an ensemble is to use the expansion of a bipartite pure state $|\psi\rangle$ in the form

$$|\psi\rangle = \sum_j |a^j\rangle \otimes |\beta^j\rangle, \quad (20)$$

where the $|a^j\rangle$ are an orthonormal basis of \mathcal{H}_a . Now think of the $\{|\beta^j\rangle\}$, which are not normalized, as forming an ensemble in the sense of (19).

◦ Note that the states $\{|a^j\rangle \otimes |\beta^j\rangle\}$ are orthogonal to each other, whereas (in general) the $\{|\beta^j\rangle\}$ are not. Hence one should not think of the $|\beta^j\rangle$ as distinct states in the absence of the corresponding $|a^j\rangle$ which serve as “tags” to distinguish them.

★ Another way to generate an ensemble. Charlie has prepared the state $|\psi_k\rangle$ using an appropriate apparatus and a random number generator that prints out k with probability p_k . When the printout is 3, Charlie prepares $|\psi_3\rangle$ and ships it to Alice in a closed box. Even if $|\psi_3\rangle$ is not orthogonal to $|\psi_2\rangle$, it is correlated with the record in Charlie’s notebook. If we do not know what is written in Charlie’s notebook (the tag), it is appropriate to describe the situation by means of an ensemble.

• Ensembles can be useful conceptual tools in quantum information, but will cause confusion if one forgets that the states have labels or tags that render them distinctive.

★ The phases of states in an ensemble are irrelevant, and thus one sometimes writes

$$\mathcal{E} = \{p_k, \rho_k\} \text{ or } \{\check{\rho}_k\}; \quad \rho_k = |\psi_k\rangle\langle\psi_k|, \quad \check{\rho}_k = |\check{\psi}_k\rangle\langle\check{\psi}_k| = p_k\rho_k, \quad (21)$$

where the ρ_k are pure-state density operators (see below). They could also represent mixed states, thus providing a more general notion of ensemble.

★ It is also sometimes useful to generalize the notion of a quantum ensemble to one in which in place of pure states $\rho_k = |\psi_k\rangle\langle\psi_k|$ the individual ρ_k are themselves mixed state density operators. The first expression in (21) makes sense if one sets $\check{\rho}_k = p_k\rho_k$.

3.2 Density operators of ensembles

★ Given an ensemble (18) or (19), one can always calculate a corresponding density operator

$$\rho = \sum_k |\check{\psi}_k\rangle\langle\check{\psi}_k| = \sum_k p_k |\psi_k\rangle\langle\psi_k|. \quad (22)$$

□ Exercise. How do you know that the operator defined in (22) is positive with trace 1?

★ Notice that in the case of the ensemble $\{|\beta^j\rangle\}$ corresponding to the expansion (20), the associated density operator in the sense of (22) is the same as the reduced density operator ρ_b given in (3). This suggests how one is to interpret the density operator associated with an ensemble: it is a pre-probability appropriate for generating probabilities of the system *if one ignores the tags*.

○ To use the classical analogy of red and green slips of paper: the density operator for system b is like the marginal probability distribution for the color of the slip of paper in Bob's envelope when no account is taken of its correlation with the color of the slip in Alice's envelope.

• By contrast, the ensemble with its tags contains additional information not present in the density operator, namely how the quantum state is correlated with the tag.

★ Therefore, one should not confuse the description given by an ensemble with the description provided by a density operator. Either description can serve as a pre-probability, but the latter is a more “coarse-grained” description than the former, with less information.

• Coarse-grained descriptions are often more useful than detailed descriptions, as anyone knows who has tried to look up a subject on the Internet. However, their limitations should be recognized. Attempting to use its density operator to discuss how a quantum system is correlated with its environment is to employ exactly the wrong tool for addressing this question.

★ To be sure, it may be an aid to the imagination to suppose that a density operator has been generated from some ensemble via (22); often an ensemble allows one to think about a problem in terms which are more physically intuitive: “Let's just suppose that what we have here is with probability 1/3 the state $|\psi\rangle$ and with probability 1/3...” There is no harm in reasoning this way, provided one keeps in mind that the same density operator could have been generated from many different ensembles.

4 Purification of a density operator

★ Sometimes it is convenient to imagine that a density operator is the result of applying a partial trace to an entangled ket $|\Psi\rangle$ defined on a tensor product of the Hilbert space of interest and a mythical “reference system.” In the notation of Sec. 1.2, think of \mathcal{H}_b as the system of interest, and \mathcal{H}_a as the reference system.

○ This trick is frequently employed in quantum information theory, and is known as *purification*.

• If $|\Psi\rangle$ is expanded in an orthonormal basis of the reference system, as in (20), the result is an ensemble $\{p_k, |\beta^k\rangle\}$, with the original density operator equal to the density operator of this ensemble in the sense of (22). Sometimes one thinks of this ensemble as arising by carrying out some measurement on the mythical reference system.

• Thinking about things in this way in no way commits one to affirming the actual existence of the reference system and the entangled state, or that measurements are actually carried out on the reference system. These are just convenient myths to aid one's thinking. No harm is done if one remembers that measurements on distant systems (mythical or not) do not have any influence on the system of interest, by what is sometimes referred to as the principle of *Einstein locality*.

5 Time development of density operators

★ If the density operator ρ refers to an *isolated* system, its unitary time development is give by

$$\rho(t) = T(t, t_0)\rho(t_0)T(t_0, t) = T(t, t_0)\rho(t_0)T^\dagger(t, t_0). \quad (23)$$

• Since $T(t, t_0)$ satisfies the Schrödinger equation

$$i\hbar\partial T(t, t_0)/\partial t = HT(t, t_0), \quad (24)$$

differentiating (23) with respect to t shows that $\rho(t)$ satisfies the equation

$$i\hbar d\rho(t)/dt = [H, \rho], \quad (25)$$

which looks very much like Schrödinger's equation for a ket, except that on the right side there is a commutator in place of $H|\psi\rangle$.

□ Exercise. Carry out the steps leading from (23) to (25).

□ Exercise. Suppose that $\rho(t) = |\psi(t)\rangle\langle\psi(t)|$. Show that (25) is a consequence of $i\hbar(d/dt)|\psi(t)\rangle = H|\psi(t)\rangle$.

★ One way to derive (23) is to suppose that $\rho(t)$ comes from an ensemble $\{p_k, |\psi^k(t)\rangle\}$, and that for each k

$$|\psi_t^k\rangle = T(t, t_0)|\psi_0^k\rangle \quad (26)$$

is a solution of Schrödinger's equation. It then follows that

$$\rho(t) = \sum_k p_k |\psi_t^k\rangle\langle\psi_t^k| = \sum_k p_k \left(T(t, t_0)|\psi_0^k\rangle\langle\psi_0^k|T(t_0, t) \right) = T(t, t_0)\rho(t_0)T(t_0, t). \quad (27)$$

★ Suppose a system consists of two subsystems a and b that do not interact with each. That is, the total Hamiltonian H on $\mathcal{H}_a \otimes \mathcal{H}_b$ is the sum $H = H_a + H_b = H_a \otimes I_b + I_a \otimes H_b$ of separate operators for the two subsystems. Then the unitary time development operator factors,

$$T(t, t') = T_a(t, t') \otimes T_b(t, t'). \quad (28)$$

○ Note that this is true if H_a and H_b are time-independent, but also if they depend on t .

□ Exercise. Show that (28) is a consequence of $H = H_a + H_b$.

• As a consequence of (28), if $|\Psi(t)\rangle$ is an entangled ket on $\mathcal{H}_a \otimes \mathcal{H}_b$ whose unitary time development is given by $T(t, t_0)|\Psi_0\rangle$, the reduced density operator

$$\rho_b(t) = \text{Tr}_a \left(|\Psi(t)\rangle\langle\Psi(t)| \right) \quad (29)$$

satisfies (23) with $T(t, t_0)$ replaced by $T_b(t, t_0)$, so once again we arrive at (25), with ρ replaced by ρ_b and H by H_b .

□ Exercise. Verify that $\rho_b(t)$ as defined in (29) satisfies (23) with $T(t, t_0)$ replaced by $T_b(t, t_0)$.

★ If an isolated composite system consists of subsystems which interact with each other, it is still possible to assign density operators as a function of time to the subsystems using (29) and its analogs. However, these reduced density operators no longer satisfy any simple time development equation, even if unitary time development applies to the system as a whole.

• In some cases it is useful to introduce *approximate* differential or differential-integral equations for the time development of the density operator of a subsystem, but this lies outside the scope of these notes.

★ It is worth keeping in mind that a density operator as a function of time provides only limited information about the stochastic time development of a quantum system. It is analogous to the “single-time” probability for a classical system discussed in CQT Sec. 9.2, and lacks information about correlations between different times. More refined stochastic descriptions can be constructed, but one needs to use histories, and $\rho(t)$ is *not* the appropriate tool for assigning probabilities to any but the simplest (two-time) histories.