

## Entropy in Canonical & Grand Canonical Ensemble

Remember the definition of the H-fct. we used to quantify entropy earlier in the lecture (to avoid confusion I will use the symbol  $\tilde{S}$  for the H-fct, because H is reserved for the Hamiltonian):

$$\tilde{S}[\{w_v\}] := - \sum_v w_v \ln w_v, \text{ where we considered the discrete case, or}$$

$$\tilde{S}[\rho] := - \int \rho \ln \rho \, dx dp, \text{ for a continuous density.}$$

Note: • Throughout these notes  $k_B = 1$  (and  $\hbar = 1$  later on)

• I write  $dx dp$  for  $d^{3N}x d^{3N}p$  (in 3D space)

•  $\int$  means  $\int_{\Gamma}$ ,  $\Gamma = \mathbb{R}^{6N}$  for example

in front of every integral, because  $\int$  means  $\int_{\Gamma}$  but actually the state space is  $\Gamma_N = \Gamma^{(N)}$ , hence by a factor of  $N!$  smaller!

### Canonical Ensemble

indistinguishability of particles

Here we are given  $Z = \frac{1}{N!} \int e^{-\beta H(x,p)} dx, dp$  and  $F(T, V, N) = -T \ln Z$ , where  $\beta = \frac{1}{T}$  inverse temperature,  $H(x,p) \equiv H$  (notation from now on) Hamiltonian, and  $F$  the free energy. We already know that in the thermodynamic limit we get back "conventional thermodynamics", implying that we must get

$$F = U - TS, \text{ where } U = \langle H \rangle = -\partial_{\beta} \ln Z = \frac{1}{N!} \int \frac{e^{-\beta H}}{Z} H \, dx dp.$$

We ask here: Is this consistent with the above definitions for  $\tilde{S}$ , the entropy defined as the H-fct.?

The (normalized) density of the canonical ensemble is

$$\rho(x, p) = \frac{e^{-\beta H(x, p)}}{Z}$$

we must have this factor in front of  $\int$ , as always

$$\text{Therefore: } \tilde{S}[\rho] = -\frac{1}{N!} \int \rho \ln \rho dx dp = - \int \frac{e^{-\beta H}}{N! Z} (-\beta H - \ln Z) dx dp$$

$$= \underbrace{\beta \int \frac{e^{-\beta H}}{N! Z} H dx dp}_{=\langle H \rangle = U} + \ln Z \underbrace{\int \frac{e^{-\beta H}}{N! Z} dx dp}_{=1 \text{ (normalized)}}$$

$$= \frac{1}{T} U - \frac{1}{T} F \Rightarrow F = U - T \tilde{S}$$

$\Rightarrow$  consistent with thermodynamics.

### Grand Canonical Ensemble

$$\text{Here: } \tilde{Z} = \sum_N \frac{1}{N!} \int e^{-\beta(H - \mu N)} dx dp \quad \text{and} \quad \Omega(T, V, \mu) = -T \ln \tilde{Z}.$$

grand potential

Since it is the same potential as we know from thermodynamics we must have:  $\Omega = U - TS - \mu N$ , where  $U = \langle H \rangle = \sum_N \frac{1}{N!} \int \frac{e^{-\beta(H - \mu N)}}{\tilde{Z}} H dx dp$

$$\text{and } N = \langle N \rangle = \sum_N N \cdot \frac{1}{N!} \int \frac{e^{-\beta(H - \mu N)}}{\tilde{Z}} dx dp = z \frac{\partial}{\partial z} \ln \tilde{Z}, \quad \text{where } z = e^{\beta \mu}$$

we have seen that fluctuations are negligible in the thermodynamic limit

For the grand canonical ensemble the (normalized) density reads:

$$\rho(x, p, N) = \frac{e^{-\beta(H - \mu N)}}{\tilde{Z}}$$

Hence:  $\tilde{S}[\rho] = - \sum_N \frac{1}{N!} \int g \ln g \, dx dp = - \sum_N \frac{1}{N!} \int \frac{e^{-\beta(H-\mu N)}}{\mathcal{Z}} (-\beta H + \beta \mu N - \ln \mathcal{Z}) \, dx dp$

$$1 = \sum_N \frac{1}{N!} \int \frac{e^{-\beta(H-\mu N)}}{\mathcal{Z}} \, dx dp = -\beta \mu \underbrace{\sum_N \frac{1}{N!} \int \frac{e^{-\beta(H-\mu N)}}{\mathcal{Z}} \, dx dp}_{\langle N \rangle = N} + \beta \underbrace{\sum_N \frac{1}{N!} \int \frac{e^{-\beta(H-\mu N)}}{\mathcal{Z}} H \, dx dp}_{\langle H \rangle = U} + \ln \mathcal{Z}$$

$$= -\frac{\mu}{T} + \frac{U}{T} - \frac{\Omega}{T} \Rightarrow \Omega = U - T \tilde{S} - \mu N$$

$\Rightarrow$  consistent.

Note: The same can be done for the micro canonical ensemble.

Then we define  $S(E, V, N) = + \ln \omega(E)$  where

$$\omega(E) = \frac{1}{N!} \int_{E \leq H(x,p) \leq E+\delta E} dx dp \quad \text{and hence the density reads } \rho(x,p) = \frac{1}{\omega(E)} \text{ for } x,p \text{ s.t. } E \leq H(x,p) \leq E+\delta E$$

$$\text{we find: } \tilde{S}[\rho] = - \frac{1}{N!} \int g \ln g \, dx dp = - \frac{1}{N!} \int_{E \leq H(x,p) \leq E+\delta E} \frac{1}{\omega(E)} \ln \frac{1}{\omega(E)} \, dx dp$$

$$= \ln \omega(E) \cdot \underbrace{\frac{1}{N!} \int_{E \leq H(x,p) \leq E+\delta E} \frac{1}{\omega(E)} \, dx dp}_{=1, \text{ normalization}} = \ln \omega(E) \equiv S(E, V, N) \Rightarrow \text{consistent.}$$

# The Density Matrix

For an informal but good introduction see:

[itp.ethz.ch/education/fs12/qm2](http://itp.ethz.ch/education/fs12/qm2) → Lidia's lecture notes  
(below)  
→ chapter 7

Most of what I tell here is borrowed from there.

Preliminary note: This is not a complete introduction to the density matrix formalism! See e.g. QIT lecture notes for that.

## Dirac Notation (Bra-ket-notation)

(Families with that?)

Consider a finite dim. quantum system with Hilbert space  $\mathcal{H} = \mathbb{C}^d$ ,  
e.g. a spin- $l$  system of dim.  $d = 2l + 1$ ,  $l = 0, \frac{1}{2}, 1, \dots$

A state of this system is a vector  $v = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix}$  in  $\mathbb{C}^d$ , normalized:  $\sum_{i=1}^d |v_i|^2 = 1$

In Dirac notation we write  $|v\rangle$  ("ket") instead of  $v$ , but essentially we mean the same. The technical difference is that  $|v\rangle$  is not an element of  $\mathcal{H}$  (strictly speaking) but is defined as a map

$$|v\rangle: \mathbb{C} \rightarrow \mathcal{H} \\ \alpha \mapsto \alpha v$$

Likewise we define the adjoint of  $|v\rangle$ ,  $\langle v| \equiv (|v\rangle)^\dagger$ , as a map

$$\langle v|: \mathcal{H} \rightarrow \mathbb{C} \quad \text{where } (\cdot, \cdot) \text{ is the scalar product in } \mathcal{H}. \\ w \mapsto (v, w)$$

Note:  $\langle v| \in \mathcal{H}^*$ , the dual of  $\mathcal{H}$ .

Furthermore we have  $\langle v|w\rangle: \mathbb{C} \rightarrow \mathbb{C}$   
 $x \mapsto x \langle v, w \rangle$

actually we mean  $\langle v|w\rangle$ , but we write  $\langle v|w\rangle$

The important thing is to note that we can also write

$$|v\rangle\langle w|: \mathcal{H} \rightarrow \mathcal{H}, \text{ hence } |v\rangle\langle w| \text{ is an operator}$$

$$u \mapsto (w, u)v$$

from  $\mathcal{H} \rightarrow \mathcal{H}$ , which is linear since  $(\cdot, \cdot)$  is bilinear.  $|v\rangle\langle w|$  takes the overlap of a vector  $u$ , with  $w$ , and maps it to  $v$  times this overlap.

From now on I will use Dirac notation to denote states of quantum systems. Furthermore, instead of  $|v\rangle\langle w|$  I will write  $|v\rangle\langle w|$ .

## Pure States & Mixed States

also called pure state

Suppose we have a quantum state  $|\psi\rangle$  on a Hilbert space  $\mathcal{H}$ ,  $\dim \mathcal{H} = d < \infty$ , with Hamiltonian operator  $H$ . From QM we know:

unitary time-evolution operator

$$|\psi(t)\rangle = e^{-iHt} |\psi(0)\rangle \equiv U_t |\psi(0)\rangle$$

If we want to measure the observable (operator), we know how to compute the probabilities to get outcome  $a_x$ .

$$A = \sum_{x=1}^d a_x \Pi_x \text{ (a self-adjoint operator)}$$

$$= \sum_{x=1}^d a_x |x\rangle\langle x|$$

spectral decomposition

projector

projector on 1-dim. subspace spanned by eigenvector  $|x\rangle$

$$P(a_x)_\psi = |\langle \psi | x \rangle|^2 = \langle \psi | x \rangle \langle x | \psi \rangle = \text{tr} \left[ \underbrace{|x\rangle\langle x|}_{\text{operator on } \mathcal{H}, \text{ linear}} |\psi\rangle\langle\psi| \right] \quad \text{is this clear why?}$$

Now imagine you are given a system, but you do not know exactly in what state it is. The only thing you know is that it is in state  $|\psi\rangle$  with probability  $p$  and in state  ~~$|\psi\rangle$~~   $|\varphi\rangle$  with prob.  $1-p$ . This is called a mixture of states.

Using basic rules to compute probabilities one can find now:

$$P(a_x) = P(\text{given } |\psi\rangle) \cdot P(a_x)_\psi + P(\text{given } |\varphi\rangle) \cdot P(a_x)_\varphi$$

$$= \text{tr} \left[ p |x\rangle\langle x| |\psi\rangle\langle\psi| + (1-p) |x\rangle\langle x| |\varphi\rangle\langle\varphi| \right]$$

tr linear

$$= \text{tr} \left[ p |x\rangle\langle x| |\psi\rangle\langle\psi| + (1-p) |x\rangle\langle x| |\varphi\rangle\langle\varphi| \right]$$

$$= \text{tr} \left[ |x\rangle\langle x| \underbrace{\left( p |\psi\rangle\langle\psi| + (1-p) |\varphi\rangle\langle\varphi| \right)}_{=: \rho} \right]$$

$$= \text{tr} \left[ |x\rangle\langle x| \rho \right]$$

We define the density matrix  $\rho := p |\psi\rangle\langle\psi| + (1-p) |\varphi\rangle\langle\varphi|$ .

It is a matrix because it is an operator on  $\mathcal{H}$ , which can always be represented as a matrix.

Why is this helpful?

Because using  $\rho$  only we can compute the probability of any outcome of any observable.

Furthermore,  $\rho$  allows us to include our ignorance about the actual quantum state

$\Rightarrow$  important to do quantum statistical physics

Def: (technical, for finite dim.)

A density matrix  $\rho$  on a space  $\mathcal{H}$  is a normalized positive operator on  $\mathcal{H}$ , i.e.  $\rho \geq 0$ ,  $\text{tr}[\rho] = 1$ .

Note: • positive operators are hermitian  $\Rightarrow$  can use spectral decomposition theorem:  $\rho = \sum_{i=1}^d p_i \pi_i$  (projectors)

• by normalization and  $\text{tr}[\pi_i] = 1$ :  $\sum_{i=1}^d p_i = 1$ .

• in general this decomposition is not unique

• let  $\{|\varphi_i\rangle\}_{i=1}^d$  be the eigenvectors of  $\rho$  (associated to the eigenvalues  $\{p_i\}_{i=1}^d$ )  $\Rightarrow$  we can write  $\rho = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|$ .

Supplementary note: (Evolution of density matrix)

We know how pure states evolve, namely as  $|\psi(t)\rangle = U_t |\psi(0)\rangle$ .

This implies that  $\langle\psi(t)| = (|\psi(t)\rangle)^\dagger = (U_t |\psi(0)\rangle)^\dagger = \langle\psi(0)| U_t^\dagger$ .

Hence for a density matrix  $\rho = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|$ :

in diagonal representation,  
we can always do that

$$\underline{\underline{\rho(t)}} = \sum_i p_i \underbrace{U_t |\varphi_i(0)\rangle}_{= |\varphi_i(t)\rangle} \underbrace{\langle\varphi_i(0)| U_t^\dagger}_{= \langle\varphi_i(t)|}$$

$$\underline{\underline{= U_t \rho(0) U_t^\dagger}}$$

Hamiltonian

Note: This also follows from the master eq.  $i \frac{\partial \rho}{\partial t} = [H, \rho]$  (see 13.13) in script

## Supplementary note: (superpositions v. mixtures)

It is easy to confuse superpositions (e.g.  $\frac{1}{\sqrt{2}}|\psi\rangle + \frac{1}{\sqrt{2}}|\phi\rangle$ ) with mixtures (e.g.  $\frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\phi\rangle\langle\phi|$ ). They are not the same thing!

Example: Consider a spin- $\frac{1}{2}$  system with the notation  $|0\rangle := |\uparrow\rangle$ ,  $|1\rangle := |\downarrow\rangle$ .

A superposition state is, for example,  $|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$ .

We will show now that this is not the same as the state  $\rho = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2}$ , even though for some measurements the cannot be distinguished.

Consider a meas. in the  $\{|0\rangle, |1\rangle\}$  basis. Then we know:

$$P(0)_{|+\rangle} = |\langle +|0\rangle|^2 = \text{tr} [ |0\rangle\langle 0| (|+\rangle\langle +|) ] = \text{tr} [ |0\rangle\langle 0| \frac{|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 1|}{2} ]$$

$$= \frac{1}{2}$$

$$P(0)_{\rho} = \text{tr} [ |0\rangle\langle 0| \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} ] = \frac{1}{2}$$

and likewise  $P(1)_{|+\rangle} = \frac{1}{2}$ ,  $P(1)_{\rho} = \frac{1}{2} \Rightarrow$  cannot be distinguished with this meas.

But if one measures in the  $\{|+\rangle, |-\rangle := \frac{|0\rangle - |1\rangle}{\sqrt{2}}\}$  basis:

$$P(+)_{|+\rangle} = |\langle +|+\rangle|^2 = 1 \quad \text{and} \quad P(-)_{|+\rangle} = 0$$

$$P(+)_{\rho} = \text{tr} [ |+\rangle\langle +| \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} ] = \text{tr} [ \frac{|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 1|}{2} \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} ]$$

$$= \frac{1}{4} \text{tr} [ \underbrace{|0\rangle\langle 0| |0\rangle\langle 0|}_{=1} ] + \frac{1}{4} \text{tr} [ \underbrace{|1\rangle\langle 1| |1\rangle\langle 1|}_{=1} ] = \frac{1}{2} \quad \text{and} \quad P(-)_{\rho} = \frac{1}{2}$$

$\Rightarrow$  can be distinguished  $\Rightarrow$  cannot be the same.



# The Density Matrix in Statistical Physics

Note. The normalization condition is sometimes neglected (e.g. in the script). That is ok, as long as one always divides by  $\text{tr}[\rho]$  when computing averages of observables or probabilities.

Example: If  $\text{tr}[\rho] \neq 1$ , then an observable  $A$  has expectation value  $\langle A \rangle = \frac{\text{tr}[A\rho]}{\text{tr}[\rho]}$  (see (3.11) in script).

Using this formalism we can incorporate our subjective knowledge in the state description.

For instance, if the only thing we know is the energy  $E$  of the system, then we can assign the state

$$\rho = \sum_{E \leq \epsilon_n \leq E + \delta E} |\psi_n\rangle\langle\psi_n| \quad (\text{unnormalized!})$$

microcanonical ensemble

to the system, where  $\epsilon_n$  and  $|\psi_n\rangle$  are the eigenenergies and eigenstates of the Hamiltonian  $H = \sum_n \epsilon_n |\psi_n\rangle\langle\psi_n|$  of the system, respectively.

Note: This state has trivial time evolution:

$$U_t \rho U_t^\dagger = e^{-iHt} \rho e^{+iHt} = \rho e^{-iHt} e^{+iHt} = \rho.$$

alternatively:

$$i \frac{\partial \rho}{\partial t} = [H, \rho] = 0.$$

$$\begin{aligned} [ \rho, H ] &= 0 \\ \Rightarrow [ \rho, e^{\pm iHt} ] &= 0 \quad \forall t \end{aligned}$$

As in classical physics:  $\rho = \ln \omega(E)$ , where  $\omega(E)$  = occupied state space.

Here:  $\omega(E) = \text{tr}[\rho] = \sum_{E \leq \epsilon_n \leq E + \delta E} \langle \psi_n | \psi_n \rangle = \# \text{ states with energy btw. } E \text{ and } E + \delta E.$