PHYSICS 391

FALL 2017

QUANTUM INFORMATION THEORY I

Jean Weill

1 Introduction

A method is more important than a discovery, since the right method will lead to new and even more important discoveries.

Lev Landau (1908-1968)

The objective of this course was to introduce ourselves to quantum information theory, a booming topic in physics with tremendous practical applications. This paper is based on a lecture I gave for the Kapitza Society.

In the first section we will learn about the density operator: a powerful computing tool in quantum mechanics. We will then have a brief overview of Lagrange multipliers and how to use them to maximize entropy and retrieve the canonical ensemble's properties. Finally, we will revealed a surprising relation between statistical and quantum mechanics.

2 The Density Operator

2.1 Quantum Mechanical Ensembles

We generally may not have perfect knowledge of a prepared quantum state. Suppose a third party, Bob, prepares a state for us and only gives a probabilistic description of it, i.e., Bob selects $|\psi_x\rangle$ with probability $p_X(x)$, where p_X is the probability distribution for the random variable X, and x is in some alphabet $\mathcal{X}$. We can summarize this information by defining an ensemble E of quantum states

$$E \equiv \{p_X(x), |\psi_x\rangle\}_x \tag{1}$$

For example, $E = \left\{\left(\frac{1}{3}, |1\rangle\right), \left(\frac{2}{3}, |3\rangle\right)\right\}$. If this ensemble is span by $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$, then the state is $|0\rangle$ with probability 0, $|1\rangle$ with probability $\frac{1}{3}$, $|2\rangle$ with probability 0, and $|3\rangle$ with probability $\frac{2}{3}$. Of course, the sum of the probabilities adds up to 1.

Consider a system with

$$E = \{p_i, |\psi_i\rangle\}_{i \leq N} \tag{2}$$

with $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. Suppose $i = 1$, then $E = \{p_1, |\psi_1\rangle\} = \{1, |\psi_1\rangle\}$. If we measure an observable $\hat{A}$ in E, then

$$\begin{aligned} \left\langle \hat{A} \right\rangle &= \langle\psi_1|\hat{A}|\psi_1\rangle \\ &= 1\langle\psi_1|\hat{A}|\psi_1\rangle \\ &= p_1\langle\psi_1|\hat{A}|\psi_1\rangle \end{aligned} \tag{3}$$

Now, what if $E = \{(p_1, |\psi_1\rangle), (p_2, |\psi_2\rangle)\}$? If you think as p_1 and p_2 as giving weight to their corresponding state's expectation value, you can guess

$$\left\langle \hat{A} \right\rangle = p_1\langle\psi_1|\hat{A}|\psi_1\rangle + p_2\langle\psi_2|\hat{A}|\psi_2\rangle$$

and finally, if $E = \{p_i, |\psi_i\rangle\}_{i \leq N}$, then

$$\left\langle \hat{A} \right\rangle = \sum_i p_i\langle\psi_i|\hat{A}|\psi_i\rangle \tag{4}$$

Claim:

$$I = \sum_n |n\rangle\langle n| \tag{5}$$

where $\{|n\rangle\}_{n\in N}$ is a complete set, i. e., $\delta_{nm} = \langle n|m\rangle$, and I is the identity operator.

Proof:

$I|m\rangle = \sum_n |n\rangle\langle n|m\rangle = \sum_n |n\rangle\delta_{nm} = |m\rangle$

QED

Thus,

$$\begin{aligned}
\langle \hat{A} \rangle &= \sum_i p_i \langle\psi_i|\hat{A}|\psi_i\rangle \\
&= \sum_i p_i \langle\psi_i|I\hat{A}|\psi_i\rangle \\
&= \sum_i p_i \langle\psi_i|\left(\sum_n |n\rangle\langle n|\right)\hat{A}|\psi_i\rangle \\
&= \sum_i p_i \sum_n \langle n|\hat{A}|\psi_i\rangle\langle\psi_i|n\rangle \\
&= \sum_n \langle n|\hat{A}\left(\sum_i p_i|\psi_i\rangle\langle\psi_i|\right)|n\rangle \\
&= \sum_n \langle n|\hat{A}\hat{\rho}|n\rangle
\end{aligned} \tag{6}$$

where

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{7}$$

is the density operator.

For some matrix B, we have $B_{ij} = \langle i|B|j\rangle$, thus $\langle \hat{A} \rangle = \sum_n (\hat{A}\hat{\rho})_{nn} = Tr(\hat{A}\hat{\rho})$, which gives the very useful and important result

$$\langle \hat{A} \rangle = Tr(\hat{A}\hat{\rho}) \tag{8}$$

Note that the trace is independent of which complete set you use. Suppose $\{|\phi_i\rangle\}_{i\in N}$ is a complete set, then using (5) we get

$$\begin{aligned}
Tr(B) = \sum_i \langle i|B|i\rangle &= \sum_i \langle i|BI|i\rangle \\
&= \sum_i \langle i|B\left(\sum_j |\phi_j\rangle\langle\phi_j|\right)|i\rangle \\
&= \sum_j \langle\phi_j|\left(\sum_i |i\rangle\langle i|\right)B|\phi_j\rangle \\
&= \sum_j \langle\phi_j|B|\phi_j\rangle
\end{aligned} \tag{9}$$

thus, just think of the complete basis $\{|\epsilon_i\rangle\}_{i\in N}$, where for ϵ_i the i^{th} component is 1 and the other components are 0, to build your intuition.

Claim 2: $Tr(\hat{\rho}) = 1$

Proof:

$$
\begin{aligned}
Tr(\hat{\rho}) = \sum_n \langle n|\hat{\rho}|n\rangle &= \sum_n \langle n| \left(\sum_i p_i |\psi_i\rangle\langle\psi_i| \right) |n\rangle \\
&= \sum_i p_i \langle\psi_i| \left(\sum_n |n\rangle\langle n| \right) |\psi_i\rangle \\
&= \sum_i p_i \langle\psi_i| I |\psi_i\rangle \\
&= \sum_i p_i \langle\psi_i|\psi_i\rangle \\
&= \sum_i p_i = 1
\end{aligned}
\tag{20}
$$

QED

where we used the normalization property of the wavefunction, and (5).

Claim 3: $\hat{\rho}$ can be diagonalized to the following

$$
\hat{\rho} = \begin{pmatrix} p_1 & 0 & \dots & & 0 \\ 0 & p_2 & 0 & \dots & 0 \\ & \vdots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & p_{|\Psi|} \end{pmatrix}
\tag{21}
$$

where $|\Psi|$ is the size of the set.

Proof:

We have that $\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, thus

$$
\hat{\rho}|\psi_k\rangle = \sum_i p_i |\psi_i\rangle \delta_{ik} = p_k |\psi_k\rangle
\tag{22}
$$

so $\hat{\rho}$ as eigenvectors $\{|\psi_k\rangle\}_k$ with eigenvalues $\{p_k\}_k$. Let

$$
M = \left(|\psi_1\rangle \; |\psi_2\rangle \; \dots \; \psi_{|\Psi|} \right)
\tag{23}
$$

Then

$$
\hat{\rho}M = \left(\hat{\rho}|\psi_1\rangle \; \hat{\rho}|\psi_2\rangle \; \dots \; \hat{\rho}\,\psi_{|\Psi|} \right) = \left(p_1|\psi_1\rangle \; p_2|\psi_2\rangle \; \dots \; p_{|\Psi|}\,\psi_{|\Psi|} \right)
\tag{24}
$$

which we can rewrite in a more enlightening way

$$
\hat{\rho}M = \left(|\psi_1\rangle \; |\psi_2\rangle \; \dots \; \psi_{|\Psi|} \right) \begin{pmatrix} p_1 & 0 & \dots & & 0 \\ 0 & p_2 & 0 & \dots & 0 \\ & \vdots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & p_{|\Psi|} \end{pmatrix} = M \begin{pmatrix} p_1 & 0 & \dots & & 0 \\ 0 & p_2 & 0 & \dots & 0 \\ & \vdots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & p_{|\Psi|} \end{pmatrix}
\tag{25}
$$

if M is invertible, we get

$$M^{-1}\hat\rho M = \begin{pmatrix} p_1 & 0 & \dots & & 0 \\ 0 & p_2 & 0 & \dots & 0 \\ \vdots & & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & p_{|\psi|} \end{pmatrix} \tag{26}$$

QED

2.3 Pure vs. Mixed Quantum States

<u>Pure:</u> We know the system is in a particular state $|\psi\rangle \Rightarrow E = \{1; |\psi\rangle\}$.
<u>Mixed:</u> We don't know the system is in a particular state $|\psi\rangle$, i.e., the system could be in several states $\{|\psi_i\rangle\}_{i\le N} \Rightarrow E = \{p_i; |\psi_i\rangle\}_{i\le N}$.

Using Claim 3, we can define pure and mixed quantum states using the density operator. A mixed state would be given by (21), and a pure state would be given by

$$\hat\rho = \begin{pmatrix} 0 & 0 & \dots & & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix} \tag{27}$$

since $p_k = 1$ and $p_i = 0$ for all $i \neq k$.
We will come back to this when we will be discussing statistical mechanics and entropy.

2.4 Why is ρ called the Density Operator

$$\langle \hat A \rangle = Tr(\hat A\hat\rho) = \sum_{m,n} \langle n|\hat A|m\rangle\langle m|\hat\rho|n\rangle$$

which becomes for a continuous basis

$$\langle \hat A \rangle = \int d^3x \int d^3x' \langle x|\hat A|x'\rangle \langle x'|\hat\rho|x\rangle \tag{28}$$

Let's focus on $\langle x'|\hat\rho|x\rangle$.

$$\begin{aligned} \langle x'|\hat\rho|x\rangle &= \left\langle x' \right| \left(\sum_i p_i |\psi_i\rangle\langle\psi_i| \right) |x\rangle \\ &= \sum_i p_i \langle x'|\psi_i\rangle \langle\psi_i|x\rangle \\ &= \sum_i p_i \psi_i(x')\psi_i^*(x) \end{aligned} \tag{29}$$

If $x = x'$, then $\langle x|\hat\rho|x\rangle = \sum_i p_i|\psi_i(x)|^2$

3 The Canonical Ensemble

3.1 Lagrange Multipliers

Let S be a surface given by $f(x;y;z) = c$ for some constant c. What if you want to find which point on this surface is the closest to the origin? To do this we need some background. A curve on this surface is defined by $r(t) = (x(t);y(t);z(t))$. Now, the derivative of this curve, $r'(t)$, is tangent to the curve at any point $P = (x(t);y(t);z(t))$. Let's take one of them $P_0 = (x(0);y(0);z(0)) = (x_0;y_0;z_0)$. At P_0, $r'(t)$ is tangent to the curve and thus tangent to S, which means it lies on the tangent plane of $f(x;y;z)$ at P_0. This means that if a vector is perpendicular to $r'(t)$ at P_0, then it is perpendicular to the tangent plane of S at P_0.

Claim: The gradient of f is perpendicular to $r'(t)$ at any point.

If we can prove that $\nabla f|_{P_0} \cdot r'(t_0) = 0$ for some arbitrary point P_0 we would be done. But

$$\frac{df}{dt}\Big|_{P_0} = \frac{\partial f}{\partial x}\Big|_{P_0}\frac{dx}{dt}\Big|_{t_0} + \frac{\partial f}{\partial y}\Big|_{P_0}\frac{dy}{dt}\Big|_{t_0} + \frac{\partial f}{\partial z}\Big|_{P_0}\frac{dz}{dt}\Big|_{t_0} = \nabla f|_{P_0} \cdot r'(t_0) \quad (30)$$

but $f(x{:}y{:}z) = c$ where c is constant, thus $\frac{df}{dt}|_{P_0} = 0$.

QED

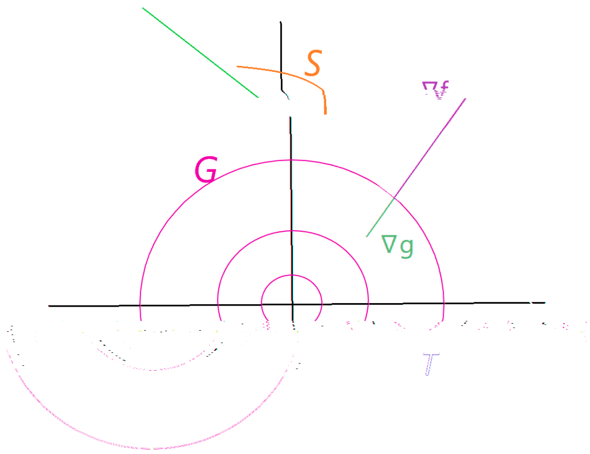

FIG. 1. Lagrange Multipliers

Let's go back to our minimization problem. Try the problem in two dimension, where S is a curve. Take a compass and start drawing circles with a common center at the origin. Start from a small one and increase the radius of the next one. Continue increasing the radius until a circle enters in contact with S

where g is the function containing the constraint, and λ is the famous Lagrange multipliers. The general case is given by

$$\nabla f = \sum_i \lambda \nabla g_i \tag{32}$$

3.2 Mixed States and Entropy

We know that the entropy of a system is given by

$$S = -k \sum_i p_i \ln p_i = -kTr(\hat{\rho} \ln \hat{\rho}) \tag{33}$$

$\hat{\rho}$ is diagonalized. What are the $\{p_i\}$ in the minimum entropy state? From (33), it's easy to see that S is a of positive quantities $\Rightarrow S \geq 0 \Rightarrow S_{min} = 0$. Recall that we need $\sum_i p_i = 1$, which means that $p_k = 1$ and $p_i = 0 \ \forall i \neq k$. This might ring a bell (no pun intended), we have in this case a pure state. Thus, in order to have minimum entropy, an ensemble must be composed of a pure state.

Now what about maximizing entropy? This is where Lagrange multipliers come handy. In this case, $f = S$, and $g = \sum_j p_j - 1$. Looking back at (32), we must have

$$\nabla \left(-k \sum_i p_i \ln p_i\right) = \lambda \nabla \left(\sum_i p_i - 1\right) \tag{34}$$

We can simplify this quite a bit

$$\begin{aligned}\nabla \left(-k \sum_i p_i \ln p_i - \lambda \left(\sum_i p_i - 1\right)\right) &= \sum_i \left(-k(\nabla p_i)\ln p_i - kp_i \frac{\nabla p_i}{p_i} - \lambda \nabla p_i\right) \\ &= \sum_i \left(-k(\ln p_i + 1) - \lambda\right) \nabla p_i\end{aligned} \tag{35}$$

but ∇p_i is arbitrary, thus

$$-k(\ln p_i + 1) - \lambda = 0 \iff p_i = e^{-\frac{\lambda}{k} - 1} \tag{36}$$

Using our constraint $\sum_i p_i = 1$, we get

$$\sum_{i=1}^{|J|} e^{-\frac{\lambda}{k} - 1} = 1 \iff e^{-\frac{\lambda}{k} - 1} = \frac{1}{|J|} \iff p_i = \frac{1}{|J|} \tag{37}$$

This means that in order to maximize entropy, we want a uniform distribution. This gives us that a state is a maximally mixed state if every eigenvalue of the density operator is equal, i.e.,

$$\hat{\rho} = \frac{1}{|J|} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 \cdots & 0 \\ \vdots & & \ddots & \\ & & & 1 \end{pmatrix}$$

This gives us

jdi+[(37d1.6254 -14.446 T84.545 0 Td [(r)]TJ/F38 .

: :C

0 1

3.3 Statistical Mechanics

To some extent, statistical mechanics is an assumption about the density matrix for a macroscopic system. The assumptions (constraints) are

(a) $\sum_i p_i = 1$.

(b) $\langle \hat{H} \rangle = E$ is known.

Constraint (b) can be expressed in a more useful way since from Schroedinger equation $\hat{H}|\psi_k\rangle = E_k|\psi_k\rangle \Rightarrow \langle\psi_k|\hat{H}|\psi_k\rangle = E_k$, which gives

$$\langle \hat{H} \rangle = \sum_k p_k \langle\psi_k|\hat{H}|\psi_k\rangle = \sum_k p_k E_k = E \tag{40}$$

Since we have two constraints, using equation (32) we have that

$$\nabla\left(-k\sum_i p_i \ln p_i\right) = \lambda_1 \nabla\left(\sum_i p_i - 1\right) + \lambda_2 \nabla\left(\sum_i p_i E_i - E\right) \tag{41}$$

which we can simplify to

$$\nabla\left(-k\sum_i p_i \ln p_i - \lambda_1\left(\sum_i p_i - 1\right) - \lambda_2\left(\sum_i p_i E_i - E\right)\right) = \sum_i\left(-k(\ln p_i + 1) - \lambda_1 - \lambda_2 E_i\right)\nabla p_i \tag{42}$$

but ∇p_i is arbitrary, thus

$$-k(\ln p_i + 1) - \lambda_1 - \lambda_2 E_i = 0 \Rightarrow p_i = e^{-\frac{\lambda_1}{k} - \frac{\lambda_2 E_i}{k} - 1} \tag{43}$$

Using our constraint $\sum_i p_i = 1$, we get

$$e^{-\frac{\lambda_1}{k} - 1} = \frac{1}{Z} \tag{44}$$

with

$$Z = \sum_i e^{-\frac{\lambda_2 E_i}{k}} \tag{45}$$

Letting $\frac{\lambda_2}{k} = \beta$, we get the well known canonical ensemble equations

$$p_i = \frac{e^{-\beta E_i}}{Z} \tag{46}$$

and

$$Z = \sum_i e^{-\beta E_i} \tag{47}$$

4 Quantum Mechanics and Statistical Mechanics

4.1 Mixed State, Pure State, and Temperature

The internal energy of a monoatomic gas is given by $E = \frac{3}{2}NkT \Rightarrow E \propto T$, but looking back at (47), we must have $\beta \propto \frac{1}{E}$ since the exponential must be unitless. These two relations imply that $\beta \propto \frac{1}{T}$. Thus,

$$T \to \infty \Rightarrow \beta \to 0^+ \Rightarrow Z = |\Omega| \Rightarrow p_i = \frac{1}{|\Omega|}$$

This means that all states are equally probable, which is what we found for the maximally mixed state. What about $T \to 0^+$? I'm sure you can guess what is about to happen

$$Z = \sum_{i=0} e^{-\beta E_i} = e^{-\beta E_0} \sum_{i=0} e^{-\beta(E_i - E_0)} = e^{-\beta E_0} + e^{-\beta E_0} \sum_{i=1} e^{-\beta(E_i - E_0)} \quad (48)$$

therefore since $E_i > E_0 \ \forall\, i > 0$, we get

$$T \to 0^+ \Rightarrow \beta \to \infty \Rightarrow Z \to e^{-\beta E_0} \quad (49)$$

this means that

$$T \to 0^+ \Rightarrow \beta \to \infty \Rightarrow p_k = \frac{e^{-\beta E_k}}{Z} \to e^{-\beta(E_k - E_0)} \quad (50)$$

thus, $p_k = 0 \ \forall\, k \neq 0$ and $p_0 = 1$. Therefore all the particles are in the ground state E_0. This exactly what we found for a pure state.

4.2 Imaginary Statistical Mechanics

Definition: (Function of a Hermitian Operator) Suppose that a Hermitian operator A has a spectral decomposition $A = \sum_i a_i |i\rangle\langle i|$ for some orthonormal basis $\{|i\rangle\}$. Then the operator $f(A)$ for some function f is defined as follows:

In quantum mechanics, we can determine how a quantum state evolves with time using the operator $U(t)$.

$$| \ (t)\rangle = U(t)| \ (0)\rangle \tag{56}$$

We can get it from looking at Schroedinger equation.

$$\hat{H}| \ (t)\rangle = ih\frac{\partial}{\partial t}| \ (t)\rangle \Rightarrow | \ (t)\rangle = e^{\frac{i\hat{H}t}{h}}| \ (0)\rangle \tag{57}$$

thus, $U(t) = e^{\frac{i\hat{H}t}{h}}$.

By making the change

$$t \to \ i \ h \tag{58}$$

we get

$$U(t) \to U(\ i \ h) = e^{\ \hat{H}} \tag{59}$$

which gives us the startling result that

$$Tr[U(\ i \ h)] = Tr \ e^{whic}$$

but then $T' = \cos\ (z \tan\ + T)$. We can simplify this by using the fact that $\sec^2\ + \tan^2\ = 1$ we get $\sec\ = \sqrt{1\ \ v^2} = \frac{1}{\cos}$. Thus