Thermodynamics and Statistical Mechanics Review

Qi Feng

University of Science and Technology of China

E-mail: feqi@mail.ustc.edu.cn

Contents

1 Thermodynamics

Thermodynamics is notorious for its multiple but equivalent formalisms and its orgy of partial derivatives. What you really need to do is building your system, which is to ensure that others can understand it. The following system belongs to me, which means you may couldn't get used to it. It's fine.

1.1 Energy and Entropy in Thermodynamics

Let's begin with a simple example. Imagine a cylinder of gas,with a piston on top. The piston exerts some pressure P and encloses a volume V of gas. We could build a map between the gas in equivalent and a point in the (P, V) plane. The gas has an internal energy U, which is a state variable. It has a unique value associated with every state. There are two ways to change U.One is to move the piston and do some mechanical work,in which case

$$dU = -PdV$$

by the law of conservation of energy.The other is to put the gas on a hot or cold plate. In this case some heat δQ can be added and we write

$$dU = \delta Q$$

which acknowledges the fact that heat is a form of energy as well.Now we have the first law of thermodynamics:

$$dU = \delta Q - PdV.$$

I use δQ and not dQ since Q is not a state variable. In calculus, we will say δQ is not a exact differential form, but dU is. It is obvious that there is no unique Q associated with a point (P, V): We can go on a closed loop in the (P, V)plane, come back to the same state, but Q would have changed by the negative of the work down, which is the area inside the loop.

The second law of the thermodynamics introduces another state variable, S, the entropy, which changes by

$$dS = \frac{\delta Q}{T},$$

when heat δQ is added reversibly. It's a state variable because it can be shown that $\oint dS = 0$ for a quasi-static cyclic process.

Now we can rewrite dU as

$$dU = TdS - PdV,$$

which tells us that

$$U = U(S, V)$$

$$T = \left(\frac{\partial U}{\partial S}\right)_V$$

$$P = -\left(\frac{\partial U}{\partial V}\right)_S$$

We could also use entropy as characteristic function to obtain other thermodynamics function. We rewrite dS as

$$dS = \frac{1}{T}dU + \frac{P}{V}dV$$

then we have

$$S = S(U, V)$$

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V$$

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_U$$

The function $U(S,V)$, also called the fundamental relation,constitutes complete thermodynamics knowledge of the system.As a example, let's consider n moles of an ideal gas for which it's known from experiments that

$$U(S,V) = C\left(\frac{e^{S/nR}}{V}\right)^{\frac{2}{3}}.$$

We apply the definition of P and T, and we obtain that

$$P = -\left(\frac{\partial U}{\partial V}\right)_S = \frac{2}{3}\frac{U}{V}$$

$$T = \left(\frac{\partial U}{\partial S}\right)_V = \frac{2U}{3nR}$$

$$PV = nRT$$

It's our familiar form of the equation of state.

1.2   Equilibrium as Maximum of Entropy

The second law of thermodynamics states that S is a maximum at equilibrium.But S, like U, is a state variable defined only in equilibrium. Though we could define S in the nonequilibrium state, this statement's meaning is indicating the evolution direction to equilibrium. Imagine a box of gas in equilibrium. It has a volume V and energy U. Now we divides the volume into two parts of size $V_1 = \alpha V$ and $V_2 = (1 - \alpha)V$. The entropy of the combined system is just the sum:

$$S = S_1(V_1) + S_2(V_2) = S_1(\alpha V) + S_2((1 - \alpha)V)$$

Suppose we now let the interface move. It's restricted by the second law.

$$0 = dS = dS_1 + dS_2 = \frac{\partial S_1}{\partial V_1}dV_1 + \frac{\partial S_2}{\partial V_2}dV_2 = \left(\frac{P_1}{T_1} - \frac{P_2}{T_2}\right)d\alpha V$$

We know $T_1 = T_2$ in equilibrium and we obtain $P_1 = P_2$. It's the correct physical meaning:in equilibrium, when S is maximized, the pressures will be equal. We will use the principle of maximum entropy as a constraint to solve the problem at the following section.

1.3   Free Energy in Thermodynamics

Now we introduce a new characteristic function F, called the free energy. It could be construct by legendre transformation:$F = U - TS$. Here are some properties:

$$dF = dU - TdS - SdT = -SdT - PdV$$

$$F = F(T,V)$$

$$-S = \left(\frac{\partial F}{\partial T}\right)_V$$

$$-P = \left(\frac{\partial F}{\partial V}\right)_T$$

Here is an example.You could have a try. If we know

$$T(S,V) = \frac{\partial U}{\partial S} = \frac{2U}{3nR}$$

we can obtain

$$F(T, V) = \frac{3nRT}{2} \left[(1 + \ln C) - \ln \left(\frac{3nRT}{2} - \frac{2}{3} \ln V \right) \right]$$

It is just the free energy of an ideal gas.

Now we may ask why we need a new function when we could describe the system with U ans S.

1.4   Equilibrium as Minimum of Free Energy

We prefer $F(T, V)$ to $U(S, V)$, because it's easier to control T than S. Start with the principle defining the equilibrium of an isolated system as the maximum of S at fixed U(when a constraint is removed). It can equally well be stated as the minimum of U at fixed S.After the Legendre transformation from U to F, there is an equivalent principle that determines equilibrium. You could image there is a constraint in the system.If you remove it, the system will evolute to the minimum F. It is similar to entropy. The principle point out the direction for the micro material, and you would find T and P are same in everywhere as the system is equalized.The proof of $P_1 = P_2$ in the same box of gas is left for you to complete.

You can see this energy from another perspective. We talked about that S is a maximum at equilibrium before. We also have the belief that the energy is a minimum at equilibrium. So there are two tendency–one wants the maximized disorder and another wants the minimal energy. We could unify these two tendency by using the free energy. Now we only consider about the free energy and we want it to be minimum. The occurrence of a phase transition is just the result of the competitive instances: the first tends to minimize the energy while the second tends to maximize the entropy. It is a good opportunity to use the free energy to describe the system.

Now we complete the review on thermodynamics.But I don't include all the things about thermodynamics. If you want to know more details or other things, like phase transition and Maxwell relationship, you could refer this. Then we will turn to statistical mechanics.

2   Classical Statistical Mechanics

2.1   Microcanonical Ensemble

Statistical mechanics is the field of physics mainly interested in the thermodynamic properties of systems made of an enormous number of particles, typically of the order of the Avogadro number $N_A \sim 10^{23}$. It is convenient to introduce the phase space Γ of the system. Let's assume that the system is made of N particles, each of them identified by a set of d coordinates q_i and d momenta p_i. The phase space Γ is the vector space of $2d \times N$ dimensions, given by the tensor product of the coordinates and momenta of all the particles. In the phase space, the system is identified at any given time by a point and its motion is associated to a curve in this space. If the system is isolated, its total energy E is conserved: in this case the motion takes place along a curve of the surface of Γ defined by the equation $H(q_i, p_i) = E$, where $H(q_i, p_i)$ is the hamiltonian of the system.

For a system with a large number of particles not only is it impossible to follow its motion but it is also useless. The only thing that matters is the possibility to predict the average properties of the system that are determined by the macroscopic constraints to which the system is subjected, such as its volume V, the total number N of particles, and its total energy E. Since there is generally a huge number of microscopic states compatible with a given set of macroscopic constraints, it is natural to assume that the system will visit all of them during its temporal evolution. Instead of considering the time evolution of the system, it is more convenient to consider an infinite number of copies of the same system, with the same macroscopic constraints. This leads to the idea of statistical ensembles. By using an analogy, this is equivalent to looking at an infinite number of snapshots of a single movie rather than the movie itself. The ensembles then provide a statistical sampling of the system.

Since each system is represented by a single point in phase space, the set of systems associated to the ensemble corresponds to a swarm of points in phase space. Because the Liouville theorem states that the density of the points at any given point remains constant during the time evolution, a probability density $\tilde{\rho}_i(q,p)$ is naturally defined in Γ. Hence, we can determine expectation values of physical quantities in terms of expectation values on the ensemble (a procedure that is relatively easy) rather than as a time average of an individual system (a procedure that is instead rather complicated). If the system is ergodic we have in fact the fundamental identity

$$\langle A \rangle = \lim_{t \to \infty} \frac{1}{t} \int_0^t d\tau A[q(\tau), p(\tau)] = \int dq dp A(q,p) \tilde{\rho}(q,p)$$

The microcanonical ensemble is defined by the following macroscopic conditions: a fixed number N of particles, a given volume V, and a given value of the energy in the range E and $E + \Delta$. In this ensemble the mean values are computed in terms of the probability density $\rho(q,p)$ defined by

$$\rho(q,p) = \begin{cases} 1 & \text{if } E < H(q,p) < E + \Delta s, \\ 0 & \text{otherwise} \end{cases}$$

i.e. for any physical quantity A we have

$$\langle A \rangle = \frac{\int dq dp A(q,p) \rho(q,p)}{\int dq dp \rho(q,p)}$$

The fundamental physical quantity in this formulation is the entropy. Once this quantities is known, one can recover all the rest of the thermodynamics. The entropy is a function of E and V, defined by

$$S(E,V) = k \log \Omega(E,V),$$

where k is the Boltzmann constant and ω is the volume in the phase space Γ of the microcanonical ensemble

$$\Omega(E,V) = \int dq dp \rho(q,p)$$

he absolute temperature is then given by

$$\frac{1}{T} = \frac{\partial S(E,V)}{\partial E},$$

while the pressure P is defined by

$$P = T \frac{\partial S(E,V)}{\partial E}.$$

2.2   Canonical Ensemble

The canonical ensemble permits us to deal with the statistical properties of a system that is in contact with a thermal bath much larger than the system itself. In this ensemble, the assigned macroscopic conditions are given by the total number N of the particles, the volume V of the system, and its temperature T. In this ensemble we cannot fix the energy, for it can be freely exchanged between the system and the thermal bath. Actually, it is more common to confirm the system's temperature than to ensure the isolation of a system and the constant value of its energy. The probability density of the canonical ensemble takes the form of the Gibbs distribution

$$\rho(q,p) = e^{-\beta H(q,p)},$$

with $\beta = 1/kT$. The partition function is given by

$$Z_N(V,T) = \int dq\, dp\, e^{-\beta H(q,p)}.$$

The mean values are computed according to the formula

$$\langle A \rangle = \frac{1}{Z_N} \int dq\, dp\, A(q,p) e^{-\beta H(q,p)}$$

Now we try to prove that the canonical ensembles are equivalent to the microcanonical. We analyze the fluctuations of the energy

$$\Delta E^2 = \langle H^2 \rangle - \langle H \rangle^2.$$

If you expand the $\langle H^2 \rangle$, you would obtain

$$\langle H^2 \rangle - \langle H \rangle^2 = -\frac{\partial \langle H \rangle}{\partial \beta} = kT^2 \frac{\partial \langle H \rangle}{\partial T} = kT^2 C_V,$$

where C_V is the specific heat. Since in a macroscopic system $\langle H \rangle \propto N$ but also $C_V \propto N$, the fluctuations of the energy are of gaussian type, namely in the limit $N \to \infty$ we have

$$\lim_{N \to \infty} \frac{\Delta E^2}{\langle H \rangle^2} = 0.$$

In other words, even though in the canonical ensemble the energy is a quantity that is not fixed but is subjected to fluctuations, as a matter of fact it assumes the same value in the utmost majority of the systems of the ensemble. This proves the equivalence between the two ensembles.

2.3   Grand Canonical Ensemble

In the grand canonical ensemble, one posits that the system can have an arbitrary number of particles; with its mean value determined by its macroscopic conditions. By introducing the quantity $z = e^{\beta \mu}$, where μ is the fugacity, the probability density of the grand canonical ensemble is given by

$$\rho(q,p,N) = \frac{1}{N!} z^N e^{-\beta H(q,p)}.$$

The term $N!$ in this formula takes into account the identity of the configurations obtained by the permutation of N identical particles. By integrating over the coordinates and the momenta, we arrive at the probability density relative to N particles. In its normalized form, it is expressed by

$$\rho(N) = \frac{1}{Z} \frac{z^N}{N!} Z_N(V,T),$$

where $Z_N(V,T)$ is the partition function of the canonical ensemble with N particles, whereas the denominator of this formula defines the grand canonical partition function

$$Z(z,V,T) = \sum_{N=0}^{\infty} \frac{z^N}{N!} Z_N(V,T).$$

The mean value of the number of particles of the system can be computed by the formula

$$\langle N \rangle = \sum_{N=0}^{\infty} N \rho(N) = z \frac{\partial}{\partial z} \ln Z(z,V,T).$$

The pressure P is linked with the partition function Z by

$$P = \frac{1}{\beta V} \ln Z(z, V, T).$$

The equivalence of this ensemble to the previous ones can be proved by showing that the fluctuations of the number of the number of particles are purely gaussian. It is easy to prove that

$$\lim_{V \to \infty} \frac{\langle N^2 \rangle - \langle N \rangle^2}{\langle N \rangle^2} = 0.$$

This equation shows that, even though the number of particles of the system is not fixed, it has the same value in almost all copies of the ensemble.

Now we can say that for macroscopic system, the three different ensembles give the same final results. The choice of one or another of them is then just a question of what is the most convenient for the problem at hand.

3 Quantum Statistical Mechanics

3.1 Properties of Quantum Systems

In quantum mechanics any observable A is associated with a hermitian operator that acts on a Hilbert space. At each time t, the state of an isolated system is identified by a vector $|\Psi(t)\rangle$ that evolves according to the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = H |\Psi(t)\rangle,$$

where H is the hamiltonian. By using the linear superposition principle, each state of the system can be expressed in terms of a complete set of states $|\psi(t)\rangle$ provided by the orthonormal eigenvectors of any observable A

$$A |\psi(n)\rangle = a_n |\psi(t)\rangle, \qquad \langle \psi_n | \psi_n \rangle = \delta_{n,m}.$$

This means that $|\Psi\rangle$ is given by

$$|\Psi\rangle = \sum_n c_n |\psi_n\rangle.$$

For the completeness relation of these states

$$\sum_n |\psi_n\rangle \langle \psi_n| = 1.$$

The coefficients c_n are expressed by the scalar product $c_n = \langle \Psi | \psi_n \rangle$, and the square of their modulus $|c_n|^2$ expressed the probability to obtain the eigenvalues a_n as a result of the measurement of the observable A on the state $|\Psi\rangle$. Hence,

$$\langle \Psi | \Psi \rangle = \sum_n |c_n|^2 = 1.$$

Now we focus on the statistical properties of quantum systems. As in the quantum case, in the presence of a large number of particles it is highly unrealistic to determine the behavior of a system by solving the Schrödinger equation. Hence, one needs to use a statistical formulation. We just take the incomplete information into consideration and predict the mean values of the observable. To do so, let's imagine that the system under study can be considered as a subsystem of a large one (external world) and in thermodynamic equilibrium. Denote by $\mathcal{H}$ the hamiltonian of such subsystem, E_n the spectrum of its eigenvalues and $|\phi_n\rangle$

its eigenvectors. We can use $|\phi_n\rangle$ to express the states of the system that

$$|\Psi(t)\rangle = \sum_n c_n(t)|\phi_n\rangle$$

So the mean value of an observable $\mathcal{O}$ is given by

$$\langle\Psi(t)|\mathcal{O}|\Psi(t)\rangle = \sum_{n,m} c_n^*(t)c_m(t)\langle\phi_n|\mathcal{O}|\phi_m\rangle = \sum_{n,m} c_n^*(t)c_m(t)\mathcal{O}_{n,m},$$

where $\mathcal{O}_{n,m} = \langle\phi_n|\mathcal{O}|\phi_m\rangle$. Since we have only partial information on the system, we have to take a statistical average. Under the hypothesis of ergodicity, this is equivalent to taking the time average of $\langle\mathcal{O}\rangle$. Defining

$$\rho_{m,n} = \overline{c_m(t)c_n^*(t)} \equiv \lim_{t\to\infty}\frac{1}{t}\int_0^t c_m(\tau)c_n^*(\tau)d\tau,$$

the statistical average of the observable $\mathcal{O}$ can be expressed by the formula

$$\overline{\langle\mathcal{O}\rangle} = \overline{\langle\Psi|\mathcal{O}|\Psi\rangle} = \sum_{n,m}\rho_{m,n}\mathcal{O}_{n,m} = \mathrm{Tr}(\rho\mathcal{O}),$$

where the operator ρ, defined before, is the density matrix. Since the trace of an operator is independent of the basis, the final result does not depend on the basis of the eigenvectors that we used to expand the state $|\Psi\rangle$.

In quantum statistical mechanics, the density matrix corresponds to the probability distribution of classical statistical mechanics. Hence, also in this case, we can introduce three different ensembles.

3.2 Ensembles

3.2.1 Microcanonical Ensemble

As in the classical case, the microcanonical ensemble is defined by the following macroscopic conditions: a fixed number N of particles, a fixed volume V, and the energy of the system in the range E and $E + \Delta$. Correspondingly, the density matrix assumes the form

$$\rho_{n,m} = \delta_{n,m}w_n, \qquad w_n = \begin{cases} 1; & E < E_n < E + \Delta, \\ 0; & \text{otherwise} \end{cases}$$

and the thermodynamics is derived starting from the entropy

$$S(E,V) = k\ln\Omega(E,V),$$

where

$$\Omega(E,v) = \mathrm{Tr}\,\rho.$$

3.2.2 Canonical Ensemble

In this ensemble the macroscopic variables are given by the fixed number N of particles, the volume V, and the temperature T. The corresponding expression the density matrix is given by

$$\rho_{n,m} = \delta_{n,m}e^{-\beta E_n},$$

with the partition function expressed by

$$Z_N(V,T) = \mathrm{Tr}\,\rho = \sum_n e^{-\beta E_n}.$$

In this ensemble, the thermodynamic is derived starting from the free energy

$$F_N(V,T) = -\beta^{-1} \ln Z_N(V,T).$$

3.2.3 Grand Canonical Ensemble

In the grand canonical ensemble the macroscopic variables are the volume V and temperature T. In this case the density matrix acts on a Hilbert space with an indefinite number of particles. Denoting by $E_{n,N}$ the n-th energy level with N particles, the density matrix is expressed by

$$\rho_{n,N} = z^N e^{-\beta E_{n,N}},$$

where $z = e^{\beta\mu}$. The equation of state is similar to the classical one

$$P = \frac{1}{\beta V} \ln Z(z,V,T),$$

where $Z(z,V,T)$ is the grand canonical partition function

$$Z(z,V,T) = \sum_{N,n} z^N e^{-\beta E_{n,N}}.$$

3.3 Indistinguishable Particles and Statistics

A central idea of quantum theory is the concept of indistinguishable particles: for a system with many identical particles, an operator that exchanges two of them, swapping their positions, leave the physics invariant. We can define the transposition operator P_{ij}:

$$P_{ij} \left| \cdots \varphi^i_{\alpha_i} \cdots \varphi^j_{\alpha_j} \cdots \right\rangle = \left| \cdots \varphi^j_{\alpha_i} \cdots \varphi^i_{\alpha_j} \cdots \right\rangle$$

This means that:

$$P_{ij}^2 = 1$$

According to the principle of the indistinguishability of identical particles, the N-particles state can be changed only in terms of a non-essential phase factor through the action of P_{ij}, which means that:

$$P_{ij} \left| \cdots \varphi^i_{\alpha_i} \cdots \varphi^j_{\alpha_j} \cdots \right\rangle = \left| \cdots \varphi^j_{\alpha_i} \cdots \varphi^i_{\alpha_j} \cdots \right\rangle = \lambda \left| \cdots \varphi^i_{\alpha_i} \cdots \varphi^j_{\alpha_j} \cdots \right\rangle$$

i.e. $\lambda = \pm 1$. It's obvious that the state of a system of identical particles are either symmetric or anti-symmetric under exchange of a pair of particles. The first is relative to Bosons and the second to Fermions. The different particles belong to different statistics. Now we talk about a system which can be described by the states of a single particle, here denoted by the index ν. Since the particles are indistinguishable at the quantum level, to specify a state of the system it is sufficient to state the occupation number n_ν of its modes. If ϵ_ν is the energy of the $\nu - th$ mode, the total energy of the system is given by

$$E = \sum_\nu n_\nu \epsilon_\nu,$$

while the total number of particles is

$$N = \sum_\nu n_\nu.$$

In the Fermi-Dirac statistics, each mode can be occupied by at most one particle (the anti-symmetry requires Pauli exclusion principle), so that the possible values of n_ν are

$$n_\nu = 0, 1 \qquad \text{Ferimi-Dirac,}$$

while, in the Bose-Einstein statistics, each mode can be occupied by an arbitrary number of particles. In this case the possible values of n_ν coincide with the natural numbers

$$n_\nu = 0, 1, 2, \ldots \qquad \text{Bose-Einstein.}$$

The most convenient ensemble to describe the thermodynamics of this system is the grand canonical one. The corresponding partition function is

$$Z(z, V, T) = \sum_{N=0}^{\infty} \sum_{\substack{\{n_\nu\} \\ \sum n_\nu = N}} z^N e^{-\beta \sum n_\nu \epsilon_\nu} = \sum_{N=0}^{\infty} \sum_{\substack{\{n_\nu\} \\ \sum n_\nu = N}} \prod_\nu (z e^{-\beta \epsilon_\nu})^\nu.$$

To perform the double sums, it is sufficient to sum independently on each index n_ν, for every term in one case appears once and only once, and vice versa. Hence

$$Z(z, V, T) = \sum_{n_0} \sum_{n_1} \cdots \left[(z e^{-\beta \epsilon_0})^{n_0} (z e^{-\beta \epsilon_1})^{n_1} \cdots \right]$$

$$= \left[\sum_{n_0} (z e^{-\beta \epsilon_0})^{n_0} \right] \left[\sum_{n_1} (z e^{-\beta \epsilon_1})^{n_1} \right] \cdots$$

$$= \prod_\nu \left[\sum_n (z e^{-\beta \epsilon_\nu})^n \right],$$

where the final sum is on the values 0,1 for the fermionic case and on all the integers for the bosonic case. In the first case we have

$$Z_F(z, V, T) = \prod_\nu [1 + z e^{-\beta \epsilon_\nu}],$$

while, in the second case, one has a geometrical series

$$Z_B(z, V, T) = \prod_\nu \left[\frac{1}{1 - z e^{-\beta \epsilon_\nu}} \right].$$

The two expressions can be unified by the formula

$$Z(z, V, T) = \prod_\nu (1 \pm z e^{-\beta \epsilon_\nu})^{\pm 1},$$

where the $+$ sign refers to Fermi-Dirac statistics whereas the $-$ sign refers to Bose-Einstein. The equation of state of both cases is

$$\beta P V \ln Z(z, V, T) = \pm \sum_\nu \ln \left(1 \pm z e^{-\beta \epsilon_n u} \right),$$

where the variable z is related to the average number of particles by the equation

$$N = z \frac{\partial}{\partial z} \ln Z(z, V, T) = \sum_\nu \frac{z e^{-\beta \epsilon_\nu}}{1 \pm z e^{-\beta \epsilon_\nu}}.$$

The last expression shows that the occupation average of each mode is given in both cases by

$$\langle n_\nu \rangle = \frac{z e^{-\beta \epsilon_\nu}}{1 \pm z e^{-\beta \epsilon_\nu}}.$$

References

[1] Ramamurti Shankar. *Quantum Field Theory and Condensed Matter*

[2] Giuseppe Mussardo. *Statistical Field Theory*

[3] Stephen J.Blundell, Katherine M. Blundell. *Concepts in Thermal Physics*