

The density operator in quantum mechanics¹

D. E. Soper²

University of Oregon

20 April 2012

I offer here some background for Chapter 3 of J. J. Sakurai, *Modern Quantum Mechanics*.

1 The density operator

Consider an ensemble of identical quantum systems. The system has probability w_i to be in quantum state $|\psi_i\rangle$. Here $\langle\psi_i|\psi_i\rangle = 1$, but the states $|\psi_i\rangle$ are not necessarily orthogonal to each other. That means that out of all the examples in the ensemble, a fraction w_i are in state $|\psi_i\rangle$. We must have

$$\begin{aligned} w_i &> 0 \quad , \\ \sum_i w_i &= 1 \quad . \end{aligned} \tag{1}$$

We don't allow $w_i = 0$ because there is no reason to include state $|\psi_i\rangle$ in the description unless there is a non-zero number of ensemble members in that state.

The expectation value for the result of a measurement represented by a self-adjoint operator A is

$$\langle A \rangle = \sum_i w_i \langle \psi_i | A | \psi_i \rangle \quad . \tag{2}$$

(Sakurai uses $[A]$ instead of $\langle A \rangle$.)

We can write the expectation value in a different way using a basis $|K\rangle$

¹Copyright, 2012, D. E. Soper

²soper@uoregon.edu

as

$$\begin{aligned}
\langle A \rangle &= \sum_i w_i \langle \psi_i | A | \psi_i \rangle \\
&= \sum_i w_i \sum_{J,K} \langle \psi_i | J \rangle \langle J | A | K \rangle \langle K | \psi_i \rangle \\
&= \sum_{J,K} \sum_i w_i \langle K | \psi_i \rangle \langle \psi_i | J \rangle \langle J | A | K \rangle \\
&= \sum_{J,K} \langle K | \rho | J \rangle \langle J | A | K \rangle \\
&= \text{Tr}[\rho A] \text{ ,}
\end{aligned} \tag{3}$$

where ρ is the operator

$$\rho = \sum_i w_i |\psi_i\rangle \langle \psi_i| \text{ .} \tag{4}$$

We call ρ the density operator. It provides a useful way to characterize the state of the ensemble of quantum systems.

In what follows, we will speak simply of a system with density operator ρ . That always means that we imagine having many copies of the system – an ensemble of systems.

2 Properties of the density operator

Several properties of ρ follow from its definition. First, its trace is 1 since $\langle 1 \rangle$ must equal 1:

$$\text{Tr}[\rho] = 1 \text{ .} \tag{5}$$

Second, it is self-adjoint:

$$\rho^\dagger = \rho \text{ .} \tag{6}$$

Because it is self-adjoint, it has eigenvectors $|J\rangle$ with eigenvalues λ_J and the eigenvectors form a basis for vector space. Thus ρ has a standard spectral representation

$$\rho = \sum_J \lambda_J |J\rangle \langle J| \text{ .} \tag{7}$$

We can express λ_J as

$$\lambda_J = \langle J | \rho | J \rangle = \sum_i w_i |\langle J | \psi_i \rangle|^2 \tag{8}$$

Since $w_i \geq 0$ and $|\langle J|\psi_i\rangle|^2 \geq 0$, we see that each eigenvalue must be non-negative

$$\lambda_J \geq 0 . \quad (9)$$

The trace of ρ is the sum of its eigenvalues, so

$$\sum_J \lambda_J = 1 . \quad (10)$$

Since each eigenvalue is non-negative, this implies

$$\lambda_J \leq 1 . \quad (11)$$

It is also useful to consider the trace of ρ^2 :

$$\text{Tr}[\rho^2] = \sum_J \lambda_J^2 \leq 1 \times \sum_J \lambda_J = 1 . \quad (12)$$

3 Pure states and mixed states

In Eq. (4), we don't include any terms with $w_i = 0$. Also, suppose there were two states $|\psi_i\rangle$, call them $|\psi_a\rangle$ and $|\psi_b\rangle$ that are the same up to a phase. Then $|\psi_a\rangle\langle\psi_a| = |\psi_b\rangle\langle\psi_b|$. That means that the $i = a$ term and the $i = b$ term in Eq. (4) can be combined in a trivial way. Let's suppose that we always combine them if necessary.

With this understanding of Eq. (4), if there is more than one term and not all of the states $|\psi_i\rangle$ are the same (up to a phase), then we say that the ensemble represents a "mixed state." If there is only one term, so that $\rho = |\psi\rangle\langle\psi|$ for a single state ψ , then we say that the ensemble represents a "pure state." Then $\rho^2 = \rho$. Thus for a pure state we have $\text{Tr}[\rho^2] = 1$.

If $\text{Tr}[\rho^2] = 1$, can we conclude that ρ represents a pure state? Yes. If we have $\lambda_J \geq 0$, $\sum_J \lambda_J = 1$ and $\sum_J \lambda_J^2 = 1$, then all of the λ_J must vanish except for one, call it λ_0 for $J = 0$, with $\lambda_0 = 1$. Then $\rho = |0\rangle\langle 0|$, so ρ is a projection onto the state $|0\rangle$ and thus represents a pure state.

Could we then be sure that there is not another representation of the form (4) in which ρ is a mixed state? Also yes, but we have to think a little more carefully. Suppose that there is another representation of the form (4). Then for any J with $J \neq 0$ we have

$$0 = \lambda_J = \langle 0|\rho|0\rangle = \sum_i w_i |\langle J|\psi_i\rangle|^2 \quad (13)$$

Since we don't allow $w_i = 0$, this means that for each i that occurs in Eq. (4), $\langle J|\psi_i\rangle = 0$. That is, $|\psi_i\rangle$ is orthogonal to each basis vector $|J\rangle$ except for $|0\rangle$. That implies that $|\psi_i\rangle \propto |0\rangle$ and

$$|\psi_i\rangle\langle\psi_i| = |0\rangle\langle 0| \quad (14)$$

That is,

$$\rho = \sum_i w_i |0\rangle\langle 0| \quad (15)$$

We have agreed that if all of the $|\psi_i\rangle$ are the same up to a phase, we combine the like terms and call the ensemble a pure state, not a mixed state. With this understanding, we see that $\text{Tr}[\rho^2] = 1$ implies that ρ represents a pure state.

4 Mixing ensembles

We can mix two ensembles. If one ensemble is represented by a density operator ρ_1 and another is represented by a density operator ρ_2 , then we can make another ensemble by taking a random member of ensemble 1 a fraction f_1 of the time and taking a random member of ensemble 2 a fraction f_2 of the time. Here we need $f_1 + f_2 = 1$. The resulting ensemble has density operator

$$\rho_{\text{tot}} = f_1\rho_1 + f_2\rho_2 \quad (16)$$

5 Spin 1/2 example

A spin 1/2 system provides a nice example of the density operator. Let ρ be a density operator for a spin 1/2 system. Since $\rho^\dagger = \rho$ and $\text{Tr}[\rho] = 1$, we can write ρ in the form

$$\rho = \frac{1}{2} [1 + \vec{a} \cdot \vec{\sigma}] \quad (17)$$

The eigenvalues of $\vec{n} \cdot \vec{\sigma}$ for a unit vector \vec{n} are ± 1 . Thus the eigenvalues of ρ are

$$\lambda_{\pm} = \frac{1}{2} [1 \pm |\vec{a}|] \quad (18)$$

Since λ_- cannot be negative, we must have

$$|\vec{a}| \leq 1 \quad (19)$$

This automatically ensures that $\lambda_+ \leq 1$.

We have

$$\text{Tr}[\rho^2] = \lambda_+^2 + \lambda_-^2 = \frac{1}{2} (1 + \vec{a}^2) . \quad (20)$$

We see that $\text{Tr}[\rho^2] = 1$ if $\vec{a}^2 = 1$. Thus the density operators that represent pure states have $\vec{a}^2 = 1$, while the density operators that represent mixed states have $\vec{a}^2 < 1$.

The value of \vec{a} tells the expectation values of σ_x , σ_y , and σ_z . We have

$$\begin{aligned} \langle \sigma_j \rangle &= \text{Tr}[\rho \sigma_j] \\ &= \frac{1}{2} \text{Tr}[\sigma_j] + \frac{1}{2} a_i \text{Tr}[\sigma_i \sigma_j] \\ &= 0 + a_i \delta_{ij} \\ &= a_j . \end{aligned} \quad (21)$$

Thus

$$\langle \vec{\sigma} \rangle = \vec{a} . \quad (22)$$

We can mix two ensembles $\rho_1 = \frac{1}{2} [1 + \vec{a}_1 \cdot \vec{\sigma}]$ and $\rho_2 = \frac{1}{2} [1 + \vec{a}_2 \cdot \vec{\sigma}]$. The mixture with fraction f_1 of the first ensemble and fraction f_2 of the second ensemble is

$$\rho_{\text{tot}} = f_1 \rho_1 + f_2 \rho_2 = \frac{1}{2} [1 + \vec{a}_{\text{tot}} \cdot \vec{\sigma}] , \quad (23)$$

with

$$\vec{a}_{\text{tot}} = f_1 \vec{a}_1 + f_2 \vec{a}_2 . \quad (24)$$

That is, \vec{a}_{tot} lies on the line from \vec{a}_1 to \vec{a}_2 a fraction f_2 of the way from \vec{a}_1 to \vec{a}_2 . This illustrates that there is more than one way to get a given density operator ρ .

The ensemble with $\rho = 1/2$, that is $\vec{a} = 0$, has $\langle \vec{\sigma} \rangle = 0$. We say that this state is unpolarized.

Exercise 5.1 Consider a statistical ensemble of spin 1/2 particles such that the density operator written in matrix form (in the conventional basis in which J_z is diagonal) is

$$\rho = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \quad (25)$$

What are the expectation values of J_x , J_y , and J_z for this ensemble? Is this a pure state or a mixture? Why?

6 Density operator for higher j

One can have a density operator for the spin space for spin j with $j > 1/2$. However, it is not so simple. With spin j , there are $N = 2j + 1$ dimensions. Thus the matrix representing ρ is an $N \times N$ self-adjoint matrix, which can be characterized with N^2 real numbers. Since we need $\text{Tr}[\rho] = 1$, we can characterize ρ with $N^2 - 1$ real numbers. Thus for spin 1, we have $N = 3$ and $N^2 - 1 = 8$. Thus we need more than $\langle S_x \rangle$, $\langle S_y \rangle$, and $\langle S_z \rangle$ to characterize ρ .

7 Entropy

With a mixed state, we have less than perfect knowledge of what the quantum state is. One uses the entropy to describe how much less. We define the entropy by

$$S = -\text{Tr}[\rho \log(\rho)] . \quad (26)$$

Actually, the definition is $S = -k_B \text{Tr}[\rho \log(\rho)]$ where k_B is the Boltzmann constant. We can get rid of k_B by using units with $k_B = 1$. Then, for instance, we measure temperatures in eV.

In terms of the eigenvalues λ_J of ρ , this is

$$S = -\sum_J [\lambda_J \log(\lambda_J)] . \quad (27)$$

We should understand here that in the case of an eigenvalue zero we take

$$\lim_{\lambda \rightarrow 0} \lambda \log(\lambda) = 0 . \quad (28)$$

For a pure state, all of the λ_J vanish except for one, which equals 1. Then, since $\log(1) = 0$, we have $S = 0$. Mixed states have $S > 0$ since $-\log(\lambda) > 0$ for $\lambda < 1$.

For instance, for our spin 1/2 example, we have

$$\begin{aligned} S &= -\lambda_+ \log(\lambda_+) - \lambda_- \log(\lambda_-) \\ &= -\frac{1}{2} [1 + |\vec{a}|] \log\left(\frac{1}{2} [1 + |\vec{a}|]\right) - \frac{1}{2} [1 - |\vec{a}|] \log\left(\frac{1}{2} [1 - |\vec{a}|]\right) . \end{aligned} \quad (29)$$

This varies from $S = 0$ at $|\vec{a}| = 1$ to $S = \log 2$ at $\vec{a} = 0$.

8 Products of vector spaces

If we have one vector space that describes one set of quantum variables and another vector space that describes another set of quantum variables, we can form the tensor product of the two vector spaces. A simple example is that we have a vector space \mathcal{V}_A that describes particle A and another vector space \mathcal{V}_B that describes particle B. Then we can form the tensor product space $\mathcal{V}_A \otimes \mathcal{V}_B$. Then if $|\phi_A\rangle \in \mathcal{V}_A$ and $|\phi_B\rangle \in \mathcal{V}_B$, then we can form

$$|\phi_A, \phi_B\rangle = |\phi_A\rangle \otimes |\phi_B\rangle \in \mathcal{V}_A \otimes \mathcal{V}_B . \quad (30)$$

Since $\mathcal{V}_A \otimes \mathcal{V}_B$ is a vector space, any linear combination of states of the form $|\phi_A\rangle \otimes |\phi_B\rangle$ is also in $\mathcal{V}_A \otimes \mathcal{V}_B$. Note that not every vector in $\mathcal{V}_A \otimes \mathcal{V}_B$ has the form of a product of a vector from \mathcal{V}_A times a vector from \mathcal{V}_B .

The inner product for product states is defined by

$$\langle \psi_A, \psi_B | \phi_A, \phi_B \rangle = \langle \psi_A | \phi_A \rangle \langle \psi_B | \phi_B \rangle . \quad (31)$$

Then the inner product for states that are linear combinations of product states is defined by the property that the inner product is linear in its ket vector and antilinear in its bra vector. We can construct a basis for $\mathcal{V}_A \otimes \mathcal{V}_B$ from product states. If the vectors $|i\rangle_A$ form a basis for \mathcal{V}_A and the vectors $|j\rangle_B$ form a basis for \mathcal{V}_B , then the vectors $|i\rangle_A \otimes |j\rangle_B$ form a basis for $\mathcal{V}_A \otimes \mathcal{V}_B$.

If \mathcal{O}_A is an operator on \mathcal{V}_A and \mathcal{O}_B is an operator on \mathcal{V}_B , then we can define an operator $\mathcal{O}_A \otimes \mathcal{O}_B$ on $\mathcal{V}_A \otimes \mathcal{V}_B$ by

$$(\mathcal{O}_A \otimes \mathcal{O}_B) (|\phi_A\rangle \otimes |\phi_B\rangle) = (\mathcal{O}_A |\phi_A\rangle) \otimes (\mathcal{O}_B |\phi_B\rangle) . \quad (32)$$

There is also a natural definition of \mathcal{O}_A as an operator on $\mathcal{V}_A \otimes \mathcal{V}_B$. We can understand \mathcal{O}_A acting on $\mathcal{V}_A \otimes \mathcal{V}_B$ as $\mathcal{O}_A \otimes 1$:

$$\mathcal{O}_A (|\phi_A\rangle \otimes |\phi_B\rangle) = (\mathcal{O}_A |\phi_A\rangle) \otimes |\phi_B\rangle . \quad (33)$$

A simple example is for two spinless particles, A and B. Particle A has position \vec{x}_A and particle B has position \vec{x}_B . We often describe a vector $|\Psi_A\rangle$ for particle A by giving its wave function $\Psi_A(\vec{x}_A) = \langle \vec{x}_A | \Psi_A \rangle$. Here we use the fact that the states $|\vec{x}_A\rangle_A$ form a basis for \mathcal{V}_A . Similarly we use the basis states $|\vec{x}_B\rangle_B$ of \mathcal{V}_B to define the wave functions $\Psi_B(\vec{x}_B) = \langle \vec{x}_B | \Psi_B \rangle$. For both particles together, we use basis states

$$|\vec{x}_A, \vec{x}_B\rangle = |\vec{x}_A\rangle_A \otimes |\vec{x}_B\rangle_B \quad (34)$$

to define wave functions

$$\Psi(\vec{x}_A, \vec{x}_B) = \langle \vec{x}_A, \vec{x}_B | \Psi \rangle . \quad (35)$$

Sometimes the wave function that represents both particles together is a product, $\Psi(\vec{x}_A, \vec{x}_B) = \Psi_A(\vec{x}_A)\Psi_B(\vec{x}_B)$. Then $|\Psi\rangle = |\Psi_A\rangle_A \otimes |\Psi_B\rangle_B$. But note that not all functions $\Psi(\vec{x}_A, \vec{x}_B)$ are products of two functions in this fashion.

Another simple example is the product of a space that carries the spin j_1 representation of the rotation group and another that carries the spin j_2 representation. The tensor product of these spaces is a vector space with basis elements

$$|j_1, j_2, m_1, m_2\rangle = |j_1, m_1\rangle \otimes |j_2, m_2\rangle . \quad (36)$$

The vectors in this space represent physically the combined j_1 and j_2 systems. The general vector in the tensor product space is a linear combination of these basis states.

9 Why do we get mixed states?

If we deliberately make a statistical ensemble of pure quantum states, then the ensemble can be described using a density operator. But can't we just stick with pure quantum states and avoid mixed states? Apparently not, because the pure states have zero entropy and we know from statistical mechanics that there is a tendency for entropy to increase. How does that happen?

The simplest answer is that we get mixed states because we don't look at everything. Consider two electrons, thinking of just their spins. Each electron has spin 1/2. A basis for the states of the two electrons together is

$$|1/2, 1/2, m_A, m_B\rangle = |1/2, m_A\rangle_A \otimes |1/2, m_B\rangle_B . \quad (37)$$

Here m_A and m_B can be either $+1/2$ or $-1/2$. A particularly interesting state of this system is the so-called spin singlet state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|1/2, +1/2\rangle_A \otimes |1/2, -1/2\rangle_B - |1/2, -1/2\rangle_A \otimes |1/2, +1/2\rangle_B \right) . \quad (38)$$

Suppose that the A electron is sent to Alice and the B electron is sent to Bob. If Bob measures S_z for his electron he gets $+1/2$ with probability $1/2$

and $-1/2$ with probability $1/2$. If Alice measures S_z for her electron she gets $+1/2$ with probability $1/2$ and $-1/2$ with probability $1/2$. However, if they both make a measurement, then Alice always gets the opposite of what Bob gets.

In fact, this state is a bit more special. Let \vec{n} be any unit vector. If Bob measures $\vec{S} \cdot \vec{n}$ for his electron he gets $+1/2$ with probability $1/2$ and $-1/2$ with probability $1/2$. If Alice measures $\vec{S} \cdot \vec{n}$ for her electron she gets $+1/2$ with probability $1/2$ and $-1/2$ with probability $1/2$. As with the case that \vec{n} is along the z axis, if they both make a measurement, then Alice always gets the opposite of what Bob gets. To see this, one just needs to write the vectors in the basis of eigenvectors of $\vec{S} \cdot \vec{n}$.

Now notice that Alice and Bob are dealing with a pure quantum state – the spin singlet state of two electron spins.

But what if Bob abandons his laboratory and joins a circus? Now nobody is looking at Bob’s electron. What does Alice see? No matter how she adjusts the axis \vec{n} of her Stern-Gerlach apparatus, she sees that her electron has a probability $1/2$ to have $\vec{S} \cdot \vec{n} = +1/2$ and a probability $1/2$ to have $\vec{S} \cdot \vec{n} = -1/2$. Thus her electron is “unpolarized,” described by a statistical ensemble with $\rho = 1/2$.

This is common. We may have a quantum system that we subject to experiments, but our quantum system is typically interacting with the environment. Thus the quantum state of our system becomes entangled with the quantum state of the environment. This means that the quantum state of both together is not just a product of the state of our system and the state of the environment, just as the states of Alice’s and Bob’s electrons is not just something with the form $|\psi_A\rangle \otimes |\psi_B\rangle$. Then if we don’t measure the quantum state of the environment, we find that the system that we are interested in must be described as a mixed state with a density operator.

In statistical mechanics, we often describe the environment as a “heat bath.” The heat bath exchanges energy with the system of interest.

10 The canonical ensemble

In quantum statistical mechanics, one finds ρ in certain standard situations. One of the most important is the following. Let $E = \langle H \rangle$ be the expectation value of the energy of the system. Let’s say that we know what E is, but we don’t know anything else about the system. Then we look for the density

operator ρ with the biggest entropy subject to the two conditions

$$\begin{aligned}\mathrm{Tr}[\rho] &= 1 \quad , \\ \mathrm{Tr}[\rho H] &= E \quad .\end{aligned}\tag{39}$$

The solution of this problem is

$$\rho = \frac{1}{\mathrm{Tr}[e^{-\beta H}]} e^{-\beta H} \quad .\tag{40}$$

This is the ‘‘canonical ensemble.’’

To get this is not so easy. The derivation in Sakurai is not really complete. Here is a derivation. First, define an operator

$$X = \log(\rho)\tag{41}$$

so that

$$\rho = e^X \quad .\tag{42}$$

Consider a small variation δX of X . There is a corresponding small variation $\delta\rho$ of ρ . We want $-\mathrm{Tr}[\rho X]$ to be maximum subject to the two conditions: $\mathrm{Tr}[\rho] = 1$ and $\mathrm{Tr}[\rho H] = E$. That is

$$\mathrm{Tr}[(\delta\rho)X] + \mathrm{Tr}[\rho(\delta X)] = 0\tag{43}$$

for every variation δX that satisfies

$$\begin{aligned}\mathrm{Tr}[\delta\rho] &= 0 \quad , \\ \mathrm{Tr}[(\delta\rho)H] &= 0 \quad .\end{aligned}\tag{44}$$

Now, we would like to say that $\delta\rho = \rho \delta X$, but that isn't right if δX does not commute with ρ . (The calculus of variations is harder with operators than with numbers.) Instead, we can write

$$e^{X+\delta X} = e^{(X+\delta X)/N} e^{(X+\delta X)/N} \dots e^{(X+\delta X)/N} \quad ,\tag{45}$$

where N is very big and there is a product of N factors of $e^{(X+\delta X)/N}$. Here $e^{(X+\delta X)/N}$ is equivalent to $1 + (X + \delta X)/N$ because N is large. Then to first order in δX we have

$$\begin{aligned}e^{X+\delta X} - e^X &= [\delta X/N] e^{X/N} \dots e^{X/N} \\ &+ e^{X/N} [\delta X/N] \dots e^{X/N} \\ &+ e^{X/N} e^{X/N} \dots [\delta X/N] \quad .\end{aligned}\tag{46}$$

Here the factor $[\delta X/N]$ appears in the first place, then in the second place, then in the third place, *etc.* We can write this as

$$\delta e^X = \sum_{J=0}^{N-1} \frac{1}{N} (e^{X/N})^{N-J-1} \delta X (e^{X/N})^J . \quad (47)$$

We have a sum here with many small terms. We can replace this by an integral over a parameter λ that is approximated by a sum with step size $\Delta\lambda = 1/N$:

$$\delta e^X = \int_0^1 d\lambda e^{(1-\lambda)X} \delta X e^{\lambda X} . \quad (48)$$

This seems a little complicated, but it is the price that we must pay to take into account that δX , which could be a variation in any direction in operator space and might not commute with X .

Using Eq. (48), we have

$$0 = \int_0^1 d\lambda \text{Tr}[e^{(1-\lambda)X} \delta X e^{\lambda X} X] + \text{Tr}[e^X \delta X] \quad (49)$$

for every variation δX that satisfies

$$\begin{aligned} 0 &= \int_0^1 d\lambda \text{Tr}[e^{(1-\lambda)X} \delta X e^{\lambda X}] , \\ 0 &= \int_0^1 d\lambda \text{Tr}[e^{(1-\lambda)X} \delta X e^{\lambda X} H] . \end{aligned} \quad (50)$$

Since $\text{Tr}[AB] = \text{Tr}[BA]$, X commutes with itself, and $\int_0^1 d\lambda = 1$, Eq. (49) can be written

$$0 = \text{Tr}[\delta X e^X X] + \text{Tr}[e^X \delta X] . \quad (51)$$

With the same sort of manipulations, we see that the required conditions (50) on the variations δX are

$$\begin{aligned} 0 &= \text{Tr}[e^X \delta X] , \\ 0 &= \text{Tr} \left[\delta X \int_0^1 d\lambda e^{\lambda X} H e^{(1-\lambda)X} \right] . \end{aligned} \quad (52)$$

Finally, since the first of the two conditions is that the second term in Eq. (51) vanishes, our revised statement of the problem is as follows. We are to find X such that

$$0 = \text{Tr}[e^X X \delta X] . \quad (53)$$

whenever the variation δX is such that

$$\begin{aligned} 0 &= \text{Tr}[e^X \delta X] , \\ 0 &= \text{Tr}[\tilde{H} \delta X] . \end{aligned} \quad (54)$$

where

$$\tilde{H} = \int_0^1 d\lambda e^{\lambda X} H e^{(1-\lambda)X} . \quad (55)$$

Now we have a classic problem in calculus of variations. We want to find a vector \vec{F} such that $\vec{F} \cdot \delta\vec{x} = 0$ whenever $\vec{A} \cdot \delta\vec{x} = 0$ and $\vec{B} \cdot \delta\vec{x} = 0$. In the classic calculus of variations problem, \vec{F} , \vec{A} and \vec{B} are the gradients of certain functions. The solution is that \vec{F} must be a linear combination of \vec{A} and \vec{B} . The coefficients are known as Lagrange multipliers. In our case, the role of the dot product between vectors, as in $\vec{F} \cdot \delta\vec{x}$, is taken by the trace of a product of operators, $\text{Tr}[F(\delta X)]$.

We conclude that $e^X X$ must be a linear combination of e^X and \tilde{H} :

$$e^X X = \alpha e^X - \beta \tilde{H} . \quad (56)$$

That is

$$X = \alpha - \beta \hat{H} , \quad (57)$$

where

$$\hat{H} = e^{-X} \tilde{H} = \int_0^1 d\lambda e^{-(1-\lambda)X} H e^{(1-\lambda)X} . \quad (58)$$

Now I claim that

$$X = \alpha - \beta H , \quad (59)$$

solves our problem. That is because with this X we have $[X, H] = 0$, so that $\hat{H} = H$. Thus our solution is

$$\rho = e^\alpha e^{-\beta H} \quad (60)$$

The requirement that $\text{Tr}[\rho] = 1$ fixes α :

$$e^\alpha = 1/\text{Tr}[e^{-\beta H}] . \quad (61)$$

Finally, $\text{Tr}[\rho H] = E$ fixes the value of β (which is $1/T$ where T is the temperature of the system, or $1/(k_{\text{B}}T)$ if you don't set $k_{\text{B}} = 1$).

This gives the canonical ensemble. There is a derivation in Sakurai, but that derivation assumes that the operator $\delta\rho$ has the same eigenvectors as H and also ρ has the same eigenvectors. Then all the operators commute and everything is simple. However, that doesn't tell us what to do for a general $\delta\rho$ and doesn't tell us how to deduce that ρ and H have the same eigenvectors.