

Cryocourse 2016
Espoo, Finland, 26.9.2016

Thermodynamics and Statistical physics

Erkki Thuneberg

Department of physics
University of Oulu

1. Introduction

These are lecture notes for a short summary of the basic principles of thermodynamics and statistical physics. It is assumed that the audience has some prior knowledge of these topics. The lectures are divided into two parts. The first part considers thermodynamics as a phenomenological theory of heat and work. The second gives basic introduction to statistical physics which, among others, can be used to justify thermodynamics.

The time allocated for the lectures was approximately 2 times 90 minutes. This allowed to go through these lecture notes until the section "variable particle number", and the rest was omitted. The present text is slightly edited version of the one presented in the lectures.

Here are some books (not carefully selected) for further reading.

- L.D. Landau and E.M. Lifshitz, *Statistical Physics, Part 2* (Pergamon, Oxford, 1980). In general the series of books by Landau and Lifshitz is very good but I experienced some difficulties understanding the basics from this book.
- K. Huang, *Statistical mechanics* (Wiley, New York 1987). Good old book.
- R.P. Feynman, *Statistical mechanics, a set of lectures* (Westview Press 1972). The beginning is great, but the rest is harder.
- R. Pathria and P. Beale, *Statistical Mechanics, 3rd ed.*. This is used widely as a textbook.
- F. Reif, *Fundamentals of Statistical and Thermal Physics*.
- J. Sethna, *Statistical Mechanics: Entropy, Order Parameters, and Complexity*. A book with modern flavor, available free online (<http://pages.physics.cornell.edu/~sethna/StatMech/>).
- R. Fitzpatrick, *Thermodynamics & Statistical Mechanics: An Intermediate Level Course*. Lecture notes (<http://farside.ph.utexas.edu/teaching/sm1/sm1.html>).
- L. Reichl, *A Modern Course in Statistical Physics*.
- D. Chandler, *Introduction to Modern Statistical Mechanics*.

In these lecture I have used mainly the logic of Huang for thermodynamics and Feynman for the derivation of the Gibbs distribution. For other parts I cannot name a definite source.

2. Thermodynamics

In this chapter we discuss thermodynamics phenomenologically. We state the laws of thermodynamics and develop the theory based on them.

Preliminary concepts

We start by listing some basic thermodynamic concepts.

Thermodynamics deals with macroscopic systems, which typically consist of a huge number of particles. It deals with macroscopic variables. The most standard ones are the pressure P , the volume V , and the temperature T , but there are also others, the number of particles N (macroscopic when counted in moles), the electric E and magnetic B fields, the chemical concentrations c_i , the flow velocity v etc. For simplicity we concentrate here on the variables P , V , and T , but more generally we could have more like X , Y etc. These quantities are well defined when the system under study is in equilibrium. In equilibrium the variables are related by a condition called equation of state. For the variables P , V , and T it has the form

$$f(P, V, T) = 0. \quad (1)$$

This defines a surface in the 3 dimensional space with axes P , V , and T . A thermodynamic transformation is a change of state. The transformation is called reversible if the change proceeds equally in the reverse direction. This is possible for transformations that proceed slowly enough so that the system stays in equilibrium states.

The energy change of the system with its surroundings is divided in two contributions. *Work* arises from macroscopic forces between the system and the surroundings. In system with variables (P, T, V) the work done by the system is

$$dW = PdV \quad (2)$$

when the volume changes by dV . For some other variables the *Heat* is the energy exchanged in the form that is not work (in the sense just defined). It arises from energy exchange by microscopic forces that does not correspond to macroscopic motion in the system. The heat capacity of is defined as the coefficient of proportionality C in the relation of the heat absorbed by the system and the change of temperature,

$$C = \frac{dQ}{dT}. \quad (3)$$

The heat capacity depends on the experimental conditions. The usual ones are C_V for constant volume and C_P for constant pressure.

Other measurable quantities of interest in the (P, T, V) system are thermal expansion

$$\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_P, \quad (4)$$

and the isothermal compressibility

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T, \quad (5)$$

and the adiabatic compressibility

$$\kappa_S = -\frac{1}{V} \left(\frac{dV}{dP} \right)_{dQ=0}. \quad (6)$$

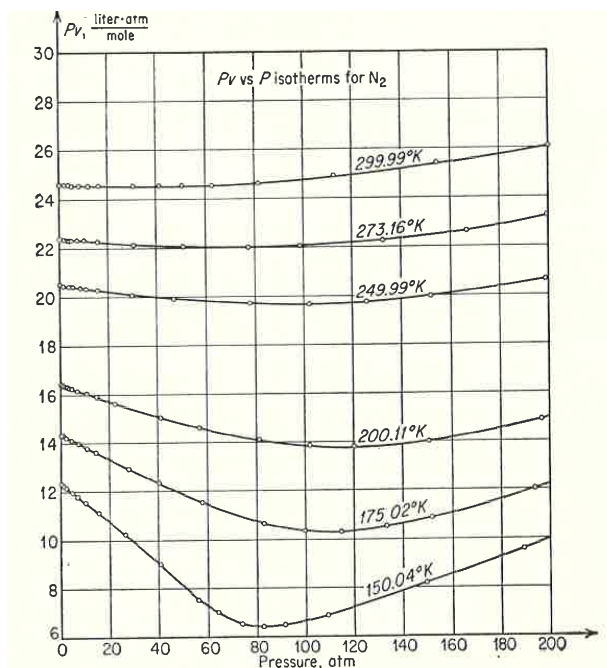
Ideal gas is an idealized thermodynamic model. It obeys the equation of state

$$PV = Nk_B T, \quad (7)$$

where $k_B = 1.380 \times 10^{-23}$ J/K is the Boltzmann constant. In fully macroscopic units one writes

$$PV = nRT, \quad (8)$$

where n is the number of moles and $R = N_A k_B = 8.31$ J/(K mole) the gas constant and N_A the Avogadro constant. Real gases obey well the ideal gas equation of state when they are not too dense. Below an example of PV of Nitrogen (A.S. Friedman 1950).



By fixing the constant in equations (7) and (8) we have defined the absolute scale for the temperature T in degrees of Kelvin. We note that in this scale the zero temperature corresponds that the gas PV/n extrapolates to zero.

Laws of thermodynamics

The basis of thermodynamics can be compressed into a few laws.

The first law. For arbitrary transformation from one state to another, we measure the heat ΔQ absorbed by the system and the work ΔW done by the system. The first law states that the quantity ΔU defined by

$$\Delta U = \Delta Q - \Delta W \quad (9)$$

depends only on the initial and final states (not on the intermediate ones).

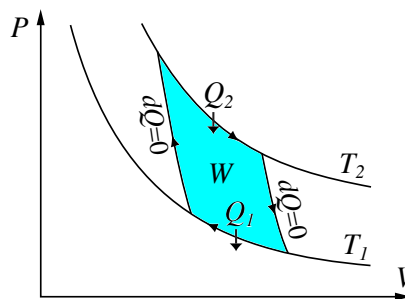
Interpretation: knowing that ΔQ and $-\Delta W$ are just two forms of energy flow to the system, we see that (9) is just the statement of energy conservation, where ΔU is the difference in the *internal energy* U between the initial and final states (and thus does not depend on the intermediate states).

The second law. The second law states that there is no process whose only effect would be a transfer of heat from a system at one temperature to a system at a higher temperature.

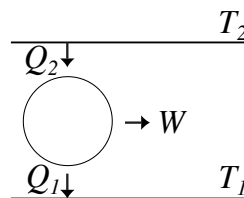
Interpretation: This restricts processes that can take place. The first law would allow a heat Q be taken from a reservoir at temperature T_1 and deliver it to another reservoir at temperature $T_2 > T_1$, but this is prohibited by the second law. An alternative formulation of the second law is that there is no process whose sole effect is to take heat from a reservoir at fixed temperature and convert it all to work. The equivalence of the two statements can be verified by contraposition: assuming one false leads that the other is false and vice versa.

The two laws look quite simple. They are like axioms in a mathematical theory. It is amazing that effectively the whole thermodynamics can be based only on these. In the following we state some of the main consequences of these laws. We skip most of derivations, they can be found in Huang's book, for example.

Consider a cyclic process of a system (an engine) that takes heat Q_2 from a reservoir at temperature T_2 , then shifts adiabatically (=no heat exchange) to a lower temperature T_1 where it gives off heat Q_1 to a second reservoir and then returns adiabatically to the starting point. The process of the engine in $P - V$ plane is illustrated below.



The process can also be represented schematically as below.



Assuming all processes are reversible, this is called the Carnot engine.

By the first law the work done by the engine is $W = Q_2 - Q_1$. By the second law $Q_1 \neq 0$. It can be shown that for $W > 0$, it has to be $Q_2 > 0$ and $Q_1 > 0$. We define the

efficiency (for $W > 0$)

$$\eta = \frac{W}{Q_2} = 1 - \frac{Q_1}{Q_2} \quad (10)$$

The efficiency of any reversible engine working between the same temperatures has to be the same. (If it were not, one could run them in series, the more efficient backwards, which leads to contradiction with the second law.) An engine not working reversibly has to have lower efficiency than the Carnot engine.

The Carnot engine allows to define temperature quantitatively. (Above we used only statements like $T_2 > T_1$.) We set a Carnot engine working between two temperatures and measure the efficiency η_{Carnot} . Then we define that the ratio of the temperatures is

$$\frac{T_1}{T_2} = 1 - \eta_{\text{Carnot}} = \frac{Q_{1,\text{Carnot}}}{Q_{2,\text{Carnot}}} \quad (11)$$

This allows to assign values for all temperatures once one temperature is chosen as reference. The temperatures defined in this way can be shown to agree with the temperatures defined by the ideal gas law (8).

As stated above, the efficiency of any engine is less than that of the Carnot engine, $\eta \leq \eta_{\text{Carnot}}$. Using equations (10) and (11) this can be written

$$\frac{Q_2}{T_2} - \frac{Q_1}{T_1} \leq 0. \quad (12)$$

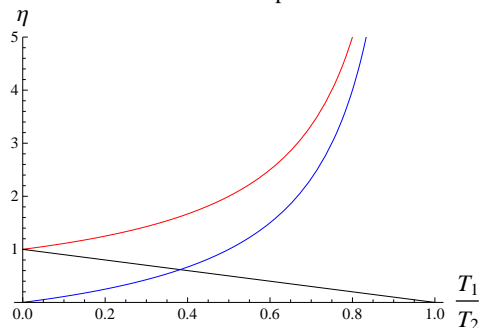
Generalizing this to arbitrary cyclic process we reach the statement

$$\oint \frac{dQ}{T} \leq 0, \quad (13)$$

where dQ is the differential of the heat absorbed by the system. In (13) the equality is achieved for a reversible process.

We derived equations (12) and (13) for an engine with $W > 0$. These equations are valid also for a reversed engine, a heat pump, where $W < 0$. We can use the heat pump as a refrigerator. For a refrigerator it is reasonable to define the coefficient of performance $\eta_r = Q_1/W$. For this we also get that its maximum is given by the Carnot efficiency, $\eta_r \leq \eta_{r,\text{Carnot}} = T_1/(T_2 - T_1)$. Alternatively we can use the heat pump as a heater. For a heater it is reasonable to define the coefficient of performance $\eta_h = Q_2/W$. Again the maximum is given by the Carnot efficiency, $\eta_h \leq \eta_{h,\text{Carnot}} = T_2/(T_2 - T_1)$. Note that the coefficient of performance can be, and often is, larger than unity. This is the advantage of heat pumps in comparison to direct heating, where $\eta_h = 1$.

Carnot coefficients of performance



Equation (13) allows to define *entropy*. The entropy of a state A relative to a reference state O is defined by the integral in a reversible transformation

$$S(A) = S(O) + \int_O^A \frac{dQ}{T} \quad (14)$$

Because of the equality in (13) for reversible transformation, the entropy is independent of the selected path in the integration of (14). The heat differential can now be expressed

$$dQ = TdS, \quad (15)$$

which is somewhat symmetric to $dW = PdV$ (2). Combining these with the first law (9) gives

$$dU = TdS - PdV. \quad (16)$$

In this relation all quantities (U, T, S, P, V) are defined for any given state. This is in contrast to W and Q , which are defined for a particular process and thus are not unique for a given state.

Another useful property of the entropy is the following.

The entropy of a closed system can only increase.

This statement is again equivalent to the second law. Proof: Suppose that the state of a system changes from A to B . We do this back by a reversible transformation from B to A . Applying (13) to the whole process and (14) to the reversible part gives

$$\int_A^B \frac{dQ}{T} + S(A) - S(B) \leq 0. \quad (17)$$

For a closed system $dQ = 0$. The vanishing of the integral in (17) gives $S(B) \geq S(A)$.

As special case we get that the equilibrium state of a closed system corresponds to the maximum of entropy.

In order to remove the unknown constant $S(O)$ in the definition of entropy (14), one can use the **third law of thermodynamics**. It states that the entropy of all substances at absolute zero temperature, $T = 0$, is the same, and can be set to vanish, $S = 0$. In U (and also F , G and H below) an unknown constant remains.

It is useful to define new state functions, the Helmholtz and Gibbs free energies and the enthalpy

$$F = U - TS, \quad (18)$$

$$G = U - TS + PV, \quad (19)$$

$$H = U + PV \quad (20)$$

Let us concentrate on F . It is called *free energy* because the work the system can do *at constant temperature* is always smaller or equal to the decrease of F

$$W \leq -\Delta F = -\Delta U + T\Delta S. \quad (21)$$

The proof follows from (17): at constant temperature the integral equals $\Delta Q/T$ and substituting the first law (9) gives (21). We also reach the conclusion that for a mechanically isolated system at constant temperature F can only decrease, and in equilibrium it reaches its minimum value.

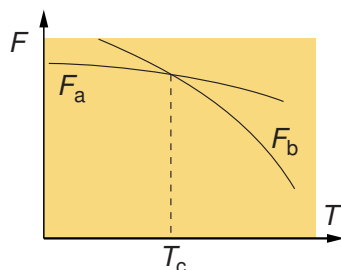
Let us restate the previous result more precisely. In non-equilibrium state $F(T, V, \lambda_1, \lambda_2, \dots)$ depends on several internal degrees of freedom λ_i . In equilibrium it is

$$F(T, V) = \min_{\lambda_1, \lambda_2, \dots} F(T, V, \lambda_1, \lambda_2, \dots). \quad (22)$$

The condition $F = U - TS = \text{minimum}$ generalizes the zero-temperature condition $U = \text{minimum}$ to finite temperature.

Example Phase equilibrium

Often the system can appear in two different phases, for example liquid and gas. Suppose that we have obtained the free energies $F_a(T, V)$ and $F_b(T, V)$ for the two phases as shown in the figure.



According to the previous result, the phase that is realized in equilibrium is the one having lower free energy. We conclude that there is a phase transition between the phases at temperature where

$$F_a = F_b. \quad (23)$$

Similar conclusions as for F apply to G under conditions of constant temperature and pressure.

Relations

Above we have defined quite many quantities. Depending on the problem, any of them can be useful. Here we derive some relations between these quantities.

Using entropy, the specific heats can be written

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_V = \left(\frac{\partial U}{\partial T} \right)_V \quad (24)$$

$$C_P = T \left(\frac{\partial S}{\partial T} \right)_P = \left(\frac{\partial H}{\partial T} \right)_P \quad (25)$$

and the adiabatic compressibility (6)

$$\kappa_S = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_S. \quad (26)$$

The internal energy differential (16) can be interpreted that it is a function of S and V : $U(S, V)$. This allows write the partial derivatives

$$T = \left(\frac{\partial U}{\partial S} \right)_V, \quad P = -\left(\frac{\partial U}{\partial V} \right)_S. \quad (27)$$

where the variable kept constant is given explicitly. Because second partial derivatives are independent of the order of derivation, we have the Maxwell relations

$$\left(\frac{\partial T}{\partial V} \right)_S = -\left(\frac{\partial P}{\partial S} \right)_V. \quad (28)$$

Correspondingly we get for the Helmholtz free energy

$$dF = -SdT - PdV, \quad F(T, V), \quad (29)$$

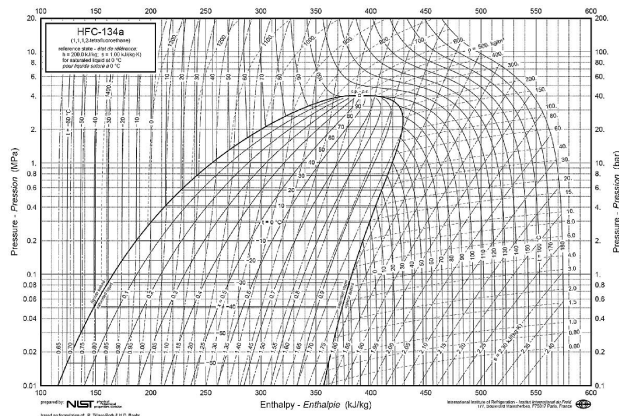
$$S = -\left(\frac{\partial F}{\partial T} \right)_V, \quad P = -\left(\frac{\partial F}{\partial V} \right)_T, \quad (30)$$

$$\left(\frac{\partial S}{\partial V} \right)_T = \left(\frac{\partial P}{\partial T} \right)_V. \quad (31)$$

and for other functions similarly.

Based on these relations it appears that we need to know one of the energies as a function of its proper variables. For example, we could know U as a function of S and V , or F as a function of T and V . From this information we then can extract all other thermodynamic variables using relations between the variables.

How can we then obtain the basic information? This can be based on experiments. Many substances have been measured.



We can also think of calculating the quantities theoretically. For this we need statistical physics.

3. Statistical physics

Statistical method

Statistical physics can be difficult for several reasons. Thinking in terms of probabilities is harder than knowing something for sure. It is not easy to imagine very large numbers. There is quite a number of different approaches: one can use either classical mechanics or quantum mechanics, there are different ensembles (microcanonical, canonical, grand canonical). There is no fixed way of presenting statistical physics. In the texts I have seen, the presentation, the order of the topics, the emphasis, and the examples vary greatly. Some approaches are good for basic understanding while others are good for true calculations. Here I have chosen one (or a few) of the many alternatives. One particular purpose is to give microscopic justification for thermodynamics, which was presented in the first part of these lecture.

Statistical physics is a branch of physics that applies statistical methods to solve physical problems. We do not know the detailed state of the system, but we make an estimate of the possible states of the system, an ensemble. Based on this we can predict the results of a measurement by studying the distribution of measured quantity in the ensemble.

Statistical methods work best for systems that are sufficiently large, for example, contain a large number of particles. Statistical physics does not predict the trajectory of a single particle, but it can very accurately predict collective properties, like the pressure or the temperature. Statistical methods can also be applied to small systems, if the measurement lasts long enough. Supposing that the system hops randomly between its possible states during the measurement time, the measured average value can be predicted by statistical methods.

Statistical methods are particularly used to study systems in equilibrium. In microscopic systems the time constants typically are short and therefore equilibrium is achieved rapidly. With growing system size, the time it takes to achieve equilibrium grows as distant parts of the system tend to be out of equilibrium. In practice, things on a microscopic scale are close to equilibrium but nonequilibrium is around us everywhere on a macroscopic scale. Therefore statistical physics works especially for systems that contain many particles but not too many.

Density matrix

We use statistical description when the state of a system is not precisely known. We use all the knowledge of the system we have in order to exclude part of the generally possible states. In a simple case we have that the remaining possible states can be represented as linear combination of a finite number of orthonormal states $|\psi_i\rangle$, where $i = 1, 2, \dots, W$. If nothing else is known of the likelihood of these states, an often used assumption is that each of these states is equally likely, i.e. occurs with probability $p =$

$1/W$. Then the expectation value $\langle \hat{A} \rangle$ of any operator \hat{A} can be calculated,

$$\langle \hat{A} \rangle = \frac{1}{W} \sum_{i=1}^W \langle \psi_i | \hat{A} | \psi_i \rangle. \quad (32)$$

A convenient technical tool is to introduce the (probability) density operator. In the general case it is defined

$$\hat{\rho} = \sum_i \rho_i |\psi_i\rangle \langle \psi_i|, \quad (33)$$

where ρ_i is the probability of the state $|\psi_i\rangle$. The expectation values are then given by

$$\langle \hat{A} \rangle = \sum_i \rho_i \langle \psi_i | \hat{A} | \psi_i \rangle = \sum_i \langle \psi_i | \hat{A} \hat{\rho} | \psi_i \rangle = \text{Tr}(\hat{A} \hat{\rho}). \quad (34)$$

In these equations the probabilities are assumed to be normalized

$$\sum_i \rho_i = 1 \quad \Leftrightarrow \quad \text{Tr} \hat{\rho} = 1. \quad (35)$$

We emphasize that the states $|\psi_i\rangle$ describe all particles in the system. The corresponding wave functions $\psi_i(\mathbf{r}_1, \mathbf{r}_2, \dots)$ depend on coordinates of all particles and only in special cases can be expressed simply in terms of products of single-particle wave functions $\phi_k(\mathbf{r})$.

As an example consider the internal energy U of the system. It is defined as the expectation value of the energy of the system,

$$U = \langle \hat{H} \rangle = \text{Tr}(\hat{H} \hat{\rho}). \quad (36)$$

Its differential is

$$dU = \text{Tr}(\hat{H} d\hat{\rho}) + \text{Tr}(\hat{\rho} d\hat{H}). \quad (37)$$

Here the second term arises from the change of H . This can be interpreted as work, for example, when the energies of the energy eigenstates change as a result of changing the volume of the system. The first term in (37) comes from changes in the probabilities of each states. This can be interpreted as heat, the energy exchange arising from changing the probabilities of the microscopic states. Thus we arrive at the microscopic definitions

$$dQ = \text{Tr}(\hat{H} d\hat{\rho}), \quad dW = -\text{Tr}(\hat{\rho} d\hat{H}), \quad (38)$$

and (37) then corresponds to the first law of thermodynamics (9). If the essential parameter determining H is the volume V , we get

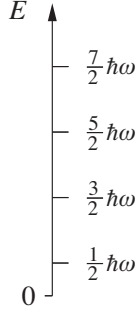
$$dW = pdV, \quad p = -\left\langle \frac{d\hat{H}}{dV} \right\rangle. \quad (39)$$

Number of states

It is important to have insight how many states there are for a given energy interval. We look simple examples. For a single harmonic oscillator the energies are

$$E_n = \hbar\omega(n + \frac{1}{2}), \quad n = 0, 1, 2, \dots \quad (40)$$

In the following we neglect the zero point energy. That is, we count the energy relative to the ground state.



For N oscillators the total energy F is

$$F = \sum_{i=1}^N \hbar\omega_i n_i \quad (41)$$

We are interested in the number of states $n_{\text{states}}(F)$ below a given energy $F \gg \hbar\omega$. For $N = 1$ this is approximately $F/\hbar\omega$. For two harmonic oscillators it is $\frac{1}{2}(F/\hbar\omega_1)(F/\hbar\omega_2) \propto F^2$. For N harmonic oscillators the number of states below energy E seems to be $c(N)F^N$ with some N dependent coefficient $c(N)$. The density of states per energy interval is

$$\eta = \frac{dn_{\text{states}}(F)}{dF} = Nc(N)F^{N-1} \quad (42)$$

For a large N (say $N \sim 10^{23}$) it is not immediately clear how this depend on F . In order to find out we look at the logarithm of η and its derivative with respect to energy

$$\ln \eta = \ln[Nc(N)F^{N-1}] \quad (43)$$

$$= \ln[Nc(N)] + (N-1) \ln F, \quad (44)$$

$$\frac{d \ln \eta}{dF} = \frac{N-1}{F} \approx \frac{1}{E_{\text{ave}}}, \quad (45)$$

where E_{ave} is the average energy per oscillator. Integrating this gives

$$\ln \eta = F/E_{\text{ave}} + c_1 \quad (46)$$

$$\eta = c_2 \exp(F/E_{\text{ave}}). \quad (47)$$

with some constants c_i .

Repeating the same with ideal gas gives

$$\eta = c_3 \exp(3F/2E_{\text{ave}}), \quad (48)$$

where E_{ave} is the average energy per particle.

Gibbs distribution

An important basic result is *Gibbs distribution*: the probability ρ_i that the state ψ_i occurs is

$$\rho_i = e^{\beta(F-E_i)}. \quad (49)$$

Alternatively, the density matrix is

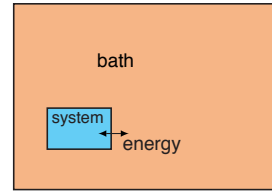
$$\hat{\rho} = e^{\beta(F-\hat{H})}. \quad (50)$$

Here the constant β can be interpreted as inverse temperature: $\beta = 1/(k_B T)$. Here is Boltzmann's constant $k_B = 1.38 \times 10^{-23}$ J/K, which is needed to express the temperature T in Kelvin units. The constant F is determined by the normalization condition (35),

$$F = -\frac{1}{\beta} \ln(\text{Tr} e^{-\beta \hat{H}}) = -\frac{1}{\beta} \ln\left(\sum_i e^{-\beta E_i}\right). \quad (51)$$

We cite what Richard Feynman says about (50). “This fundamental law is the summit of statistical mechanics, and the entire subject is either the slide-down from this summit, as the principle is applied to various cases, or the climb-up where the fundamental law is derived and the concepts of thermal equilibrium and temperature T clarified.” (R.P. Feynman, *Statistical mechanics, a set of lectures*, 1972)

The Gibbs distribution can be derived under the following main assumptions: 1) the system we study interacts with a much larger surroundings called bath. 2) all states of the whole system (system under study + bath) occur with equal probability within some energy interval.



Neglecting the interaction energy between the system and the surroundings, the states of the composite system can be chosen in product form

$$|\Psi_{i,k}\rangle = |\psi_i\rangle |\psi_k^b\rangle \quad (52)$$

where $|\psi_k^b\rangle$ is the state of the bath. The total energy of such a state is $E_{i,k} = E_i + F_k$. As all composite states in the allowed energy range are equally likely, the likelihood of a given system state $|\psi_i\rangle$ is directly proportional to the number γ of bath states that it is allowed to couple with. We assume that for any large bath this is of the form given in (47) and (48). Because of conservation of the total energy we have

$$\rho_i \propto \gamma \propto e^{\beta F} \propto e^{\beta(\text{constant} - E_i)} \propto e^{-\beta E_i}. \quad (53)$$

This gives the claimed exponential distribution (49). The only thing that remains is to relate β to the absolute temperature. This can be done by applying the Gibbs distribution to some known thermometer, for example the ideal gas thermometer.

Connection to thermodynamics

Let us calculate

$$\begin{aligned} \frac{1}{T} dQ &= \frac{1}{T} \text{Tr}(\hat{H} d\hat{\rho}) = \frac{1}{T} \text{Tr}[(\hat{H} - F) d\hat{\rho}] \\ &= -k_B \text{Tr}(\ln \hat{\rho} d\hat{\rho}) = d[-k_B \text{Tr}(\hat{\rho} \ln \hat{\rho})], \end{aligned} \quad (54)$$

where we have used the normalizing condition (35), the microscopic definition of heat (38), and the Gibbs distribution (50) with $\beta = 1/(k_B T)$. Based on comparing this

with the thermodynamic relation (15), we give a microscopic definition of the entropy

$$S = -k_B \langle \ln \hat{\rho} \rangle = -k_B \sum_i \rho_i \ln \rho_i \quad (55)$$

Using this, (35), (36), and (50) gives

$$F = U - ST, \quad (56)$$

which allow to identify F in the Gibbs distribution (50) as the Helmholtz free energy (18).

It is of interest to look at the definition of entropy. Boltzmann invented the definition

$$S = k_B \ln W \quad (57)$$

when the number of microstates is W . Taking the density matrix (33) as equal weight distribution of W states and substituting into (55) gives indeed (57). It can be argued that (55) is the natural generalization of (57) to case when the availability of some states is restricted. This has also close connection to information theory, i.e. to the question of how many bits of information can be coded when the availability of some values of the bits is limited.

In order to derive thermodynamics, we still need to prove the second law. For that consider some non-equilibrium state. For example, consider two independent macroscopic systems with density matrices ρ_1 and ρ_2 . The composite system then has the density matrix $\rho = \rho_1 \rho_2$. The entropy can be calculated from (55) with the result $S = S_1 + S_2$. The argument is now that if the two subsystems are at different temperatures, S is small compared to the case that the subsystems could exchange energy so that they would reach mutual equilibrium (where the temperature would be the same for both). What this means is that the number of microscopic states corresponding to the two subsystems out of equilibrium is much smaller than the number of microscopic states in equilibrium. Thus after joining the subsystems, a nonequilibrium state is highly unlikely compared to the states in equilibrium so that it practically does not appear. It is easy to verify (using Lagrange multipliers) that the Gibbs distribution (50) corresponds to the maximum of entropy (55) under the constraints of constant $U = \langle \hat{H} \rangle$ and (35).

We can also verify the third law. This is straightforward if the ground state (the lowest energy state) is nondegenerate. Even for a degenerate ground state, the entropy at $T = 0$ is small compared to the contributions at $T > 0$, and thus can be neglected in comparison.

Usually there is no upper bound for the energy eigenvalues E_i . In order to the Gibbs distribution (49) to be reasonable, we must have $\beta > 0$, which implies $T \geq 0$. This agrees with our discussion of thermodynamics.

For some systems there is a maximum energy. For example consider a two-state system or their combinations (spin systems). In this cases the Gibbs distribution (49) is reasonable also at negative β , implying negative T according to $\beta = 1/(k_B T)$. Such negative temperature systems

have higher internal energy than the positive temperature states. Thus negative absolute temperatures can be interpreted to be higher temperatures than the positive ones.

Variable particle number

Often it is mathematically easier to study the case, where the particle number is not fixed. This can be achieved by thinking the system as connected to a ‘‘particle bath’’, an ideal reservoir of particles at constant energy μ , which is called the chemical potential. In Gibbs distribution (50) we can generalize

$$\hat{H} \rightarrow \hat{H} + \mu \hat{N}_b = \hat{H} - \mu \hat{N} + \text{constant} \quad (58)$$

(because $N_{\text{tot}} = N + N_b = \text{constant}$). (Here \hat{N} is the particle-number operator of the system: $\hat{N} \Psi_i = N_i \Psi_i$, where N_i is the number of particles in state Ψ_i .) In making the substitution (58) one also replaces the constant F (Helmholtz free energy) with another constant Ω (grand potential). Therefore

$$\hat{\rho} = e^{\beta(\Omega - \hat{H} + \mu \hat{N})}. \quad (59)$$

In the same way as for F , one can derive for Ω the definition (also nonequilibrium, $N = \langle \hat{N} \rangle$)

$$\Omega = U - \mu N - ST, \quad (60)$$

the equilibrium expression

$$\Omega = -\frac{1}{\beta} \ln \left[\text{Tr} e^{-\beta(\hat{H} - \mu \hat{N})} \right] = -\frac{1}{\beta} \ln \left[\sum_i e^{-\beta(E_i - \mu N_i)} \right], \quad (61)$$

the differential for equilibrium states

$$d\Omega = -SdT - pdV - Nd\mu \quad (62)$$

and the time development

$$\frac{d\Omega}{dt} \leq 0 \quad (T, V \text{ and } \mu \text{ constants}). \quad (63)$$

In addition we deduce from equation (62)

$$S = - \left(\frac{\partial \Omega}{\partial T} \right)_{V, \mu}, \quad p = - \left(\frac{\partial \Omega}{\partial V} \right)_{T, \mu}, \quad N = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T, V}. \quad (64)$$

Ideal Fermi gas

Above we formally discussed the many-body wave functions Ψ_i . The calculation of these is possible only in very special cases. One case is an ideal gas, where we assume that there are no interactions between the particles.

The natural choice for wave functions of a single free particle are plane wave states

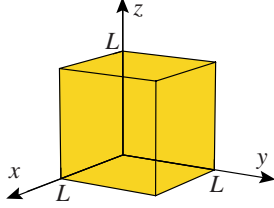
$$\phi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad (65)$$

where the wave vector \mathbf{k} appears as a parameter. The energy of these states is $\epsilon_{\mathbf{k}} = \hbar^2 k^2 / 2m$. In order to count the

states, it is most simple to require that the wave functions are periodic in a cube of volume $V = L^3$, which allows the wave vectors \mathbf{k} (n_x, n_y and n_z integers)

$$k_x = \frac{2\pi n_x}{L}, \quad k_y = \frac{2\pi n_y}{L}, \quad k_z = \frac{2\pi n_z}{L}. \quad (66)$$

We suppose that the volume V is very large. Then we can take the limit $V \rightarrow \infty$ in quantities that do not essentially depend on V .



In addition to the location \mathbf{r} , the wave function of a fermion depends on spin index σ , which describes the component of spin angular momentum on some chosen z axis. For spin- $\frac{1}{2}$ fermions this can have two values. These can be denoted by $\sigma = \pm\frac{1}{2}$, or alternatively by \uparrow and \downarrow . For free particles we can thus choose “spin-up levels”

$$\phi_{\mathbf{k}\uparrow}(\mathbf{r}, \sigma) = \begin{cases} \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} & \text{if } \sigma = \frac{1}{2} \\ 0 & \text{if } \sigma = -\frac{1}{2} \end{cases}. \quad (67)$$

and “spin-down levels”

$$\phi_{\mathbf{k}\downarrow}(\mathbf{r}, \sigma) = \begin{cases} 0 & \text{if } \sigma = \frac{1}{2} \\ \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} & \text{if } \sigma = -\frac{1}{2} \end{cases}. \quad (68)$$

Note that (following book AM) we call the single-particle $\phi_{\mathbf{k}\uparrow}$ and $\phi_{\mathbf{k}\downarrow}$ “levels” in order to clearly distinguish them from the “states” Ψ_i of the many body system.

There are several ways to present many-body states Ψ_i , as will be discussed later. One useful way to think of these states is first to list all levels (67)-(68) in some arbitrary order, for example

$$\phi_{\mathbf{0}\uparrow}, \phi_{\mathbf{0}\downarrow}, \phi_{\mathbf{k}_1\uparrow}, \phi_{\mathbf{k}_1\downarrow}, \phi_{\mathbf{k}_2\uparrow}, \phi_{\mathbf{k}_2\downarrow}, \dots \quad (69)$$

Then the basis states of the many-body space can be expressed by telling how many particles is in any of the levels,

$$|\Psi_i\rangle = |n_1, n_2, n_3, \dots, n_\infty\rangle. \quad (70)$$

Here n_α is the number of particles in the α 'th level (69). [In practice the writing of the state (70) is difficult because there is an infinite number of levels and thus the great majority of the numbers n_α are zeros.] The energy of the many-body state (70) is $E = \sum_\alpha n_\alpha \epsilon_\alpha$. Fermions obey the Pauli exclusion principle and thus all occupations n_α are either 0 or 1.

In order to determine the thermal equilibrium state, it is easiest to use the formulas for variable particle number. Starting from equation (61) we get

$$e^{-\beta\Omega} = \text{Tr} e^{-\beta(\hat{H} - \mu\hat{N})}$$

$$\begin{aligned} &= \sum_{n_1} \sum_{n_2} \dots \langle n_1, n_2, \dots | e^{-\beta(\hat{H} - \mu\hat{N})} | n_1, n_2, \dots \rangle \\ &= \sum_{n_1} \sum_{n_2} \dots e^{-\beta(\epsilon_1 - \mu)n_1} e^{-\beta(\epsilon_2 - \mu)n_2} \dots \\ &= \sum_{n_1} e^{-\beta(\epsilon_1 - \mu)n_1} \sum_{n_2} e^{-\beta(\epsilon_2 - \mu)n_2} \dots \\ &= \prod_\alpha \sum_{n_\alpha} e^{-\beta(\epsilon_\alpha - \mu)n_\alpha} \\ &= \prod_\alpha \left[1 + e^{-\beta(\epsilon_\alpha - \mu)} \right]. \end{aligned} \quad (71)$$

Thus

$$\Omega = -\frac{1}{\beta} \ln \prod_\alpha \left[1 + e^{-\beta(\epsilon_\alpha - \mu)} \right] = -\frac{1}{\beta} \sum_\alpha \ln \left[1 + e^{-\beta(\epsilon_\alpha - \mu)} \right]. \quad (72)$$

From this we can calculate all thermodynamic quantities. Especially the particle number (62) is

$$\begin{aligned} N &= -\left(\frac{\partial \Omega}{\partial \mu} \right)_{T, V} = \frac{1}{\beta} \sum_\alpha \frac{e^{-\beta(\epsilon_\alpha - \mu)} \beta}{1 + e^{-\beta(\epsilon_\alpha - \mu)}} \\ &= \sum_\alpha \frac{1}{e^{\beta(\epsilon_\alpha - \mu)} + 1}. \end{aligned} \quad (73)$$

Here we see that the average occupation probability of each level depends on its energy ϵ and is

$$f(\epsilon) = \frac{1}{e^{\beta(\epsilon - \mu)} + 1}. \quad (74)$$

This is the familiar Fermi-Dirac distribution.

In a similar fashion we can derive for an ideal Bose gas (possible occupation numbers $n_\alpha = 0, 1, 2, \dots, \infty$) the Bose-Einstein distribution

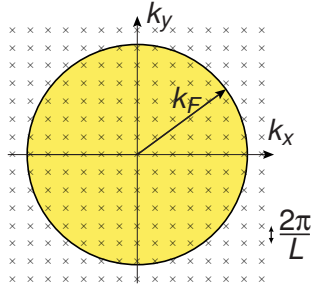
$$f(\epsilon) = \frac{1}{e^{\beta(\epsilon - \mu)} - 1}. \quad (75)$$

The purpose above was to show that ideal Bose and Fermi distributions can be derived from the more general Gibbs distribution, which can be applied to arbitrary interacting systems as well.

Let us remind about the main features of Fermi distribution. When the temperature $T \rightarrow 0$, the occupation becomes a step function

$$f(\epsilon) = \begin{cases} 1 & \text{for } \epsilon < \mu \\ 0 & \text{for } \epsilon > \mu, \end{cases} \quad (76)$$

where all levels below the chemical potential μ are filled. The kinetic energy at highest filled level is called *Fermi energy* ϵ_F , and expressed in temperature units it is called *Fermi temperature* T_F : $\mu(T = 0) = \epsilon_F = k_B T_F$. We also define the *Fermi wave vector* k_F and the *Fermi momentum* $p_F = \hbar k_F$ corresponding to the Fermi energy, $\hbar^2 k_F^2 / 2m = \epsilon_F$. In momentum space all levels inside ($k < k_F$) of the Fermi surface ($k = k_F$) are occupied, and the ones outside are empty.



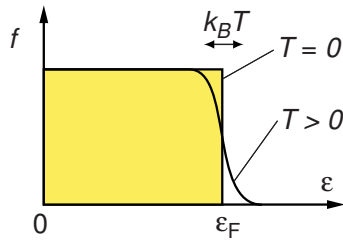
Equation (73) gives at zero temperature

$$N = 2 \sum_{k < k_F} 1 = 2 \frac{\frac{4}{3}\pi k_F^3}{(2\pi/L)^3},$$

where the factor 2 comes from spin. From this we get a relation between the Fermi wave vector and the particle density,

$$\frac{N}{V} = \frac{k_F^3}{3\pi^2}. \quad (77)$$

When $T > 0$, the occupation $f(\epsilon)$ gets rounded so that the change from $f \approx 1$ to $f \approx 0$ takes place in the energy interval $\approx k_B T$.



Conclusion

This finishes our introduction to thermodynamics and statistical physics. I hope there has been some new aspect to your previous knowledge, and that this basic knowledge would be helpful in further lectures of the Cryocourse.