Quantum Mechanics Entropy and a Quantum Version of the H-Theorem

Paul Bracken
Department of Mathematics,
University of Texas, Edinburg, TX
USA

1. Introduction

Entropy is a fundamental concept which emerged along with other ideas during the development of thermodynamics and statistical mechanics Landau and Lifshitz (1978); Lieb and Yngvason (1999). Entropy has developed foremost out of phenomenological thermodynamical considerations such as the second law of thermodynamics in which it plays a prominent role Wehrl (1978). With the intense interest in the investigation of the physics of matter at the atomic and subatomic quantum levels, it may well be asked whether this concept can emerge out of the study of systems at a more fundamental level. In fact, it may be argued that a correct definition is only possible in the framework of quantum mechanics, whereas in classical mechanics, entropy can only be introduced in a rather limited and artificial way. Entropy relates macroscopic and microscopic aspects of nature, and ultimately determines the behavior of macroscopic systems. It is the intention here to present an introduction to this subject in a readable manner from the quantum point of view. There are many reasons for undertaking this. The intense interest in irreversible thermodynamics Grössing (2008), the statistical mechanics of astrophysical objects Padmanabhan (1990); Pathria (1977), quantum gravity and entropy of black holes Peres & al. (2004), testing quantum mechanics Ballentine (1970) and applications to condensed matter and quantum optics Haroche & al. (2006); Raimond & al. (2001) are just a few areas which are directly or indirectly touched on here.

Let us begin by introducing the concept of entropy from the quantum mechanical perspective, realizing that the purpose is to focus on quantum mechanics in particular. Quantum mechanics makes a clear distinction between observables and states. Observables such as position and momentum are mathematically described by self-adjoint operators in a Hilbert space. States, which are generally mixed, can be described by a density matrix, which is designated by ρ throughout. This operator ρ is Hermitean, has trace one and yields the expectation value of an observable A in the state ρ through the definition

$$\langle A \rangle = Tr\,(\rho A). \tag{1.1}$$

Entropy is not an observable, so there does not exist an operator with the property that its expectation value in some state would be the entropy. In fact, entropy is a function of state. If

the given state is described by the density matrix ρ, its entropy is defined to be

$$S(\rho) = -k_B \, Tr(\rho \log(\rho)). \tag{1.2}$$

This formula is due to von Neumann von Neumann (1955), and generalizes the classical expression of Boltzmann and Gibbs to the quantum regime. Of course, k_B is Boltzmann's constant, and the natural logarithm is used throughout. If k_B is put equal to one, the entropy becomes dimensionless. Thus, entropy is a well-defined quantity, no matter what size or type of system is considered. It is always greater than or equal to zero, and equal to zero exactly for pure states.

It will be useful to give some interpretation of von Neumann's formula. The discovery for which Boltzmann is remembered is his formula for entropy which appeared in 1877, namely,

$$S = k_B \log(W), \tag{1.3}$$

This form for S was established by making a connection between a generalization S of thermostatic entropy and the classical H-function. The identification of the constant on the right of (1.3) as Boltzmann's was proposed by Planck. Equation (1.3) taken in conjunction with the H-theorem, interprets the second law, $\Delta S \geq 0$, simply as the tendency of an isolated system to develop from less probable states to more probable states; that is, from small W to large W. Thermostatic equilibrium corresponds to the state in which W attains its maximum value. In fact, equation (1.3) has had far reaching consequences. It led Planck, for example, to his quantum hypothesis, which is that the energy of radiation is quantized, and then from there to the third law of thermodynamics. The H-theorem provided an explanation in mechanical terms of the irreversible approach of macroscopic systems towards equilibrium. By correlating entropy with the H-function and thermodynamic probability, Boltzmann revealed the statistical character of the second law. Of course, Boltzmann was restricted to a classical perspective. The question as to whether the number of microstates makes literal sense classically has been discussed as an objection to his approach. As stated by Pauli Pauli (2000), a microstate of a gas for example is defined as a set of numbers which specify in which cell each atom is located, that is, a number labeling the atom, an index for the cell in which atom s is located and a label for the microstate. The macrostate is uniquely determined by the microstate, however the converse does not hold. For every macrostate there are very many microstates, as will be discussed. Boltzmann's fundamental hypothesis is then: All microstates are equally probable.

However, as Planck anticipated, in quantum mechanics such a definition immediately makes sense. There is no ambiguity at all, as there is a natural idea of microstate. The number of microstates may be interpreted as the number of pure states with some prescribed expectation values. Suppose there are W different pure states in a system, each occuring with the same probability. Then the entropy is simply $S = \log(W)$. However, the density matrix of the system is given by $\rho = (1/W)\mathcal{P}$, where $\mathcal{P}$ is a W-dimensional projection operator. Thus, the correspondence follows immediately, that is, $\log W = -Tr\,[\rho \log \rho]$.

Each density matrix can be diagonalized Wehrl (1978),

$$\rho = \sum_k p_k \, |k\rangle \langle k|, \tag{1.4}$$

where $|k\rangle$ is a normalized eigenvector corresponding to the eigenvalue p_k and $|k\rangle\langle k|$ is a projection operator onto $|k\rangle$ with $p_k \geq 0$ and $\sum_k p_k = 1$. Here the coefficients are positive probabilities and not complex amplitudes as in a quantum mechanical superposition. Substituting (1.4) into (1.2) finally yields,

$$S(\rho) = -\sum_k p_k \log(p_k). \tag{1.5}$$

There is a more combinatorial approach Wehrl (1978). This will come up again subsequently when ensembles take the place of a density operator. If N measurements are performed, one will obtain as a result that for large N, the system is found $p_1 N$ times in $|1\rangle$, $p_2 N$ times in state $|2\rangle$ and so on, all having the same weight. By straightforward counting, there results

$$W_N = \frac{N!}{(p_1 N)!(p_2 N)! \cdots}. \tag{1.6}$$

When $N \to \infty$, Stirling's formula can be applied to the logarithm of (1.6) so the entropy is

$$\log \frac{N!}{n_1! n_2! \cdots} = N \log(N) - N - \sum_j (n_j \log n_j - n_j) = -N \sum_j p_j \log(p_j). \tag{1.7}$$

Dividing both sides of (1.7) by N, then as $N \to \infty$ (1.5) is recovered. It should also be noted that (1.5) is of exactly the same form as Shannon entropy, which can be thought of as a measure of unavailable information.

Of course, another way to look at this is to consider N copies of the same Hilbert space, or system, in which there are microstates $|1\rangle \otimes |2\rangle \cdots$ such that $|1\rangle$ occurs $p_1 N$ times, $|2\rangle$ occurs $p_2 N$ times, and so forth. Again (1.6) is the result, and according to Boltzmann's equation, one obtains $\log(W_N)$ for the entropy as in (1.5). In (1.5), S is maximum when all the p_j are equal to $1/N$.

By invoking the constraint $\sum_k p_k = 1$, (1.5) takes the form

$$S = -\sum_{k=1}^{N-1} p_k \log(p_k) - p_N \log(p_N), \tag{1.8}$$

where $p_N = 1 - \sum_{k=1}^{N-1} p_k$, and all other p_k are considered to be independent variables. Differentiating S in (1.8), it is found that

$$\frac{\partial S}{\partial p_k} = -\log(p_k) + \log(p_N).$$

This vanishes of course when $p_k = p_N = N^{-1}$ and this solution is the only extremum of S.

To summarize, entropy is a measure of the amount of chaos or lack of information about a system. When one has complete information, that is, a pure state, the entropy is zero. Otherwise, it is greater than zero, and it is bigger the more microstates exist and the smaller their statistical weight.

2. Basic properties of entropy

There are several very important properties of entropy function (1.5) which follow from simple mathematical considerations and are worth introducing at this point Peres (1995).

The first point to make is that the function $S(\mathbf{p})$ is a concave function of its arguments $\mathbf{p} = (p_1, \cdots, p_N)$. For any two probability distributions $\{p_j\}$ and $\{q_j\}$, and any $\lambda \in [0,1]$, S defined in (1.5) satisfies the following inequality

$$S(\lambda \mathbf{p} + (1 - \lambda \mathbf{q})) \geq \lambda S(\mathbf{p}) + (1 - \lambda)S(\mathbf{q}), \qquad \lambda \in [0,1]. \tag{2.1}$$

This can be proved by differentiating S twice with respect to λ to obtain,

$$\frac{d^2 S(\lambda \mathbf{p} + (1 - \lambda)\mathbf{q})}{d\lambda^2} = -\sum_j \frac{(p_j - q_j)^2}{\lambda p_j + (1 - \lambda)q_j} \leq 0. \tag{2.2}$$

This is a sufficient condition for a function to be concave. Equality holds only when $p_j = q_j$, for all j. The physical meaning of inequality (2.1) is that mixing different probability distributions can only increase uniformity.

If N is the maximum number of different outcomes obtainable in a test of a given quantum system, then any test that has exactly N different outcomes is called a maximal test, called T here. Suppose the probabilities p_m for the outcomes of a maximal test T which can be performed on that system are given. It can be shown that this entropy never decreases if it is elected to perform a different maximal test. The other test may be performed either instead of T, or after it, if test T is repeatable.

To prove this statement, suppose the probabilities for test T are $\{p_m\}$ and those for a subsequent test are related to the $\{p_m\}$ by means of a doubly stochastic matrix $P_{\mu m}$. This is a matrix which satisfies $\sum_\mu P_{\mu m} = 1$ and $\sum_m P_{\mu m} = 1$. In this event,

$$q_\mu = \sum_m P_{\mu m} p_m$$

are the probabilities for the subsequent test. The new entropy is shown to satisfy the inequality $S(\mathbf{q}) \geq S(\mathbf{p})$. To prove this statement, form the difference of these entropies based on (1.5),

$$\sum_m p_m \log(p_m) - \sum_\mu q_\mu \log(q_\mu) = \sum_m p_m (\log(p_m) - \sum_\mu P_{\mu m} \log(q_\mu))$$

$$= \sum_{m\mu} p_m (P_{\mu m} \log(p_m) - P_{\mu m} \log(q_\mu))$$

$$= \sum_{m\mu} p_m P_{\mu m} \log\left(\frac{p_m}{q_\mu}\right).$$

In the second line, $\sum_\mu P_{\mu m} = 1$ has been substituted to get this result. Using the inequality $\log x \geq 1 - x^{-1}$, where equality holds when $x = 1$, and the fact that S has a negative sign, it follows that

$$S(\mathbf{q}) - S(\mathbf{p}) \geq \sum_{m\mu} p_m P_{\mu m}(1 - \frac{q_\mu}{p_m}) = \sum_{m\mu}(p_m P_{\mu m} - q_\mu P_{\mu m}) = \sum_\mu (q_\mu - q_\mu \sum_m P_{\mu m}) = 0. \tag{2.3}$$

The equality sign holds if and only if $P_{\mu m}$ is a permutation matrix, so the sets are identical.

After a given preparation whose result is represented by a density matrix ρ, different tests correspond to different sets of probabilities, and therefore to different entropies. The entropy of a preparation can be defined as the lowest value attained by (1.5) for any complete test performed after that preparation. The optimal test which minimizes S is shown to be the one that corresponds to the orthonormal basis v_μ given by the eigenvectors of the density matrix ρ

$$\rho\, v_\mu = w_\mu v_\mu.$$

In this basis, ρ is diagonal and the eigenvalues w_μ satisfy $0 \le w_\mu \le 1$ and $\sum_\mu w_\mu = 1$.

A basic postulate of quantum mechanics asserts that the density matrix ρ completely specifies the statistical properties of physical systems that were subjected to a given preparation. All the statistical predictions that can be obtained from (1.1) for an operator are the same as if we had an ordinary classical mixture, with a fraction w_μ of the systems with certainty in the state v_μ. Therefore, if the maximal test corresponding to the basis v_μ is designed to be repeatable, the probabilities w_μ remain unchanged and entropy S remains constant. The choice of any other test can only increase the entropy, as in the preceding result. This proves that the optimal test, which minimizes the entropy, is the one corresponding to the basis that diagonalizes the density matrix.

The entropic properties of composite systems obey numerous inequalities as well. Let $\{v_m\}$ and $\{e_\mu\}$ be two orthonormal basis sets for the same physical system. Let $\rho = \sum w_m |v_m\rangle\langle v_m|$ and $\sigma = \sum \omega_\mu |e_\mu\rangle\langle e_\mu|$ be two different density matrices. Their relative entropy $S(\sigma|\rho)$ is defined to be

$$S(\sigma|\rho) = Tr[\rho(\log \rho - \log \sigma)]. \tag{2.4}$$

Let us evaluate $S(\sigma|\rho)$ in (2.3) in the $|v_\mu\rangle$ basis where ρ is diagonal. The diagonal elements of $\log \sigma$ are

$$(\log \sigma)_{mm} = \langle v_m, \sum_\mu \log \omega_\mu |e_\mu\rangle\langle e_\mu| v_m\rangle = \sum_\mu \log \omega_\mu |\langle e_\mu, v_m\rangle|^2 = \sum_\mu \log \omega_\mu\, P_{\mu m}. \tag{2.5}$$

The matrix $P_{\mu m}$ is doubly stochastic, so as in (2.3) we have

$$S(\sigma|\rho) = \sum_m w_m(\log w_m - \sum_\mu P_{\mu m} \log \omega_\mu) = \sum_{\mu m} w_m P_{\mu m} \log(\frac{w_m}{\omega_\mu}) \ge 0. \tag{2.6}$$

Equality holds in (2.6) if and only if $\sigma = \rho$.

Inequality (2.6) can be used to prove a subadditivity inequality. Consider a composite system, with density matrix ρ, then the reduced density matrices of the subsystems are called ρ_1 and ρ_2. Then matrices ρ, ρ_1 and ρ_2 satisfy,

$$S(\rho) \le S(\rho_1) + S(\rho_2). \tag{2.7}$$

This inequality implies that a pair of correlated systems involves more information than the two systems separately.

To prove this, suppose that w_m, ω_μ and $W_{m\mu} = w_m \omega_\mu$ are the eigenvalues of ρ_1, ρ_2 and $\rho_1 \otimes \rho_2$, respectively, then

$$\sum_m w_m \log w_m + \sum_\mu \omega_\mu \log \omega_\mu = \sum_{m\mu} W_{m\mu} \log W_{m\mu}.$$

This has the equivalent form,

$$S(\rho_1) + S(\rho_2) = S(\rho_1 \otimes \rho_2).$$

Consider now the relative entropy

$$S(\rho_1 \otimes \rho_2|\rho) = Tr[\rho(\log\rho - \log\rho_1 \otimes \rho_2)] = Tr[\rho(\log\rho - \log\rho_1 - \log\rho_2)].$$

It has just been shown that relative entropy is nonnegative, so it follows from this that

$$Tr(\rho\log\rho) \geq Tr(\rho\log\rho_1) + Tr(\rho\log\rho_2).$$

Since $Tr(\rho\log\rho_1) = \sum_{m\mu n\nu} \rho_{m\mu,n\nu}(\log\rho_1)_{nm}\delta_{\nu\mu} = Tr(\rho_1\log\rho_1)$, and similarly for $Tr(\rho\log\rho_2)$, it follows that

$$Tr(\rho\log\rho) \geq Tr(\rho_1\log\rho_1) + Tr(\rho_2\log\rho_2).$$

Now using (1.2), it follows that

$$-S(\rho) \geq -S(\rho_1) - S(\rho_2).$$

Multiplying both sides by minus one, (2.7) follows.

3. Entanglement and entropy

The superposition principle applied to composite systems leads to the introduction of the concept of entanglement Mintet & al. (2005); Raimond & al. (2001), and provides an important application for the density matrix. A very simple composite object is a bipartite quantum system S which is composed of two parts A and B. The states of A and B belong to two separate Hilbert spaces called $\mathcal{H}_A$ and $\mathcal{H}_B$ which are spanned by the bases $|i_A\rangle$ and $|i_B\rangle$, and may be discrete or continuous. If A and B are prepared independently of each other and are not coupled together at some point, S is described by the tensor product $|\psi_S\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. Each subsystem is described by a well-defined wave function. Any manipulation of one part leaves the measurement prediction for the other part unchanged. System S can also be prepared by measuring joint observables, which act simultaneously on A and B. Even if S has been prepared by measuring separate observables, A and B can become coupled by means of an interaction Hamiltonian. In this instance, it is generally impossible to write the global state $|\psi_S\rangle$ as a product of partial states associated to each component of S.

This is what the expression *quantum entanglement* means. The superposition principle is at the heart of the most intriguing features of the microscopic world. A quantum system may exist in a linear superposition of different eigenstates of an observable, suspended between different classical realities, as when one says a particle can be at two positions at the same time. It seems to be impossible to get a classical intuitive representation of superpositions. When the superposition principle is applied to composite systems, it leads to the concept of entanglement. Moreover, as Bell has shown, entanglement cannot be consistent with any local theory containing hidden variables.

Even if the state S cannot be factorized according to the superposition principle, it can be expressed as a sum of product states $|i_A\rangle \otimes |\mu_B\rangle$, which make up a basis of the global Hilbert

space, $\mathcal{H}_S$. Consequently, an entangled state can be expressed as

$$|\psi_S\rangle = \sum_{i,\mu} \alpha_{i\mu} |i_A\rangle \otimes |\mu_B\rangle \neq |\psi_A\rangle \otimes |\psi_B\rangle, \tag{3.1}$$

where the $\alpha_{i\mu}$ are complex amplitudes. The states $|\psi_S\rangle$ contain information not only about the results of measurements on A and B separately, but also on correlations between these measurements. In an entangled state, each part loses its quantum identity. The quantum content of the global state is intricately interwoven between the parts. Often it is the case that there is interest in carrying out measurements on one part without looking at another part. For example, what is the probability of finding a result when measuring observable O_A attached to subsystem A, without worrying about B. The complete wave function $|\psi_S\rangle$ can be used to predict the experimental outcomes of the measurement of $O_A \otimes \mathbf{1}_B$. This can also be done by introducing the density operator ρ_S of a system described by the quantum state $|\psi_S\rangle$, which is just the projector

$$\rho_S = |\psi_S\rangle\langle\psi_S|. \tag{3.2}$$

It has the same information content as $|\psi_S\rangle$, and for all predictions on S, all quantum rules can be expressed in such a fashion; for example, the expectation values of an observable O_S of S is found by (1.1). The probability of finding the system in $|i\rangle$ after a measurement corresponding to the operator $\rho_i = |i\rangle\langle i|$ is given by $|\langle i|\psi_S\rangle|^2$ in the quantum description and $Tr(\rho_i\rho_S)$ in terms of the density matrix.

The density operator approach is very advantageous for describing one subsystem, A, without looking at B. A partial density operator ρ_A can be determined which has all the predictive information about A alone, by tracing ρ_S over the subspace of B

$$\rho_A = Tr_B(\rho_S) = \sum_{i,i',\mu} \alpha_{i\mu}\alpha^*_{i'\mu}|i_A\rangle\langle i_A|. \tag{3.3}$$

Thus, the probability of finding A in state $|j_A\rangle$ is found by computing the expectation value of the projector $\rho_j = |j_A\rangle\langle j_A|$, which is $\pi_j = Tr(\rho_A\rho_j)$. Predictions on A can be done without considering B. The information content of ρ_A is smaller than in ρ_S, since correlations between A and B are omitted. To say that A and B are entangled is equivalent to saying that ρ_A and ρ_B are not projectors on a quantum state. There is however a basis in $\mathcal{H}_A$ in which ρ_A is diagonal. Let us call it $|j_A\rangle$, so that ρ_A is given by

$$\rho_A = \sum_j \lambda_j |j_A\rangle\langle j_A|. \tag{3.4}$$

In (3.4), λ_j are positive or zero eigenvalues which sum to one. By neglecting B, there is acquired only a statistical knowledge of state A, with a probability λ_j of finding it in $|j_A\rangle$.

It is possible to express the state for S in a representation which displays the entanglement. The superposition (3.1) claims nothing as to whether the state can be factored. To put this property in evidence, choose a basis in $\mathcal{H}_A$, called $|j_A\rangle$ in which ρ_A is diagonal. Then (3.1) is written

$$|\psi_S\rangle = \sum_j |j_A\rangle|\tilde{j}_B\rangle, \tag{3.5}$$

where state $|\tilde{j}_B\rangle$ is given by

$$|\tilde{j}_B\rangle = \sum_{\mu} \alpha_{j\mu} |\mu_B\rangle. \tag{3.6}$$

The $|\tilde{j}_B\rangle$ are mirroring in $\mathcal{H}_B$ the basis of orthonormal states in $\mathcal{H}_A$ in which ρ_A is diagonal. These mirror states are also orthogonal to each other as can be seen by expressing the fact that ρ is diagonal

$$\langle j_A | \rho_A | j'_A \rangle = \lambda_j \delta_{jj'} = \langle \tilde{j}_B | \tilde{j}'_B \rangle.$$

At this point, the mirror state can be normalized by means of the transformation $|\hat{j}_B\rangle = |j_B\rangle / \sqrt{\lambda_j}$ giving rise to the Schmidt expansion,

$$|\psi_S\rangle = \sum_{j} \sqrt{\lambda_j} |j_A\rangle |\hat{j}_B\rangle. \tag{3.7}$$

The sum over a basis of product mirror states exhibits clearly the entanglement between A and B. The symmetry of this expression shows that ρ_A and ρ_B have the same eigenvalues. Any pure entangled state of a bipartite system can be expressed in this way.

Now a measure of the degree of entanglement can be defined using the density matrix. As the λ_j become more spread out over many non-zero values, more information is lost by concentrating on one system and disregarding correlations between A and B. This loss of mutual information can be linked to the degree of entanglement. This information loss could be measured by calculating the von Neumann entropy of A or B from (1.5)

$$S_A = S_B = - \sum_{j} \lambda_j \log(\lambda_j) = -Tr(\rho_A \log \rho_A) = -Tr(\rho_B \log \rho_B). \tag{3.8}$$

This is the entropy of entanglement $S_e = S_A = S_B$, and it expresses quantitatively the degree of disorder in our knowledge of the partial density matrices of the two parts of the entangled system S.

If the system is separable, then one λ_j is non-zero and $S_e = 0$, so maximum information on the states of both parts obtains. As soon as two λ_j are non-zero, S_e becomes strictly positive and A and B are entangled. The maximum entropy, hence maximum entanglement obtains when the λ_j are equally distributed among the A and B subspaces. It is maximal and equal to $\log N_A$, when ρ_A is proportional to $\mathbf{1}_A$, that is $\rho_A = \mathbf{1}_A / N_A$. In a maximally entangled state, local measurements performed on one part of the system are not predictable at all. What can be predicted are the correlations between the measurements performed on both parts. For example, consider a bipartite system in which one part has dimension two. There are only two λ-values in the Schmidt expansion, and satisfy $\lambda_1 + \lambda_2 = 1$. Then from (1.5), the entropy when $\lambda_1 \in (0,1)$ is,

$$S_e = -\lambda_1 \log(\lambda_1) - (1 - \lambda_1) \log(1 - \lambda_1). \tag{3.9}$$

The degree of entanglement is equal to zero when $\lambda_1 = 0$ or 1 and passes through a maximum at $\lambda_1 = 1/2$ at which $S_e = 1$. The degree of entanglement measured by the von Neumann entropy is invariant under local unitary transformations acting on A or B separately, a direct consequence of the invariance of the spectrum of the partial density operators.

Consider the case of a two-level system with states $|0\rangle$ and $|1\rangle$, where the density matrix is a two-by-two hermitean matrix given by

$$\rho_A = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}. \tag{3.10}$$

The entropy can be calculated for this system. Its positive diagonal terms are the probabilities of finding the system in $|0\rangle$ or $|1\rangle$ and they sum to one. The nondiagonal terms satisfy $\rho_{01} = \rho_{10}^*$ and are zero for a statistical mixture of $|0\rangle$ and $|1\rangle$. Since ρ_A is a positive operator

$$|\rho_{10}| = |\rho_{01}| \leq \sqrt{\rho_{00}\rho_{11}}.$$

is satisfied, and the upper bound is reached for pure states.

The density matrix ρ_A can be expanded with real coefficients onto the operator basis made up of the identity matrix $\mathbf{I}$ and the Pauli matrices $\boldsymbol{\sigma}_i$

$$\rho_A = \frac{1}{2}(\mathbf{I} + \mathbf{R} \cdot \boldsymbol{\sigma}), \tag{3.11}$$

where $\mathbf{R} = (u, v, w)$ is three-dimensional and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. The components of $\mathbf{R}$ are linked to the elements of the density matrix as follows

$$u = \rho_{10} + \rho_{01}, \qquad v = i(\rho_{01} - \rho_{10}), \qquad w = \rho_{00} - \rho_{11}.$$

The modulus R of $\mathbf{R}$ satisfies $R \leq 1$, equality holding only for pure states. This follows from $Tr(\rho_A^2) \leq 1$. If nonlinear functions of an observable A are defined as $f(A) = \sum f(a_k)|e_k\rangle\langle e_k|$, the von Neumann entropy of ρ is

$$S = -\frac{1+R}{2}\log(\frac{1+R}{2}) - \frac{1-R}{2}\log(\frac{1-R}{2}). \tag{3.12}$$

To each density matrix ρ_A, the end of the vector $\mathbf{R}$ can be located on the surface of a sphere. The surface of the sphere $R = 1$ is the set of pure states with $S = 0$. The statistical mixtures correspond to inside the sphere $R < 1$. The closer the point to the center, the larger the von Neumann entropy. The center of the sphere corresponds to the totally unpolarized maximum entropy state.

Any mixed state can be represented in an infinite number of ways as a statistical mixture of two pure states, since any $\mathbf{P}$ with its end inside the sphere can be expressed as a vector sum of a $\mathbf{P}_1$ and $\mathbf{P}_2$ whose ends are at the intersection of the sphere with an arbitrary line passing by the extreme end of $\mathbf{P}$, so one can write $\mathbf{P} = \lambda\mathbf{P}_1 + (1-\lambda)\mathbf{P}_2$ for $0 < \lambda < 1$. The density matrix which is a linear function of $\mathbf{P}$ is then a weighted sum of the projectors on the pure states $|u_1\rangle$ and $|u_2\rangle$ corresponding to $\mathbf{P}_1$ and $\mathbf{P}_2$,

$$\rho_A = \frac{1}{2}[\mathbf{I} + \lambda\mathbf{P}_1 \cdot \boldsymbol{\sigma} + (1-\lambda)\mathbf{P}_2 \cdot \boldsymbol{\sigma}] = \lambda|u_1\rangle\langle u_1| + (1-\lambda)|u_2\rangle\langle u_2|. \tag{3.13}$$

Thus, there exists an ambiguity of representation of the density operator which, if $P \neq 0$, can be lifted by including the condition that $|u_1\rangle$ and $|u_2\rangle$ be orthogonal.

Before finishing, it is worth discussing the following application, which seems to have very important ramifications. A violation of the second law arises if nonlinear modifications are

introduced into Schrödinger's equation Weinberg (1989). A nonlinear Schrödinger equation does not violate the superposition principle in the following sense. The principle asserts that the pure states of a physical system can be represented by rays in a complex linear space, but does not demand that the time evolution obeys a linear equation. Nonlinear variants of Schrödinger's equation can be created with the property that if $u(0)$ evolves to $u(t)$ and $v(0)$ to $v(t)$, the pure state represented by $u(0) + v(0)$ does not evolve into $u(t) + v(t)$, but into some other pure state.

The idea here is to show that such a nonlinear evolution violates the second law of thermodynamics. This is provided the other postulates of quantum mechanics remain as they are, and that the equivalence of the von Neumann entropy to ordinary entropy is maintained. Consider a mixture of quantum systems which are represented by a density matrix

$$\rho = \lambda \Pi_u + (1 - \lambda) \Pi_v, \tag{3.14}$$

where $0 < \lambda < 1$ and Π_u, Π_v are projection operators on the pure states u and v. In matrix form the density matrix is represented as

$$\rho = \begin{pmatrix} \lambda & \lambda \langle v | u \rangle \\ (1 - \lambda)\langle u | v \rangle & 1 - \lambda \end{pmatrix}.$$

The eigenvalues are found by solving the polynomial $\det(\rho - w\mathbf{1}) = 0$ for the eigenvalues w. Setting $x = |\langle u, v \rangle|^2$, they are given by

$$w_j = \frac{1}{2} \pm [\frac{1}{4} - \lambda(\lambda - 1)(1 - x)]^{1/2}, \qquad j = 1, 2. \tag{3.15}$$

The entropy of this mixture is found by putting w_j into (1.5)

$$S = -w_1 \log(w_1) - w_2 \log(w_2). \tag{3.16}$$

The polynomial $p(\lambda) = 4\lambda(1 - \lambda)$ has range $(0,1)$ when $\lambda \in (0,1)$, so it follows that $s = 4\lambda(1 - \lambda)(1 - x) \in (0,1)$ as well. Setting $f = 1 - s$, then when $s \in (0,1)$, the derivative of (3.16) is given by

$$\frac{\partial S}{\partial x} = -\frac{\lambda(1 - \lambda)}{\sqrt{1 - 4\lambda(1 - \lambda)(1 - x)}} \log\left(\frac{(1 + \sqrt{f})^2}{4\lambda(1 - \lambda)(1 - x)}\right)$$

$$= -\frac{\lambda(1 - \lambda)}{\sqrt{1 - s}} \log\left(\frac{(1 + \sqrt{1 - s})^2}{s}\right) < 0.$$

Consequently, if pure quantum states evolve as $u(0) \rightarrow u(t)$ and $v(0) \rightarrow v(t)$, the entropy of the mixture ρ shall not decrease provided that $x(t) \leq x(0)$, or in terms of the definition of x, $|\langle u(t), v(t) \rangle|^2 \leq |\langle u(0), v(0) \rangle|^2$. If say $\langle u(0), v(0) \rangle = 0$, then also $\langle u(t), v(t) \rangle = 0$, so orthogonal states remain orthogonal. Consider now a complete orthogonal set u_k. For every v,

$$\sum_k |\langle u_k, v \rangle|^2 = 1.$$

If there exists m such that $|\langle u_m(t), v(t) \rangle|^2 < |\langle u_m(0), v(0) \rangle|^2$, there must also exist some n for which the reverse holds, $|\langle u_n(t), v(t) \rangle|^2 > |\langle u_n(0), v(0) \rangle|^2$. In this event, the entropy of a

mixture of u_n and v will spontaneously decrease in a closed system, which is in violation of the second law of thermodynamics. To retain the law, $|\langle u(t), v(t)\rangle|^2 = |\langle u(0), v(0)\rangle|^2$ must hold for every u and v. From Wigner's theorem, the mapping $v(0) \rightarrow v(t)$ is unitary, so Schrödinger's equation must be linear if the other postulates of quantum mechanics remain fixed.

4. Ensemble methods in quantum mechanics

In classical mechanics, one relinquishes the idea of a description of the microscopic mechanical states of trillions of microscopic interacting particles by instead computing averages over a virtual ensemble of systems which replicate the real system. Quantum theory is faced with a similar problem, and the remedy takes the form of the Gibbs ensemble. This last section will take a slightly different track and discusses ensemble theory in quantum mechanics. Two of the main results will be to produce a quantum version of the H-Theorem, and to show how the quantum mechanical canonical ensemble can be formulated.

An astronomic number of states, or of microstates, is usually compatible with a given set of macroscopic parameters defining a macrostate of a thermophysical system. Consequently, a virtual quantum mechanical ensemble of systems is invoked, which is representative of the real physical system. The logical connection between a physical system and ensemble is made by requiring the time average of a mechanical property G of a system in thermodynamic equilibrium equal its ensemble average calculated with respect to an ensemble made up of $N^* \rightarrow \infty$ systems representing the actual system

$$\bar{G} = \langle G \rangle. \tag{4.1}$$

The ensemble average $\langle G \rangle$ is the ordinary mean of G over all the systems of the ensemble. If N_r^* systems are in a state with eigenvalue G_r corresponding to G,

$$N^* \langle G \rangle = \sum_r N_r^* G_r, \tag{4.2}$$

where the sum is over all allowed states.

Adopt as a basic set the states $\psi_{jrm\cdots}$ uniquely identifiable by the quantum numbers $j, r, m, \cdots$ referring to a set of compatible properties. A particular system of the ensemble will not permanently be in one of these states $\psi_{jrm\cdots}$, as there exists only a probability to find a system in any one. Let us compress the basic states to read ψ_{jr} if we let r stand for the entire collection of quantum numbers $r, m, \cdots$. These cannot strictly be eigenstates of the total energy, since a system occupying a particular eigenstate of its total Hamiltonian H at any one moment will remain in this state forever. The state of the real system, which the ensemble is to represent, is a superposition of eigenstates belonging to the same or different values of the energy. To obtain an ensemble where the individual members are to change, we suppose the basic set ψ_{jr} is made up of eigenstates of the unperturbed Hamiltonian H^0. Assume it is possible to write

$$H = H^0 + H^1, \tag{4.3}$$

such that H^1 is a small perturbation added to the unperturbed Hamiltonian H^0, and vary with the physical system considered.

Suppose E_j^0 are the eigenvalues of the unperturbed H^0 and ψ_{jr}^0 the eigenstates corresponding to them, where r again denotes a set of compatible quantum numbers. Introducing H^1 now changes the energy eigenvalues and energy eigenfunctions by an amount E_{jr}^1 and ψ_{jr}^1, which should be very small compared with the unperturbed values. It is precisely the eigenstates ψ_{jr}^0 of H^0 rather than H that are used as basic states for the construction of the ensemble. Since these for the most part will appear in what follows, we continue to omit the superscript for both the eigenfunctions ψ_{jr} and eigenvalues E_{jr} whenever the situation indicates that unperturbed quantities are intended. A perturbed system finding itself initially in any one of the unperturbed states ψ_{jr} does not remain indefinitely in this state, but will continually undergo transitions to other unperturbed states ψ_{ks} due to the action of the perturbation H^1. In analogy with a classical system, a quantum ensemble is described by the number of systems N_{jr}^* in each state ψ_{jr}. The probability P_{jr} of finding a system, selected at random from the ensemble, in the state ψ_{jr} is clearly

$$P_{jr} = \frac{N_{jr}^*}{N^*}. \tag{4.4}$$

The quantities N_{jr}^* must sum up to N^*,

$$\sum_{jr} N_{jr}^* = N^*, \qquad \sum_{jr} P_{jr} = 1. \tag{4.5}$$

An ensemble can be representative of a physical system in thermodynamic equilibrium only in this context if the occupation numbers N_{jr}^* are constants. A more general picture could consider the occupation numbers as functions of time $N_{jr}^* = N_{jr}^*(t)$. The ensemble corresponds to a system removed from equilibrium. Let us ask then how do the N_{jr}^* vary with time.

Quantum mechanics claims the existence of $A_{ks}^{jr}(t)$ which determine the probability of a system in state ψ_{jr} at time zero to be in ψ_{ks} at time t. The final state could correspond to the initial state. Since $N_{jr}^*(0)$ systems are in a state specified by quantum numbers jr at $t = 0$, $A_{ks}^{jr}(t)N_{jr}^*(0)$ systems will make the transition from jr to ks during $(0, t)$ The number of systems in ks at time t will be

$$N_{ks}^*(t) = \sum_j \sum_r A_{ks}^{jr}(t) \, N_{jr}^*(0). \tag{4.6}$$

The $A_{ks}^{jr}(t)$ must satisfy the condition $\sum_j \sum_r A_{ks}^{jr}(t) = 1$. Multiplying this by $N_{ks}^*(0)$ and subtracting from (4.6) gives

$$N_{ks}^*(t) - N_{ks}^*(0) = \sum_j \sum_r A_{ks}^{jr}(t)[N_{jr}^*(0) - N_{ks}^*(0)]. \tag{4.7}$$

This is the change in occupation number over $(0, t)$. Dividing (4.7) by N^* and using (4.4) gives

$$P_{ks}(t) - P_{ks}(0) = \sum_j \sum_r A_{ks}^{jr}[P_{jr}(0) - P_{ks}(0)]. \tag{4.8}$$

A stationary ensemble or one in statistical equilibrium defined as $N_{ks}^*(t) = N_{ks}^*(0)$ for all ks holds when $N_{jr}^*(0) = N_{ks}^*(0)$, at least when $A_{ks}^{jr}(t) \neq 0$. The contribution to the right side of (4.8) comes from an extremely narrow interval $\Delta E = 2\hbar/t$ centered at $E_j = E_k$, as indicated by

perturbation theory. In this interval, it can be assumed $P_{jr}(0)$ depends on the j-index weakly enough that we can use $P_{ks}(0)$ in their place, so the term in brackets in (4.8) does not depend on j. The energy spectrum is very nearly continuous for a thermophysical system, so the sum over j can be approximated by an integral over E. This implies an approximation of the form

$$\sum_j A_{ks}^{jr}(t) = t W_{sr}^{(k)}. \tag{4.9}$$

The quantities $W_{sr}^{(k)}$ are time independent provided H^1 is time independent. Consequently, they are nonnegative and depend only on the displayed indices. Substituting (4.9) and $P_{jr}(0) = P_{kr}(0)$ into (4.8) gives

$$\frac{1}{t}[P_{ks}(t) - P_{ks}(0)] = \sum_r W_{sr}^{(k)}[P_{kr}(0) - P_{ks}(0)]. \tag{4.10}$$

In the limit when t becomes arbitrarily small, (4.10) can be approximated by expanding about $t = 0$ on the left to give the final result for the time rate of change of the probability P_{ks},

$$\dot{P}_{ks} = \sum_r W_{sr}^{(k)}[P_{kr}(0) - P_{ks}(0)]. \tag{4.11}$$

This equation was first derived by *W. Pauli*, and will lead to a quantum version of the *H*-Theorem next. It signifies that of the $N^* P_{kr}(0)$ systems occupying state kr at $t = 0$, $N^* P_{kr}(0) W_{sr}^{(k)}$ will, per unit time, go over to ks. Thus, the $W_{ks}^{(k)}$ are interpreted as transition probabilities per unit time that the system will go from state kr to ks. They must satisfy $W_{sr}^{(k)} \geq 0$ and the symmetry conditions $W_{rs}^{(k)} = W_{sr}^{(k)}$. This is also referred to as the principle of microscopic reversibility.

4.1 A quantum H-theorem

The ensemble which represents a real physical system is determined by the thermodynamic state and environment of the actual system. The virtual ensemble has constituents which must duplicate both aspects. Of great practical interest and the one considered here is the case of isolated systems. An isolated system is characterized not only by a fixed value of the energy E, but also by a definite number of particles and volume V. Under these conditions, a quantum H-theorem can be formulated Yourgrau et al. (1966). Classically the error with which the energy of the real system can be specified can be theoretically reduced to zero. However, quantum theory claims there is a residual error specified by the uncertainty relation. All members of the ensemble cannot be said then to occupy eigenstates belonging to the same energy. It must be assumed the systems are distributed over energy levels lying within a finite range, ΔE. The following restrictions on the occupation numbers of the ensemble are imposed for an isolated system

$$N_{jr}^* \neq 0, \qquad E_j \in I_{\Delta E} = (E - \frac{1}{2}\Delta E, E + \frac{1}{2}\Delta E), \qquad N_{jr}^* = 0, \qquad E_j \notin I_{\Delta E}. \tag{4.12}$$

It will be shown that the ensemble specified by (4.12) exhibits a one-directional development in time ending ultimately in equilibrium.

Pauli's equation can be used to obtain the rate of change of the quantum mechanical H-function which is defined to be

$$H^*(t) = \sum_s P_s \log(P_s). \tag{4.13}$$

The summation in (4.13) is extended over the group of states whose energies are approximately E, that is in the interval $I_{\Delta E}$ for example. Now, differentiate (4.13) with respect to t and use the fact that $\sum_s P_s = 1$ to get

$$\dot{H}^* = \sum_s \dot{P}_s \log(P_s) + \sum_s \dot{P}_s = \sum_s \dot{P}_s \log(P_s). \tag{4.14}$$

Now requiring that $\dot{P}_s$ be determined by (4.11), $\dot{H}^*$ in (4.14) becomes

$$\dot{H}^*(t) = \sum_s \sum_r W_{sr}(P_r - P_s) \log(P_s). \tag{4.15}$$

Interchanging r and s and using the symmetry property $W_{rs} = W_{sr}$, this is

$$\dot{H}^*(t) = \sum_r \sum_s W_{rs}(P_s - P_r) \log(P_r) = -\sum_r \sum_s W_{sr}(P_r - P_s) \log(P_r). \tag{4.16}$$

Adding (4.15) and (4.16) yields the following result,

$$\dot{H}^*(t) = -\frac{1}{2} \sum_r \sum_s W_{sr}(P_r - P_s)(\log(P_r) - \log(P_s)). \tag{4.17}$$

Recalling that $W_{sr} \geq 0$ as well as the inequality $(x - y)(\log x - \log y) \geq 0$ for each (r, s), it follows that $(P_r - P_s)(\log(P_r) - \log(P_s)) \geq 0$. Consequently, each term in the sum in (4.17) is either zero or positive, hence $H^*(t)$ decreases monotonically with time,

$$\dot{H}^*(t) \leq 0. \tag{4.18}$$

Equality holds if and only if $P_s = P_r$ for all pairs (r, s) such that $W_{sr} \neq 0$. Thus H^* decreases and statistical equilibrium is reached only when this condition is fulfilled. Originally enunciated by Boltzmann in a classical context, (4.18) constitutes a quantum mechanical version of the H-theorem.

4.2 Quantum mechanical canonical ensemble

Let us devise an ensemble which is representative of a closed isothermal system of given volume, or characterized by definite values of the parameters T, V and N. This approach brings us back to one of the ways entropy was formulated in the introduction, and need not rely on the specification of a density matrix. Suppose there are N^* members of the ensemble each with the same values of V and N as the real system. However, they are not completely isolated from each other, so each is surrounded by a surface that does not permit the flow of particles but is permeable to heat. The collection of systems can be packed into the form of a lattice and the entire construction immersed in a heat reservoir at temperature T until equilibrium is attained. The systems are isothermal such that each is embedded in a heat reservoir composed of the remaining $N^* - 1$.

Once the ensemble is defined, it can be asked which fraction of the N^* systems occupies any particular eigenstate of the unperturbed Hamiltonian of the experimental system. Let us study the ensemble then which is regarded as a large thermophysical system having energy E^*, volume $V^* = N^*V$ and made up of N^*N particles. The quantum states of this large supersystem belonging to energy E^* are to be enumerated. The thermal interaction energy is assumed to be so small that a definite energy eigenstate can be assigned to each individual system at any time. As energy can be exchanged between constituent systems, the eigenstates accessible to them do not pertain to one value of energy. The energy eigenstates of a system are written $E_1, E_2, \cdots, E_j, \cdots$ with $E_{j+1} \geq E_j$. Only one system-state j belongs to energy eigenvalue E_j. An energy eigenstate of the supersystem is completely defined once the energy eigenstate occupied by each system is specified.

It is only needed to stipulate the number N_j^* of systems occupying every system state j. Any set of values of the occupation numbers $N_1^*, N_2^*, \cdots$ define a quantum mechanical distribution. Clearly the W^* supersystem states calculated by

$$W(N_1^*, N_2^*, \cdots) = \frac{N^*!}{N_1^*! N_2^*! \cdots} \tag{4.19}$$

are compatible with a given distribution $N_1^*, N_2^*, \cdots$. Not all sets of N_j^* are admissible. The physically relevant ones satisfy the two constraints

$$\sum_j N_j^* = N^*, \qquad \sum_j N_j^* E_j = E^*. \tag{4.20}$$

The supersystem then consists of a number N^* of fixed but arbitrary systems with a constant energy E^*.

The number of physically possible supersystem states is clearly given as

$$\Omega^*(E^*, N^*) = \sum_C W^*(N_1^*, N_2^*, \cdots), \tag{4.21}$$

where the summation is to be extended over all N_j^* satisfying constraints (4.20). According to the earlier postulate, all allowed quantum states of an isolated system are equiprobable. Consequently, from this principle all states which satisfy (4.20) occur equally often. The probability P^* that a particular distribution $N_1^*, N_2^*, \cdots$ is actualized is the quotient of W^* and Ω^*,

$$P^*(N_1^*, N_2^*, \cdots) = \frac{W^*(N_1^*, N_2^*, \cdots)}{\Omega^*(E^*, N^*)}. \tag{4.22}$$

With respect to P^* in (4.22), the average value of the occupation number N_k^* is given quite simply by

$$\bar{N}_k^* = \sum_k N_k^* P^*(N_1^*, N_2^*, \cdots). \tag{4.23}$$

Substituting P^* into (4.23), it can be written as

$$\Omega^*(E^*, N^*)\bar{N}_k^* = N^* \sum_k \frac{(N^* - 1)!}{N_1^*! N_2^*! \cdots (N_k^* - 1)! \cdots}. \tag{4.24}$$

To obtain a more useful expression for $\bar{N}_k^*$, the right-hand side can be transformed to a set of primed integers. To this end, define

$$N^{*'} = N^* - 1, \qquad N_k^{*'} = N_k^* - 1, \qquad N_j^{*'} = N_j^*, \qquad j \neq k.$$

Using these, constraints (4.20) get transformed into

$$\sum N_j^{*'} = N_1^* + N_2^* + \cdots + (N_j^* - 1) + \cdots = N^* - 1,$$
$$\sum N_j^{*'} E_j = N_1^* E_1 + \cdots + (N_k^* - 1) E_k + \cdots = E^* - E_k. \tag{4.25}$$

Consequently,

$$\Omega^*(E^*, N^*) = N^* \sum{}' \frac{N^{*'}!}{N_1^{*'}! N_2^{*'}! \cdots}, \tag{4.26}$$

where the prime means the sum extends over all $N_k^{*'}$ which satisfy constraints (4.25). Comparing (4.26) with (4.20), the right-hand side of (4.26) is exactly $N^* \Omega^*(E^* - E_k, N^* - 1)$. Thus dividing by $\Omega^*(E^*, N^*)$, we have

$$\bar{N}_k^* = \frac{N^* \Omega^*(E^* - E_k, N^* - 1)}{\Omega^*(E^*, N^*)}.$$

Dividing this by N^* and taking the logarithm of both sides results in the expression,

$$\log \frac{\bar{N}_k^*}{N^*} = \log \Omega^* (E^* - E_k, N^* - 1) - \log \Omega^*(E^*, N^*). \tag{4.27}$$

The result in (4.27) can be expanded in a Taylor series to first order if we take $N^* \gg 1$ and $E^* \gg E_k$,

$$\log \frac{\bar{N}_k^*}{N^*} = -\frac{\partial \log \Omega^*}{\partial E^*} E_k^* - \frac{\partial \log \Omega^*}{\partial N^*} = -\beta E_k^* - \alpha. \tag{4.28}$$

From the constraint $N^* = \sum_j \bar{N}_j^* = N^* e^{-\alpha} \sum e^{-\beta E_j}$, e^α can be obtained. Replacing this back in (4.28) and exponentiating gives

$$\bar{N}_k^* = N^* \frac{e^{-\beta E_k}}{\sum_k e^{-\beta E_j}}. \tag{4.29}$$

The result in (4.29) gives what the average distribution of systems over system states will be in a supersystem at equilibrium. The instantaneous distribution will fluctuate around this distribution. The relative fluctuations of the occupation numbers for large enough N^* are negligible, so to this accuracy, $\bar{N}_k^*/N^*$ can be equated to P_k. Setting $Z = \sum_j e^{-\beta E_j}$, the instantaneous probability that an arbitrarily chosen system of this supersystem will be in system state k can be summarized as follows

$$P_k = Z^{-1} e^{-\beta E_k}. \tag{4.30}$$

This distribution is the quantum version of the canonical distribution in phase space, and is referred to as the quantum mechanical canonical ensemble. The function Z so defined is called the partition function.

In effect, this formalism has permitted the construction of a type of measuring device. Let us show that the microscopic ideas which have led to these results immediately imply consequences at the macroscopic level. To this end, it will be established what the exact form of the connection between Z and the Helmholtz free energy F actually is. The starting point is the second part of (4.20). Putting $U = E^*/N^*$, it implies

$$U = \sum_j P_j E_j. \tag{4.31}$$

Formula (4.31) is in agreement with the postulate maintaining that the energy U of the physical system must be identified with the ensemble average $\langle E \rangle$ of the energy.
Begin by considering the change dU of the energy U when the experimental system remains closed but undergoes an infinitesimal reversible process. Equation (4.31) implies that

$$dU = \sum_j (E_j\, dP_j + P_j\, dE_j). \tag{4.32}$$

Now (4.30) can be solved for E_j in the form $E_j = -\beta^{-1}(\log Z + \log P_j)$. Consequently, since $\sum_j P_j = 1$, it is found that $\sum_j dP_j = 0$. Combining these it then follows that

$$-\sum_j E_j dP_j = \beta^{-1} \sum_j (\log Z + \log P_j)\, dP_j = \beta^{-1} \sum_j \log P_j\, dP_j = \beta^{-1} d(\sum_j P_j \log P_j). \tag{4.33}$$

Further, with $-\,\mathrm{d}W$ the work done on the system during the given process, we have that

$$\sum_j P_j\, dE_j = -\,\mathrm{d}W, \tag{4.34}$$

Combining (4.33) and (4.34), we get the result

$$dU = -\beta^{-1} d(\sum_j P_j \log P_j) - \,\mathrm{d}W. \tag{4.35}$$

Comparing (4.35) with the first law $dU = \mathrm{d}Q - \mathrm{d}W$, it is asserted that

$$\beta\,\mathrm{d}Q = -d(\sum_j P_j \log P_j). \tag{4.36}$$

Since the right-hand side of (4.36) is an exact differential, it is concluded that β is an integrating factor for $\mathrm{d}Q$. By the second law of thermodynamics, β must be proportional to T^{-1} and the proportionality constant must be the reciprocal of k_B. With β of this form, when combined with the second law $\mathrm{d}Q = TdS$, we have

$$dS = -k_B d(\sum_j P_j \log P_j). \tag{4.37}$$

This can be easily integrated to give

$$S = -k_B \sum_j P_j \log P_j + C, \tag{4.38}$$

where the integrating constant C is independent of both T and V. In fact the additive property of entropy requires that $C = 0$.

This complicated procedure has returned us in some sense to where we began with (1.5), but by a different route. To get a relation between Z and F, use (4.30), (4.31) and (4.38) to write

$$TS = -k_B T \sum_j P_j \log P_j = k_B T \log Z + \sum_j P_j E_j = k_B T \log Z + U. \qquad (4.39)$$

Consequently, since $F = U - TS$, (4.39) implies the following result

$$F = -k_B T \log Z. \qquad (4.40)$$

Through the construction of these ensembles at a fundamental quantum level, a formalism has been obtained which will allow us to obtain concrete predictions for many equilibrium thermodynamic properties of a system once the function $Z = Z(T, V, N)$ is known. In fact, it follows from the thermodynamic equation

$$dF = -S\, dT - p\, dV + \mu\, dN, \qquad (4.41)$$

where μ is the chemical potential per molecule, that

$$S = -(\frac{\partial F}{\partial T})_{V,N} = k_B \log Z + k_B T(\frac{\partial Z}{\partial T})_{V,N}, \qquad (4.42)$$

$$p = -(\frac{\partial F}{\partial V})_{N,T} = k_B T(\frac{\partial Z}{\partial V})_{N,T},$$

$$\mu = (\frac{\partial F}{\partial N})_{T,V} = -k_B T(\frac{\partial \log Z}{\partial N})_{T,V},$$

$$U = F + TS = k_B T^2(\frac{\partial \log Z}{\partial T})_{V,N}. \qquad (4.43)$$

As an application of these results, consider the one-dimensional harmonic oscillator which has quantum mechanical energy eigenvalues given by

$$\epsilon_n = (n + \frac{1}{2})\hbar\omega, \qquad n = 0, 1, 2, \cdots. \qquad (4.44)$$

The single-oscillator partition function is given by

$$z(\beta) = \sum_{n=0}^{\infty} e^{-\beta(n+\frac{1}{2})\hbar\omega} = (2\sinh(\frac{1}{2}\beta\hbar\omega))^{-1}.$$

The N-oscillator partition function is then given by

$$Z_N(\beta) = [z(\beta)]^N = (2\sinh(\frac{1}{2}\beta\hbar\omega))^{-N}. \qquad (4.45)$$

The Helmholtz free energy follows from (4.40),

$$F = N k_B T \log(2\sinh(\frac{1}{2}\beta\hbar\omega)).$$

By means of F, (4.42) and (4.43) imply that $\mu = F/N$, $p = 0$ and the entropy and energy are

$$S = Nk_B\left[\frac{\beta\hbar\omega}{e^{\beta\hbar\omega}-1} - \log(1 - e^{-\beta\hbar\omega})\right], \quad U = N\left[\frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\beta\hbar\omega}-1}\right]. \tag{4.46}$$

5. Conclusions

It has been seen that formulating the concept of entropy at the microscopic level can be closely related to studying the foundations of quantum mechanics. Doing so provides a useful formalism for exploring many complicated phenomena such as entanglement at this level. Moreover, predictions can be established which bridge a gap between the microscopic and the macroscopic realm. There are many other topics which branch out of this introduction to the subject. For example there is a great deal of interest now in the study of the quantization of nonintegrable systems Gutzwiller (1990), which has led to the field of quantum chaos. There are many indications of links in this work between the areas of nonintegrability and the kind of ergodicity assumed in statistical mechanics which should be pursued.

6. References

Ballentine L. (1970). The Statistical Interpretation of Quantum Mechanics, *Reviews of Modern Physics*, 42, 358-381.

Grössing G. (2008). The vacuum fluctuation theorm: Exact Schrödinger equation via nonequilibrium thermodynamics, *Physics Letters*, B 372, 4556-4563.

Gutzwiller M. (2009). Chaos in Classical and Quantum Mechanics, Springer-Verlag, NY.

Haroche S. and Raimond J. (2006). *Exploring the Quantum*, Oxford University Press.

Landau L. and Lifshitz E. (1978). *Statistical Physics*, Pergamon Press.

Lieb E. and Yngvason J. (1999). The physics and mathematics of the second law of thermodynamics, *Physics Reports*, 310, 1-96.

Mintet F., Carvalho A., Kees M. and Buchleitner A. (2005). Measures and dynamics of entangled states, *Physics Reports*, 415, 207-259.

von Neumann J. (1955). *Mathematical Foundations of Quantum Mechanics*, translated by R. T. Beyer, Princeton University Press, Princeton, NJ.

Padmanabhan T. (1990). Statistical Mechanics of self gravitating systems, *Physics Reports*, 188, 285-362.

Pathria R. (1977). *Statistical Mechanics*, Pergamon Press.

Pauli W. (2000). *Statistical Mechanics, Pauli Lectures on Physics*, vol. 4, Dover, Mineola, NY.

Peres A. (1995). *Quantum Theory: Concepts and Methods*, Klewer Academic Publishers, Dordrecht, The Netherlands.

Peres A. and Terno D. (2004). Quantum Information and Relativity Theory, *Reviews of Modern Physics*, 76, 93-123.

Raimond J., Brune M. and Haroche S. (2001). Manipulating quantum entanglement with atoms and photons in a cavity, *Reviews of Modern Physics*, 73, 565-582.

Rajeev S. (2008). Quantization of contact manifolds and thermodynamics, *Annals of Physics*, 323, 768-782.

Wehrl A. (1978). General Properties of Entropy, *Reviews of Modern Physics*, 50, 221-260.

Weinberg S. (1989). Testing Quantum Mechanics, *Annals of Physics*, 194, 336-386.

Yourgrau W., van der Merwe A. and Raw G., (1982). *Treatise on Irreversible and Statistical Thermophysics*, Dover, Mineola, NY.
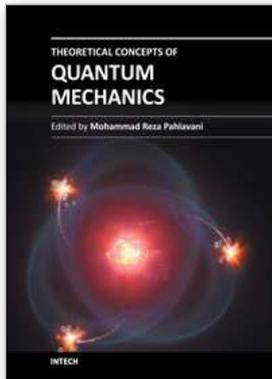
Theoretical Concepts of Quantum Mechanics

Edited by Prof. Mohammad Reza Pahlavani

Quantum theory as a scientific revolution profoundly influenced human thought about the universe and governed forces of nature. Perhaps the historical development of quantum mechanics mimics the history of human scientific struggles from their beginning. This book, which brought together an international community of invited authors, represents a rich account of foundation, scientific history of quantum mechanics, relativistic quantum mechanics and field theory, and different methods to solve the Schrodinger equation. We wish for this collected volume to become an important reference for students and researchers.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Paul Bracken (2012). Quantum Mechanics Entropy and a Quantum Version of the H-Theorem, Theoretical Concepts of Quantum Mechanics, Prof. Mohammad Reza Pahlavani (Ed.), ISBN: 978-953-51-0088-1, InTech, Available from: http://www.intechopen.com/books/theoretical-concepts-of-quantum-mechanics/quantum-mechanical-ensembles-and-the-h-theorem

INTECH
open science | open minds