

7. Mixed states

7.1 Mixed States

7.2 Dynamics of mixed states and operators

7.2.1 Heisenberg picture

7.2.2 Interaction picture

7.3 Partial Trace

7.3.1 Examples

7.4 Entanglement measurement

7.5 Mixed States and interpretation of the density matrix

7.5.1 Classical Macro-states

7.5.2 Quantum Macro-states

7.5.3 Example: Spin- $\frac{1}{2}$

7.1 Mixed States

Until now we have considered systems whose state was unequivocally described by a state vector. Although the result of an observable measurement on the state is probabilistic, until now the state of the system was well defined and evolved in a deterministic way. When we presented the fundamental concepts of QM we defined the state as a complete description of the set of probabilities for all observables. In particular, we put this into the context of the preparation step of an experiment. Since in order to obtain information about a system, the experiment has to be repeated many times, often we deal with an ensemble of systems (either an ensemble of *copies* of the same systems, or an ensemble *in time* of the same system). In many cases, when we repeat in experiment, it might be difficult to prepare the system in exactly the same state (or prepare perfectly identical copies), thus there is some uncertainty on the initial state.

To describe this situation in more abstract terms, we are thus interested in the case where our information regarding the system is not complete. Thus we will associate the concept of state of a system with an ensemble of similarly prepared systems. By this, we mean an ensemble of systems that could have been prepared in principle, we do not need to refer to a concrete set of systems that coexist in space.

The first postulate now reads: to each state corresponds a unique state operator ρ . The dynamical variable X over the ensemble represented by the state operator ρ has expectation value given by: $\langle X \rangle = \text{Tr} \{ \rho X \} / \text{Tr} \{ \rho \} = \sum_i \langle i | \rho X | i \rangle$ (Notice that here the summation is done over some basis, but any basis is equivalent as it gives the same result). If we impose to ρ to have trace 1, the expectation value of X is just $\langle X \rangle = \text{Tr} \{ \rho X \}$. We impose further constraints on ρ :

- $\text{Tr} \{ \rho \} = 1$ as said.
- ρ is self-adjoint $\rho^\dagger = \rho$, so that $\langle X \rangle$ is real.
- ρ is non-negative $\langle u | \rho | u \rangle \geq 0$.

These properties will allow us to associate a probability meaning to ρ . The state operator ρ can be expressed as the sum of projectors: $\rho = \sum_{n=1}^N \rho_n |u_n\rangle \langle u_n|$, where N is the dimension of the space (that is, ρ has a spectral representation in terms of projectors). With the properties established above, we have: $\sum_n \rho_n = 1$, $\rho_n = \rho_n^*$, that is, the coefficients are real: $0 \leq \rho_n \leq 1$.

If the system can also be described by a state vector $|\psi\rangle$, the state operator is given by: $\rho = |\psi\rangle \langle \psi|$. A state that can be written in this way is called **pure state**.

Since the state operator for a pure state is a projector, it is an idempotent: $\rho^2 = \rho$ (Proof: $(|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|$). Therefore, the eigenvalues of ρ and ρ^2 are the same, or $\rho_n^2 = \rho_n$ and they must be either 0 or one. Since we know that the sum of the eigenvalues, which is equal to the trace, must be one, we can deduce that the state operator for a pure state has just one eigenvalue equal one and all the other are zero. This is the definition of a pure state, a state with only one non-zero eigenvalue (and equal to 1). An equivalent formulation is to say that $\text{Tr}\{\rho^2\} = 1$.

A more general state operator can be written as a convex sum of pure states. To define a convex sum, let's consider a set of state operators $\{\rho_i\}$ and the operator $\rho = \sum a_i \rho_i$. If $0 \leq a_i \leq 1 \forall i$ and $\sum a_i = 1$, the sum is said to be convex and ρ is a good state operator.

? Question: Show that the representation as a convex sum of pure states is not unique.

Consider $\rho = a|\psi\rangle\langle\psi| + (1-a)|\varphi\rangle\langle\varphi|$ with $0 \leq a \leq 1$. Now define

$$|x\rangle = \sqrt{a}|\psi\rangle + \sqrt{1-a}|\varphi\rangle$$

$$|y\rangle = \sqrt{a}|\psi\rangle - \sqrt{1-a}|\varphi\rangle$$

By substitution, $\rho = \frac{1}{2}|x\rangle\langle x| + \frac{1}{2}|y\rangle\langle y|$. □

There is actually an infinite number of ways of representing ρ . A state operator that is not pure, is called **mixed state**. The properties of a mixed state are that $\text{Tr}\{\rho^2\} < 1$ and it cannot be expressed in terms of one pure state only.

As said, the state operator for a pure state is the outer product of the pure state vector and its dual: $\rho = |\psi\rangle\langle\psi|$. The expectation value of an observable is therefore $\langle X \rangle = \text{Tr}\{|\psi\rangle\langle\psi|X\} = \text{Tr}\{\langle\psi|X|\psi\rangle\}$ since the trace is invariant under permutation. We find the known result: $\langle X \rangle = \langle\psi|X|\psi\rangle$.

Imagine we have two state operators in the same Hilbert space. We have:

$$0 \leq \text{Tr}\{\rho_1\rho_2\} \leq 1$$

the equality $\text{Tr}\{\rho_1\rho_2\} = 1$ is reached only if the two state operator are equal and pure.

7.2 Dynamics of mixed states and operators

For a pure state, the evolution is dictated by the Schrödinger equation:

$$i\frac{d|\psi\rangle}{dt} = \mathcal{H}|\psi\rangle$$

which has formal solution: $|\psi(t)\rangle = U(t,0)|\psi(0)\rangle$. The unitary operator U (the propagator) that gives the evolution is the solution of the equation:

$$i\frac{dU}{dt} = \mathcal{H}U(t,0)$$

If the Hamiltonian is time-independent, the propagator has the form: $U(t,0) = e^{-i\mathcal{H}t}$. The dynamics of a pure state in state operator form ($\rho = |\psi\rangle\langle\psi|$) is simply given by:

$$\rho(t) = |\psi(t)\rangle\langle\psi(t)| = U(t,0)|\psi(0)\rangle\langle\psi(0)|U^\dagger(0) = U(t,0)\rho(0)U^\dagger(t,0)$$

The equivalent of the Schrödinger equation for the state operators is the Liouville equation:

$$\frac{d\rho}{dt} = -i[\mathcal{H}, \rho],$$

which can be easily derived from the evolution of vector states described by Schrödinger equation.

? Question: Derive the Liouville equation.

Given the definition of density matrix as a convex sum of pure states:

$$\rho = \sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle \langle \psi_{\alpha}|$$

where each vector state obeys Schrödinger equation:

$$i\hbar|\dot{\psi}\rangle = \mathcal{H}|\psi\rangle$$

we obtain, by taking the derivative of the first equation and inserting the second one:

$$\begin{aligned} i\hbar\dot{\rho} &= i\hbar \sum_{\alpha} p_{\alpha} (|\dot{\psi}_{\alpha}\rangle \langle \psi_{\alpha}| + |\psi_{\alpha}\rangle \langle \dot{\psi}_{\alpha}|) \\ &= \sum_{\alpha} p_{\alpha} (\mathcal{H}|\psi_{\alpha}\rangle \langle \psi_{\alpha}| + |\psi_{\alpha}\rangle \langle \psi_{\alpha}| \mathcal{H}) = [\mathcal{H}, \rho] \end{aligned}$$

□

The solution of the Liouville equation is:

$$\rho(t) = U(t)\rho(0)U^{\dagger}(t)$$

7.2.1 Heisenberg picture

As the Liouville equation is more general than the Schrödinger equation, we would like to reformulate the QM dynamics starting from it. We are thus interested in obtaining the evolution of the observables in the Heisenberg picture starting from the Liouville equation.

The expectation value of an observable O at time t is given by the trace: $\langle O(t) \rangle = \text{Tr} \{ \rho(t) O \} = \text{Tr} \{ U(t, 0) \rho(0) U^{\dagger}(t) O \} = \text{Tr} \{ \rho(0) U^{\dagger}(t) O U(t) \}$. Notice that using the invariance of the trace under cyclic permutation it is possible to assign the time dependence either to the state operator (Schrödinger picture) or to the operator (Heisenberg picture). In the first one, the state evolves forward in time while the observable operator is time-independent. In the Heisenberg picture instead, the observable evolves "backward" (since as we saw $U^{\dagger} = U(-t)$, at least for time-independent hamiltonian) and the state operator is fixed. With this last picture we can follow the evolution of the observable without having to establish a state operator, that is, we can generalize this evolution to a class of state operators.

The operator in the Heisenberg picture at time t is given by: $O_H(t) = U^{\dagger}(t, 0) O U(t, 0)$ and it evolves following the equation:

$$\frac{dO_H}{dt} = i [\mathcal{H}(t), O_H(t)] + \left(\frac{\partial O}{\partial t} \right)_H$$

The observable expectation value must be the same in the two pictures:

$$\frac{d\langle O(t) \rangle}{dt} = \text{Tr} \left\{ \frac{d\rho}{dt} O + \rho \frac{\partial O}{\partial t} \right\} = \text{Tr} \left\{ i\rho(t) [\mathcal{H}, O] + \rho(t) \frac{\partial O}{\partial t} \right\}$$

and:

$$\frac{d\langle O(t) \rangle}{dt} = \text{Tr} \left\{ \rho(0) \frac{dO_H}{dt} \right\} = \text{Tr} \left\{ i\rho(0) [\mathcal{H}, O_H] + \rho(0) \left(\frac{\partial O}{\partial t} \right)_H \right\}$$

7.2.2 Interaction picture

We revisit the interaction picture also in the context of the Liouville equation. Assume that the overall Hamiltonian of the system can be written as $\mathcal{H} = \mathcal{H}_0 + V$ (where we separate the known, trivial part \mathcal{H}_0 from the interesting one, V). The transformation to the interaction picture is operated by the propagator $U_I(t) = e^{-i\mathcal{H}_0 t}$, such that $|\psi\rangle_I = U_I^{\dagger} |\psi\rangle$ and $A_I = U_I^{\dagger} A U_I$.

The evolution of the density matrix in the interaction picture $\rho_I = U_I^{\dagger} \rho U_I$, is then:

$$i\dot{\rho}_I = i\dot{U}_I^{\dagger} \rho(t) U_I + iU_I^{\dagger} \dot{\rho} U_I + iU_I^{\dagger} \rho(t) \dot{U}_I$$

with

$$i\dot{U}^\dagger_I = -\mathcal{H}_0 U^\dagger_I(t) \quad \text{and} \quad i\dot{U}_I = U_I(t)\mathcal{H}_0$$

We obtain therefore:

$$\begin{aligned} & -\mathcal{H}_0 U^\dagger \rho(t) U + U^\dagger [\mathcal{H}, \rho(t)] U + U^\dagger \rho(t) U \mathcal{H}_0 \\ &= -[\mathcal{H}_0, \rho_I(t)] + [U^\dagger(t)\mathcal{H}(t)U(t), \rho_I(t)] \\ &= [\mathcal{H}_I, \rho_I(t)] \end{aligned}$$

where $\mathcal{H}_I = U^\dagger_I(t) V U_I(t)$.

A. Example: rf Hamiltonian in the rotating wave approximation

The interaction picture is particularly useful when the Hamiltonian is composed by a large part time independent (\mathcal{H}_0) and a small, time-dependent part $\mathcal{H}_1(t)$. The interaction picture is defined by the operator $U(t)_I = e^{-i\mathcal{H}_0 t}$, which would give the evolution of the state operator if \mathcal{H}_1 were zero. The interaction picture allows to make more evident the effect of the perturbation on the system, by isolating it and often by simplifying the calculations. Let for example consider the following Hamiltonian acting on a two level system ¹⁹:

$$\mathcal{H} = \underbrace{\omega_0 \sigma_z}_{\mathcal{H}_0} + \underbrace{\omega_1 e^{-i\omega_0 t \sigma_z} \sigma_x e^{i\omega_0 t \sigma_z}}_{\mathcal{H}_1}, \quad \omega_0 \gg \omega_1, \quad \rho(0) = (\mathbb{1} + \epsilon \sigma_z)/2$$

Since $[\mathcal{H}_0, \sigma_z] = 0$, in the absence of the perturbation the system does not evolve, it is a constant of the motion. Let us define an unitary operator $R = e^{-i\omega_0 t \sigma_z}$ that operates the transformation to the interaction picture. We can rewrite the Hamiltonian as: $\mathcal{H} = \omega_0 \sigma_z + R \omega_1 \sigma_x R^\dagger$.

The state operator in the interaction picture is given by: $\rho(t)_I = R^\dagger(t) \rho(t) R(t)$. Its evolution is therefore:

$$\frac{d\rho_I}{dt} = \frac{dR^\dagger}{dt} \rho R(t) + R^\dagger(t) \frac{d\rho}{dt} R + R^\dagger(t) \rho \frac{dR}{dt}$$

Notice that $\frac{dR}{dt} = -i\omega_0 \sigma_z R$ and $\frac{dR^\dagger}{dt} = i\omega_0 \sigma_z R^\dagger$. We obtain:

$$\frac{d\rho_I}{dt} = i[\omega_0 \sigma_z, \rho] + R^\dagger(t) \frac{d\rho}{dt} R(t)$$

and using Liouville equation we have:

$$\frac{d\rho_I}{dt} = i[\omega_0 \sigma_z, \rho] - iR^\dagger [\mathcal{H}, \rho] R = -iR^\dagger [\mathcal{H}_1, \rho] R = \omega_1 [R^\dagger (R \sigma_x R^\dagger) R, R^\dagger \rho R] \Rightarrow \frac{d\rho_I}{dt} = -i[\omega_1 \sigma_x, \rho_I(t)]$$

Notice that this is true in general:

$$\frac{d\rho_I}{dt} = -i[\tilde{\mathcal{H}}_1, \rho_I], \quad \tilde{\mathcal{H}} = U^\dagger_I(t) \mathcal{H}_1(t) U_I(t)$$

7.3 Partial Trace

We define the partial trace of a bipartite system on $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ as a linear map $\text{Tr}_B \{ \cdot \}$ from $\mathcal{H}_{AB} \rightarrow \mathcal{H}_A$ (or \mathcal{H}_B) that is determined by the equation

$$\text{Tr}_B \{ A \otimes B \} = A \text{Tr} \{ B \}$$

(where A, B are operators on $\mathcal{H}_A, \mathcal{H}_B$ respectively). This can be extended to more general composite (multipartite) systems. As for the trace, the partial trace is independent of the basis.

Why do we define the partial trace? Consider a composite system composed of two parts, A and B, and an observable of the first system only O_A . The expectation value of the observable on the system A alone is given by:

¹⁹ It could be a nuclear spin in a magnetic field under the action of a weaker rf field

$\langle O_A \rangle = \text{Tr} \{O_A \rho_A\}$ and on the composite system: $\langle O_A \rangle = \text{Tr} \{(O_A \otimes \mathbb{1}_B) \rho_{AB}\}$. We can rewrite this last equation as $\langle O_A \rangle = \text{Tr}_A \{O_A \text{Tr}_B \{\rho_{AB}\}\}$ where Tr_B denote the partial trace on the B system. Thus, to obtain information about observables of a subsystem we can first take the partial trace of the state density operator and then use that to calculate the expectation value.

We use also the partial trace to reduce the dimensionality of the system: $\rho_A = \text{Tr}_B \{\rho_{AB}\}$.

To calculate the partial trace, write ρ as a sum of tensor products $\rho = \sum_{ijkl} m_{ijkl} |a_i\rangle \langle a_j| \otimes |b_k\rangle \langle b_l|$ ²⁰ and for each term we have: $\text{Tr}_B \{|a_i\rangle \langle a_j| \otimes |b_k\rangle \langle b_l|\} = |a_i\rangle \langle a_j| \text{Tr} \{|b_k\rangle \langle b_l|\}$.

We are often interested in describing a particular system inside a larger space and we would like to just describe the state of this system ρ_S without having to describe or know the overall system. The larger system containing the subsystem which we are interested in, can be the environment, a cavity, a field. By doing a partial trace over the environment degrees of freedom we discard the knowledge about them. In general we will obtain a state operator that describes a mixed state (that as we saw, describe some lack of knowledge on the system). The state operator can thus be seen as resulting from the reduction of a larger system to a smaller one, via the partial trace. If the initial multipartite system was entangled, the reduced system is left in a mixed state, since some information was lost. The partial trace reveals the level of entanglement of a state.

7.3.1 Examples

1) Pure product state (separable): $\rho_{AB} = \rho_A \otimes \rho_B$. The reduced density matrix is therefore: $\rho_r^A = \text{Tr}_B \{\rho_{AB}\} = \rho_A$. No information is lost about the A state.

2) Pure entangled state: Bell State. $\rho = (|00\rangle + |11\rangle) \otimes (|00\rangle + |11\rangle)/2 = (|00\rangle \langle 00| + |00\rangle \langle 11| + |11\rangle \langle 00| + |11\rangle \langle 11|)/2$. The partial trace over B picks up only the diagonal terms and it gives the reduced matrix: $\rho_r^A = \text{Tr}_B \{\rho\} = (|0\rangle \langle 0| + |1\rangle \langle 1|)/2$. All the information about the system A is now lost, since it is now in the maximally mixed state (the identity).

7.4 Entanglement measurement

We have seen examples of entangled states, but we haven't given a formal definition of entanglement yet. This is because it is not easy to give such a definition in the most general case. It is however possible to do so in the simplest case of bipartite pure systems. In that case we say that a state is entangled if it cannot be written as $|\psi\rangle = |a\rangle \otimes |b\rangle$. If such a decomposition exists, the state is called a *separable* or *product* state. The **Schmidt** decomposition can be used to check if the state is separable.

■ **Theorem:** For any vector v on the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ of two Hilbert spaces, there exist orthonormal sets on each space $\{u_i^1\}, \{u_i^2\}$ such that v can be written as $v = \sum_{i=1}^m a_i u_i^1 \otimes u_i^2$ with a_i non-negative.

The proof is obtained from the singular value decomposition²¹.

The number m of the vectors needed for the decomposition is called the *Schmidt rank* and the a_i are the Schmidt coefficients. If the Schmidt rank of a vector is one, the associate state is separable. Note that a_i^2 are the eigenvalues of the reduced density matrix obtained by taking the partial trace over the other system. As such, the rank is easily calculated by taking the partial trace.

The Schmidt rank is sometimes used to quantify entanglement for pure, bipartite systems. There exists many other measure of entanglement, however they coincide at least for this simplest case. For more complex cases, multi-partite, mixed states, the measures are not equivalent and sometimes ill-defined.

A. Concurrence

One of the most used metrics for pure bipartite states is the concurrence. It can be operatively defined as: $C = 2|\alpha\delta - \beta\gamma|$, where the 4 coefficients are defined as: $|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$. This metric has the following properties:

²⁰ Notice that by the Schmidt theorem (see [later](#)) we can always find such decomposition.

²¹ The proof is presented in M. Nielsen & I. L. Chuang, *Quantum computation and quantum information* Cambridge University Press (2000).

1. The concurrence is bounded by 0 and 1: $0 \leq C \leq 1$.
2. $C = 0$ iff the state is separable.
3. $C = 1$ for any maximally entangled state.

The four Bell States are maximally entangled states. They correspond to the triplet and singlet manifolds:

$$\begin{aligned} |\varphi^+\rangle &= (|00\rangle + |11\rangle)/2 & |\varphi^-\rangle &= (|00\rangle - |11\rangle)/2 \\ |\psi^+\rangle &= (|01\rangle + |10\rangle)/2 & |\psi^-\rangle &= (|01\rangle - |10\rangle)/2 \end{aligned}$$

We can go from one of the Bell State to another with simple local operations (e.g. $\sigma_x^1 |\varphi^+\rangle = |\psi^+\rangle$), but local operations (that is, operations on single qubit) cannot change the degree of entanglement.

The concurrence can be used to calculate the entanglement even for a mixed state of two qubits. For mixed qubit, an equivalent (more general) definition is given by

$$C(\rho) \equiv \max(0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4})$$

in which $\lambda_1, \dots, \lambda_4$ are the eigenvalues of

$$A = \rho(\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y)$$

in decreasing order (ρ^* is the complex conjugate of the density matrix).

B. Entropy

The von Neumann entropy is defined as

$$S(\rho) = -\text{Tr} \{\rho \log \rho\}$$

The entropy of the reduced density matrix is a good measure of entanglement:

$$\mathcal{E} \rightarrow S(\rho_A) = -\text{Tr} \{\rho_A \log \rho_A\}$$

where $\rho_A = \text{Tr}_B \{\rho\}$. We can prove that this quantity is the same independently of which subsystem we trace over first.

C. Purity

We can also consider the purity of the reduced state as a measure of entanglement

$$\mathcal{E} \rightarrow \text{Pur}(\rho_A) = -\text{Tr} \{\rho_A^2\}.$$

Reference

Dagmar Bruss, *Characterizing entanglement*, Journal of Mathematical Physics, **43**, 9 (2002)

7.5 Mixed States and interpretation of the density matrix

We have seen how a mixed state emerged naturally from tracing over one part of a composite system, when the two parts were entangled. Now we can also introduce a density operator as a probabilistic description of a system, instead of the reduced system of a larger one. We consider an ensemble of systems: this ensemble can arise either because there are many copies of the same system (as for example in NMR, where there are 10^{18} molecules in the sample) or because we are making many experiments on the same system (for example in a photon counting experiment from the same molecule). In this last case we have an ensemble over the time. The requirements on the ensemble are

1. that the elements of the ensemble do not interact with each other (first type of ensemble), and
2. that the system does not have memory (ensemble over time).

With these requirements, the physical ensembles we are considering are equivalent to a more abstract concept of ensemble, as seen at the beginning of the chapter.

7.5.1 Classical Macro-states

In classical statistical mechanics, equilibrium properties of macroscopic bodies are phenomenologically described by the laws of thermodynamics²². The macro-state M depends on a relatively small number of thermodynamic coordinates. To provide a more fundamental derivation of these thermodynamic properties, we can examine the dynamics of the many degrees of freedom N , comprising a macroscopic body. Description of each micro-state μ , requires an enormous amount of information, and the corresponding time evolution is usually quite complicated. Rather than following the evolution of an individual (pure) micro-state, statistical mechanics examines an ensemble of micro-states corresponding to a given (mixed) macro-state. It aims at providing the probabilities $p_M(\mu)$, for the equilibrium ensemble.

A. Microcanonical ensemble

Our starting point in thermodynamics is a mechanically and adiabatically isolated system. In the absence of heat or work input to the system, the internal energy E , and the generalized coordinates x , are fixed, specifying a macro-state $M = (E, x)$. The corresponding set of individual micro-states form the **microcanonical** ensemble. All micro-states are confined to the surface $H(\mu) = E$ in phase space. The probability distribution function for a microstate μ of Hamiltonian \mathcal{H} is thus just given by the number of accessible states $\Omega(E)$ at the fixed energy E :

$$p_E(\mu) = \frac{1}{\Omega(E, x)} \delta(\mathcal{H}(\mu) - E)$$

B. Canonical ensemble

Instead of fixing the energy of the system, we can consider an ensemble in which the temperature of the system is specified and its internal energy is then deduced. This is achieved in the canonical ensemble where the macro-states, specified by $M = (T, x)$, allow the input of heat into the system, but no external work. The system \mathcal{S} is maintained at a constant temperature through contact with a reservoir \mathcal{R} . The reservoir is another macroscopic system that is sufficiently large so that its temperature is not changed due to interactions with \mathcal{S} . The probability distribution function (p.d.f.) for a microstate μ of Hamiltonian \mathcal{H} in the canonical ensemble is

$$p_T(\mu) = \frac{e^{-\beta\mathcal{H}(\mu)}}{Z(T, x)},$$

where the normalization $Z(T, x) = \sum_{\{\mu\}} e^{-\beta\mathcal{H}(\mu)}$ is the partition function and $\beta = 1/k_bT$ (with k_b the Boltzmann factor). Unlike in a microcanonical ensemble, the energy of a system exchanging heat with a reservoir is a random variable, and it is e.g. possible to define a probability distribution for the energy itself (by changing variables from μ to $\mathcal{H}(\mu)$ in the p.d.f. above.)

C. Gibbs and Grand-canonical ensemble

A generalization of the canonical ensemble is to allow the energy to vary by both the addition of heat and work. The Gibbs canonical ensemble describes a system where (mechanical) work is done (which changes the internal variables x). In the Grand-canonical ensemble instead chemical work is performed (which varies the number of particles). Thus the chemical potential μ_c is fixed and N can vary.

[The chemical potential of a thermodynamic system is the amount by which the energy of the system would change if an additional particle were introduced, with the entropy and volume held fixed. The chemical potential is a fundamental parameter in thermodynamics and it is conjugate to the particle number.]

²² (Note: this section and the next one is taken from Prof. Kardar 8.333 “Statistical Mechanics I” notes as available on OCW, in some points with only small changes).

D. Entropy

Given a probability distribution, we can define the entropy S as

$$S = -k_b \sum_a p_a \log(p_a)$$

(with the convention that $x \log(x) \rightarrow 0$ for $x \rightarrow 0$) where p_a describe the probability distribution ($0 \leq p_a \leq 1$, $\sum_a p_a = 1$). It is a measure of our knowledge about the state of the system.

For example, if $p_j = 1$, $p_i = 0$, $\forall i \neq j$, $S = 0$ (minimum entropy, maximum knowledge). If instead we have a uniform distribution $p_i = 1/N$, $\forall i$, S is maximum:

$$S = -k_b \frac{1}{N} \sum_i \log\left(\frac{1}{N}\right) = k_b \log(N).$$

In the ensemble interpretation of the density matrix, the entropy $S(\rho) = -k_b \text{Tr}\{\rho \log \rho\}$ can be seen to have the same meaning as in classical statistics, since we give a probabilistic meaning to the density matrix. Given the decomposition into pure states: $\rho = \sum p_i |\psi_i\rangle \langle \psi_i|$ we obtain that $S(\rho) = -k_b \sum_i p_i \log(p_i)$. In particular the entropy is maximized for the identity state.

The entropy S describes the lack of knowledge in the system and it can also be used to quantify subjective estimates of probabilities. In the absence of any information, the best unbiased estimate is that all N outcomes are equally likely. This is the distribution of maximum entropy. If additional information is available, the unbiased estimate is obtained by *maximizing* the entropy subject to the constraints imposed by this information. The entropy maximization method corresponds to finding the best unbiased estimate by minimizing the amount of information that we introduce in the estimate (given what we know about the distribution).

For example, in the canonical ensemble, we maximize the entropy given a fixed average energy. The canonical ensemble can in fact exchange energy with a large heat bath, so that the system is thermalized and the energy kept fixed. The microcanonical ensemble instead describes an isolated system, where the possible states of the system have the same energy and the probability for the system to be in any given state is the same.

7.5.2 Quantum Macro-states

We can as well formulate a statistical theory for QM. In QM we have seen already that micro-states are described by vectors in Hilbert spaces, evolving unitarily according to the Schrödinger equation. Unlike in classical mechanics, the value of an operator O is not uniquely determined for a particular micro-state. It is instead a random variable, whose average in a state $|\psi\rangle$ is given by $\langle O \rangle = \langle \psi | O | \psi \rangle$.

As in the classical case, we can define quantum macro-states describing ensembles of micro-states. Macro-states of the system depend on only a few thermodynamic functions. We can form an ensemble of a large number N of micro-states μ_a corresponding to a given macrostate. The different micro-states occur with probabilities p_a . (For example $p_a = 1/N$ in the absence of any other information.) When we no longer have exact knowledge of the microstate, it is said to be in a mixed state.

A mixed quantum state is obtained from a set of possible states $\{|\psi_a\rangle\}$, with probabilities $\{p_a\}$. The ensemble average of the quantum mechanical expectation value of an observable O is thus

$$\overline{\langle O \rangle} = \sum_a p_a \langle \psi_a | O | \psi_a \rangle = \sum_{m,n,a} p_a \langle \psi_a | n \rangle \langle n | O | m \rangle \langle m | \psi_a \rangle = \sum_{m,n,a} \langle m | \psi_a \rangle p_a \langle \psi_a | n \rangle \langle n | O | m \rangle = \text{Tr}\{\rho O\}$$

where we defined the density matrix:

$$\rho = \sum_a p_a |\psi_a\rangle \langle \psi_a|$$

with the properties seen above (trace normalization to 1, hermiticity, positivity). We have also already seen that the density matrix obeys the Liouville equation:

$$i\hbar \frac{d\rho}{dt} = [\mathcal{H}, \rho]$$

Equilibrium requires time independent averages, and suggests $\frac{d\rho}{dt} = 0$. This condition is satisfied by choosing $\rho = \rho(\mathcal{H})$, so that $[\rho(\mathcal{H}), \mathcal{H}] = 0$. ρ may also depend on conserved quantities such that $[\mathcal{H}, L] = 0$. Various equilibrium quantum density matrices can now be constructed in analogy to classical statistical mechanics. For example, it is possible to use this minimization of the entropy to calculate the density matrix describing a mixed state.

A. Microcanonical ensemble:

As the internal energy has a fixed value E , a density matrix that includes this constraint is

$$\rho(E) = \frac{\delta(\mathcal{H} - E)}{\Omega(E)}$$

In the matrix representation this can be written as

$$\rho_{n,m} = \langle n | \rho | m \rangle = \sum_a p_a \langle m | \psi_a \rangle \langle \psi_a | n \rangle = \frac{1}{\Omega} \delta(E_n - E) \delta_{n,m},$$

where $\mathcal{H} |n\rangle = E_n |n\rangle$. Thus, only eigenstates of the correct energy can appear in the quantum wave-function and (for $p_a = 1/N$) such states on average have the same amplitude, $|\langle n | \psi_a \rangle|^2 = 1/\Omega$. This is equivalent to the classical postulate of equal a priori equilibrium probabilities. The Ω eigenstates of energy E are combined in a typical microstate with independent random phases. Note that the normalization condition $\text{Tr} \{\rho\} = 1$, implies that $\Omega(E) = \sum_n \delta(E - E_n)$ is the number of eigenstates of \mathcal{H} with energy E .

Notice that we can also obtain the same result by using the maximization of the entropy method. For a microcanonical ensemble, we have no other knowledge on the system than the normalization constraint ($\text{Tr} \{\rho\} = 1$). We thus want to find an unbiased estimate that reflects this minimum knowledge by maximizing the entropy. We thus calculate the density matrix by posing:

$$\begin{cases} \max(S) \\ \text{Tr} \{\rho\} = 1 \end{cases}$$

We can use the Lagrangian multiplier method to solve this problem. Define a function $\mathcal{L} = S - \lambda [\text{Tr} \{\rho\} - 1]$, where λ is a coefficient that multiply the constraint condition. The constrained maximum is found at the maximum of the function \mathcal{L} :

$$\begin{cases} \frac{d\mathcal{L}}{d\rho} = 0 & \rightarrow -k_b \text{Tr} \{\log_2 \rho + \mathbb{1}\} - \lambda \text{Tr} \{\mathbb{1}\} = 0 \\ \frac{d\mathcal{L}}{d\lambda} = 0 & \rightarrow \text{Tr} \{\rho\} = 1 \end{cases}$$

We therefore find $\rho \propto \mathbb{1}$, since $\log(\rho) \propto \mathbb{1}$ from the first equation. From the normalization condition we obtain: $\rho_{ii} = 1/N$, where N is the dimension of the Hilbert space. This expresses the same condition as above (although for a discrete system).

B. Canonical ensemble:

A canonical ensemble describes a system with a fixed temperature. A fixed temperature $T = 1/(k_B \beta)$ can be achieved by putting the system in contact with a reservoir. The canonical density matrix is then obtained by maximizing the system entropy under the constrain of a given average energy.

If the average energy is fixed we have another condition, $\langle E \rangle = \text{Tr} \{\rho \mathcal{H}\}$ in addition to normalization. Therefore:

$$\mathcal{L} = -k_B \text{Tr} \{\rho \log_2 \rho\} - \lambda_1 [\text{Tr} \{\rho \mathcal{H}\} - \langle E \rangle] - \lambda_2 [\text{Tr} \{\rho\} - 1]$$

We can now calculate the maximum of \mathcal{L} :

$$k_b \text{Tr} \{\log_2 \rho + 1\} - \lambda_1 \text{Tr} \{\mathcal{H}\} - \lambda_2 \text{Tr} \{\mathbb{1}\} = 0 \rightarrow \log_2 \rho = -\lambda_1 \mathcal{H} + K \mathbb{1}$$

The density matrix is therefore an exponential: $\rho = e^{-\beta \mathcal{H}} / Z$, where $\beta = 1/(k_B T)$ and Z is the partition function, determined by the normalization condition:

$$Z = \text{Tr} \{e^{-\beta \mathcal{H}}\} = \sum_n e^{-\beta E_n}$$

(where the last expression is calculated in the energy eigenbasis).
 We can calculate the average energy and the entropy:

$$\langle E \rangle = \text{Tr} \{ \mathcal{H} e^{-\beta \mathcal{H}} / Z \} = - \frac{\partial}{\partial \beta} (\ln Z)$$

$$S = -k_B \text{Tr} \{ \rho \log_2 \rho \} = k_B \beta \langle E \rangle + k_B \ln Z$$

In general, any macroscopic observable can be calculated from the partition function.

C. Grand Canonical ensemble

In the Grand Canonical ensemble the number of particles N , is no longer fixed. Quantum micro-states with indefinite particle number span a space called *Fock space* (we will come back to this concept when studying the e.m. field). The density matrix can be obtained as before, where we maximize the entropy, subjected now to conditions on the energy and the particle number. It can be shown (although we only mention it here) that

$$\rho(\beta, \mu) = \frac{e^{-\beta \mathcal{H} + \beta \mu N}}{\mathcal{Q}},$$

where the normalization is:

$$\mathcal{Q}(\beta, \mu) = \text{Tr} \{ e^{-\beta \mathcal{H} + \beta \mu N} \} = \sum_{N=0}^{\infty} e^{\beta \mu N} Z_N(\beta)$$

7.5.3 Example: Spin- $\frac{1}{2}$

Consider a spin- $\frac{1}{2}$ system in a magnetic field along z. The Hamiltonian is then $\mathcal{H} = \frac{\hbar}{2} \gamma B \sigma_z = \hbar \omega \sigma_z$. At thermal equilibrium, the density matrix is

$$\rho = \frac{e^{-\beta \hbar \omega \sigma_z / 2}}{Z}, \quad Z = \text{Tr} \{ e^{-\beta \hbar \omega \sigma_z / 2} \}$$

We find $Z = e^{-\beta \hbar \omega / 2} + e^{\beta \hbar \omega / 2}$ and the expectation values:

$$\langle S_x \rangle = \langle S_y \rangle = 0. \quad \langle S_z \rangle = -\frac{\hbar}{2} \tanh \left(\frac{\beta \hbar \omega}{2} \right)$$

In the high temperature approximation, we can expand the exponential to find $\rho = \frac{1}{2} + \frac{\beta \hbar \omega}{2} \sigma_z$. This is the expression that is used for example in NMR.

MIT OpenCourseWare
<http://ocw.mit.edu>

22.51 Quantum Theory of Radiation Interactions
Fall 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.