Quantum statistical mechanics is statistical mechanics applied to quantum mechanical systems. In quantum mechanics a statistical ensemble (probability distribution over possible quantum states) is described by a density operator $S$, which is a non-negative, self-adjoint, trace-class operator of trace 1 on the Hilbert space $\mathcal{H}$ describing the quantum system. This can be shown under various mathematical formalisms for quantum mechanics. One such formalism is provided by quantum logic.

**Expectation**

From classical probability theory, we know that the expectation of a random variable $X$ is completely determined by its distribution $D_X$ by

$$\mathbb{E}(X) = \int_{\mathbb{R}} \lambda \, d\, D_X(\lambda)$$

assuming, of course, that the random variable is integrable or that the random variable is non-negative. Similarly, let $A$ be an observable of a quantum mechanical system. $A$ is given by a densely defined self-adjoint operator on $\mathcal{H}$. The spectral measure of $A$ defined by

$$E_A(U) = \int_U \lambda \, d\, E(\lambda),$$

uniquely determines $A$ and conversely, is uniquely determined by $A$. $E_A$ is a boolean homomorphism from the Borel subsets of $\mathbb{R}$ into the lattice $Q$ of self-adjoint projections of $\mathcal{H}$. In analogy with probability theory, given a state $S$, we introduce the distribution of $A$ under $S$ which is the probability measure defined on the Borel subsets of $\mathbb{R}$ by

$$D_A(U) = \text{Tr}(E_A(U)S).$$

Similarly, the expected value of $A$ is defined in terms of the probability distribution $D_A$ by

$$\mathbb{E}(A) = \int_{\mathbb{R}} \lambda \, d\, D_A(\lambda).$$
Note that this expectation is relative to the mixed state $S$ which is used in the definition of $D_A$.

**Remark.** For technical reasons, one needs to consider separately the positive and negative parts of $A$ defined by the Borel functional calculus for unbounded operators. One can easily show:

$$E(A) = \text{Tr}(AS) = \text{Tr}(SA).$$

Note that if $S$ is a pure state corresponding to the vector $\psi$, then:

$$E(A) = \langle \psi | A | \psi \rangle.$$

**Von Neumann entropy**

Of particular significance for describing randomness of a state is the von Neumann entropy of $S$ formally defined by

$$H(S) = - \text{Tr}(S \log_2 S).$$

Actually, the operator $S \log_2 S$ is not necessarily trace-class. However, if $S$ is a non-negative self-adjoint operator not of trace class we define $\text{Tr}(S) = +\infty$. Also note that any density operator $S$ can be diagonalized, that it can be represented in some orthonormal basis by a (possibly infinite) matrix of the form

$$\begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 & \cdots \\
0 & \lambda_2 & \cdots & 0 & \cdots \\
& & \ddots & & \\
0 & 0 & \cdots & \lambda_n & \cdots \\
& & & \ddots & \\
& & & & \cdots & \\
\end{bmatrix}$$

and we define

$$H(S) = - \sum_i \lambda_i \log_2 \lambda_i.$$

The convention is that $0 \log_2 0 = 0$, since an event with probability zero should not contribute to the entropy. This value is an extended real number (that is in $[0, \infty]$) and this is clearly a unitary invariant of $S$. 
Remark. It is indeed possible that $H(S) = +\infty$ for some density operator $S$. In fact $T$ be the diagonal matrix

$$
T = \begin{bmatrix}
\frac{1}{2\log^2 2} & 0 & \cdots & 0 & \cdots \\
0 & \frac{1}{3\log^2 3} & \cdots & 0 & \cdots \\
\vdots & \vdots & \ddots & \vdots & \ddots \\
0 & 0 & \cdots & \frac{1}{n\log^2 n} & \cdots \\
\vdots & \vdots & \ddots & \vdots & \ddots
\end{bmatrix}
$$

$T$ is non-negative trace class and one can show $T \log_2 T$ is not trace-class.

Theorem. Entropy is a unitary invariant.

In analogy with classical entropy (notice the similarity in the definitions), $H(S)$ measures the amount of randomness in the state $S$. The more dispersed the eigenvalues are, the larger the system entropy. For a system in which the space $\mathcal{H}$ is finite-dimensional, entropy is maximized for the states $S$ which in diagonal form have the representation

$$
\begin{bmatrix}
\frac{1}{n} & 0 & \cdots & 0 \\
0 & \frac{1}{n} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{n}
\end{bmatrix}
$$

For such an $S$, $H(S) = \log_2 n$. The state $S$ is called the maximally mixed state. Recall that a pure state is one of the form

$$
S = |\psi\rangle \langle \psi|,
$$

for $\psi$ a vector of norm 1.

Theorem. $H(S) = 0$ if and only if $S$ is a pure state.

For $S$ is a pure state if and only if its diagonal form has exactly one non-zero entry which is a 1.
Entropy can be used as a measure of quantum entanglement.

**Gibbs canonical ensemble**

Consider an ensemble of systems described by a Hamiltonian $H$ with average energy $E$. If $H$ has pure-point spectrum and the eigenvalues $E_n$ of $H$ go to $+\infty$ sufficiently fast, $e^{-rH}$ will be a non-negative trace-class operator for every positive $r$.

The **Gibbs canonical ensemble** is described by the state

$$S = \frac{e^{-\beta H}}{\text{Tr}(e^{-\beta H})}.$$  

Where $\beta$ is such that the ensemble average of energy satisfies

$$\text{Tr}(SH) = E$$

and

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

This is called the **partition function**; it is the quantum mechanical version of the canonical partition function of classical statistical mechanics. The probability that a system chosen at random from the ensemble will be in a state corresponding to energy eigenvalue $E_m$ is

$$\frac{e^{-\beta E_m}}{\sum_n e^{-\beta E_n}}.$$

Under certain conditions, the Gibbs canonical ensemble maximizes the von Neumann entropy of the state subject to the energy conservation requirement.\[clarification needed\]

**Grand canonical ensemble**

Main article: grand canonical ensemble

For open systems where the energy and numbers of particles may fluctuate, the system is described by the **grand canonical ensemble**, described by the density matrix
\[ \rho = \frac{e^{-\beta (H + \mu_1 N_1 + \mu_2 N_2 + \cdots)}}{\text{Tr}(e^{-\beta (H + \mu_1 N_1 + \mu_2 N_2 + \cdots)})}, \]

where the \( N_1, N_2, \ldots \) are the particle number operators for the different species of particles that are exchanged with the reservoir. Note that this is a density matrix including many more states (of varying \( N \)) compared to the canonical ensemble.

The grand partition function is
\[ Z(\beta, \mu_1, \mu_2, \cdots) = \text{Tr}(e^{-\beta (H + \mu_1 N_1 + \mu_2 N_2 + \cdots)}) \]

References
