

FYS4130 - STATISTICAL MECHANICS

UNIVERSITY OF OSLO



Obligatory assignment

Even Marius Nordhagen

April 13, 2018

1 CLUSTER SIZE

This problem aims at a one dimensional lattice with N sites at a given temperature. There is an atom at each site that can either have energy $+\epsilon$ or $-\epsilon$, and a chain of sites with energy $+\epsilon$ is called a cluster.

1.1 Onebody probabilities

Since each atom can be in two possible energy states, the one-particle partition function has two terms

$$Z_1 = \sum_i e^{-\beta\epsilon_i} = e^{-\beta\epsilon} + e^{\beta\epsilon} \quad (1)$$

and the probability of an atom having energy $+\epsilon$ and $-\epsilon$ is

$$P_+ = \frac{e^{-\beta\epsilon}}{Z_1} = \frac{e^{-\beta\epsilon}}{e^{-\beta\epsilon} + e^{\beta\epsilon}} \quad \text{and} \quad P_- = \frac{e^{\beta\epsilon}}{Z_1} = \frac{e^{\beta\epsilon}}{e^{-\beta\epsilon} + e^{\beta\epsilon}} \quad (2)$$

respectively.

1.2 Cluster number density

Here we will find the cluster number density, i.e, the probability, P_L , that a site belongs to a cluster of length L . The probability that a cluster of length L occurs is the joint probability consisting of the probabilities of each site in the chain that creates the cluster,

$$n(L, P_+) = P_-^2 \cdot P_+^L = (1 - P_+)^2 P_+^L, \quad (3)$$

because the site to the left of the cluster has energy $-\epsilon$ and so for the site to the right. This is also the probability that a site is at a specific position in the cluster, and since that can happen in L ways, we obtain

$$P_L = L(1 - P_+)^2 P_+^L = L \left(\frac{e^{\beta\epsilon}}{e^{-\beta\epsilon} + e^{\beta\epsilon}} \right)^2 \left(\frac{e^{-\beta\epsilon}}{e^{-\beta\epsilon} + e^{\beta\epsilon}} \right)^L \quad (4)$$

1.3 Normalization of the cluster number density

We will now ensure that the sum over all cluster probabilities P_L is P_+ , which is a criterion since the probability of finding a site in a cluster of any size should be P_+ .

$$\sum_{L=1}^{\infty} P_L = \sum_{L=1}^{\infty} L(1 - P_+)^2 P_+^L = (1 - P_+)^2 P_+ \sum_{L=1}^{\infty} L P_+^{L-1} \quad (5)$$

where P_+^L is splitted up. We will now do a trick

$$\sum_{x=1}^{\infty} x p^{x-1} = \frac{d}{dp} \sum_{x=0}^{\infty} p^x = \frac{d}{dp} \left(\frac{1}{1-p} \right) = \frac{1}{(1-p)^2} \quad (6)$$

where we also have used a common series. Setting back in equation [5], we find

$$(1 - P_+)^2 P_+ \frac{1}{(1 - P_+)^2} = P_+ \quad (7)$$

A cluster of length $L = 0$ is not defined as a cluster, and the corresponding probability is therefore zero.

1.4 Average length of a cluster

The average length of a cluster is given by

$$\langle L \rangle = \sum_{L=0}^{\infty} L^2 (1 - P_+)^2 P_+^L \quad (8)$$

from the general average statement for discrete values. We can let the sum run from $L = 0$ because this just adds a zero term to the sum.

$$\langle L \rangle = (1 - P_+)^2 \sum_{L=0}^{\infty} L^2 P_+^L \quad (9)$$

$$= (1 - P_+)^2 \left(\frac{P_+(P_+ + 1)}{(1 - P_+)^3} \right) \quad (10)$$

$$= \frac{P_+(P_+ + 1)}{1 - P_+} \quad (11)$$

$$(12)$$

where we again have replaced the sum with the known answer. We can now insert the onebody probabilities. Observe that the denominator is P_- and $P_+/P_- = e^{-2\beta\epsilon}$, and get

$$\langle L \rangle = e^{-2\beta\epsilon} \left(\frac{e^{\beta\epsilon} + 2e^{-\beta\epsilon}}{e^{\beta\epsilon} + e^{-\beta\epsilon}} \right). \quad (13)$$

The average length is plotted as a function of the temperature in figure [1], and we see that the average length decreases rapidly when $T \rightarrow 0$. The low-temperature limit is not very informative when we already know that

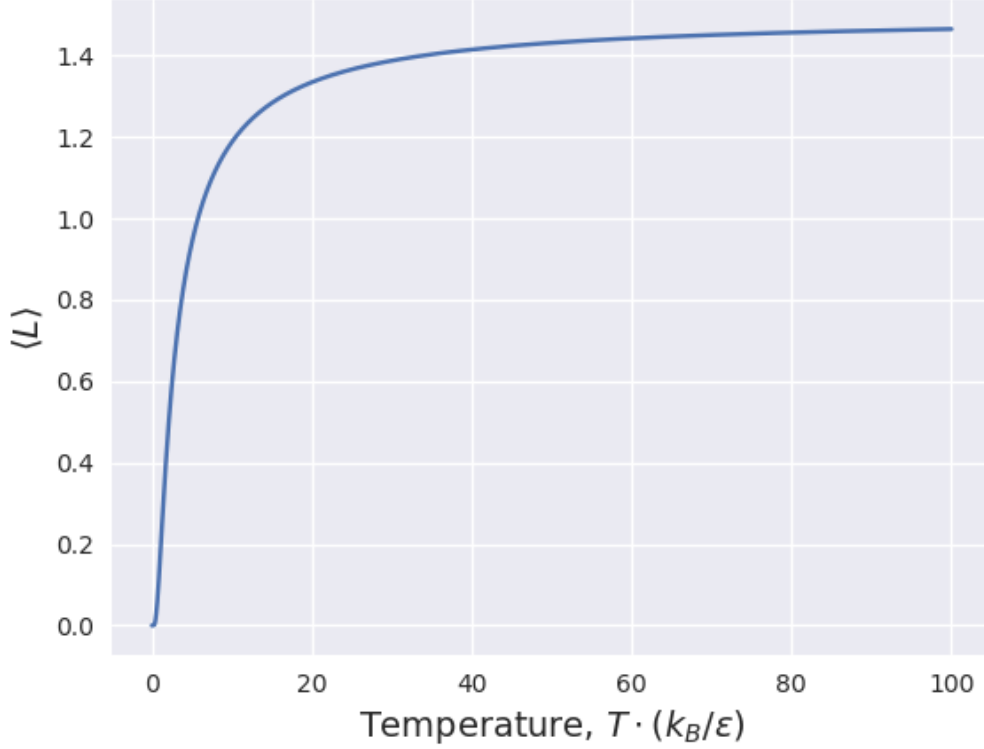


Figure 1: Average length of cluster as a function of temperature. In this plot the temperature spans from 0 to 100 given in units of ϵ/k_B where k_B is Boltzmanns constant.

the average length goes to zero when the temperature approaches zero, but it is interesting to see what the high-temperature asymptote is.

The high-temperature limit is

$$\lim_{T \rightarrow \infty} \langle L \rangle = 1 \left(\frac{1+2}{1+1} \right) = \frac{3}{2} \quad (14)$$

which matches the plot. The low-temperature limit is

$$\lim_{T \rightarrow \infty} \langle L \rangle = \lim_{T \rightarrow \infty} \frac{1 + 2e^{-2\epsilon/kT}}{1 + e^{2\epsilon/kT}} = \frac{1+0}{1+\infty} = 0. \quad (15)$$

as expected. We can connect these numbers to the onebody probabilities, which can be rewritten as

$$P_+ = \frac{1}{1 + e^{2\beta\epsilon}} \quad \text{and} \quad P_- = \frac{1}{1 + e^{-2\beta\epsilon}}. \quad (16)$$

When $T \rightarrow 0$ all the sites will have energy $-\epsilon$, and we will not have any clusters (therefore the average cluster length goes to zero). On the other hand, when $T \rightarrow \infty$, an equal number of sites have energy ϵ as having $-\epsilon$, and the most likely configurations give an average cluster length between 1 and 2.

2 ELECTRON GAS IN A MAGNETIC FIELD

Here we study an electron gas in a magnetic field, where each particular electron has energy $\varepsilon_T = \varepsilon \mp \mu_B H$ depending on whether its magnetic moment is parallel or antiparallel to the field. ε_T is the total energy of an electron, ε is its kinetic energy and $\pm \mu_B H$ is the magnetic energy.

2.1 Average number of particles

For continuous energies, the average number of fermions with infinitesimal energy $d\varepsilon$ is

$$dN(\varepsilon) = g(\varepsilon)\bar{n}(\varepsilon_T)d\varepsilon \quad (17)$$

and the average total number of particles is

$$\langle N \rangle = \int_0^\infty d\varepsilon g(\varepsilon)\bar{n}(\varepsilon_T). \quad (18)$$

Notice that the density function $g(\varepsilon)$ takes the kinetic energy, while the distribution function $\bar{n}(\varepsilon_T)$ takes the total energy. The electrons are Fermi-Dirac distributed and in our case the density of states is $g(\varepsilon) = AV\varepsilon^{1/2}$ with $A = (2/\sqrt{\pi})(2\pi m/\hbar^2)^{3/2}$, such that

$$\langle N \rangle = AV \int_0^\infty d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon_T - \mu)} + 1}. \quad (19)$$

When inserting for the total energy, we need to separate the average depending on whether the magnetic energy is positive or negative. We get

$$\langle N_\pm \rangle = AV \int_0^\infty d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon \mp \mu_B H - \mu)} + 1} \quad (20)$$

where we can define $x = \mu \pm \mu_B H$. The integral is quite tricky, but we can express it in terms of the Sommerfeld expansion:

$$\int_0^\infty d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon - x)} + 1} = \frac{2}{3}x^{3/2} + \frac{\pi^2}{12}(kT)^2x^{-1/2} + \dots \quad (21)$$

shown in the appendix. We are left with

$$\langle N_\pm \rangle \cong \frac{2V}{\sqrt{\pi}} \left(\frac{2\pi m}{\hbar^2} \right)^{3/2} \left(\frac{2}{3}(\mu \pm \mu_B H)^{3/2} + \frac{\pi^2}{12}(kT)^2(\mu \pm \mu_B H)^{-1/2} \right) \quad (22)$$

which is the average number of particles with magnetic moment parallel and antiparallel to the magnetic field at temperature T . As we can see in figure

[2], both these averages increase rapidly when the temperature is increased, which is allowed in the grand-canonical ensemble. What is interesting is the difference we observe for really low temperatures, where we still have some electrons with magnetic moment antiparallel to the field.

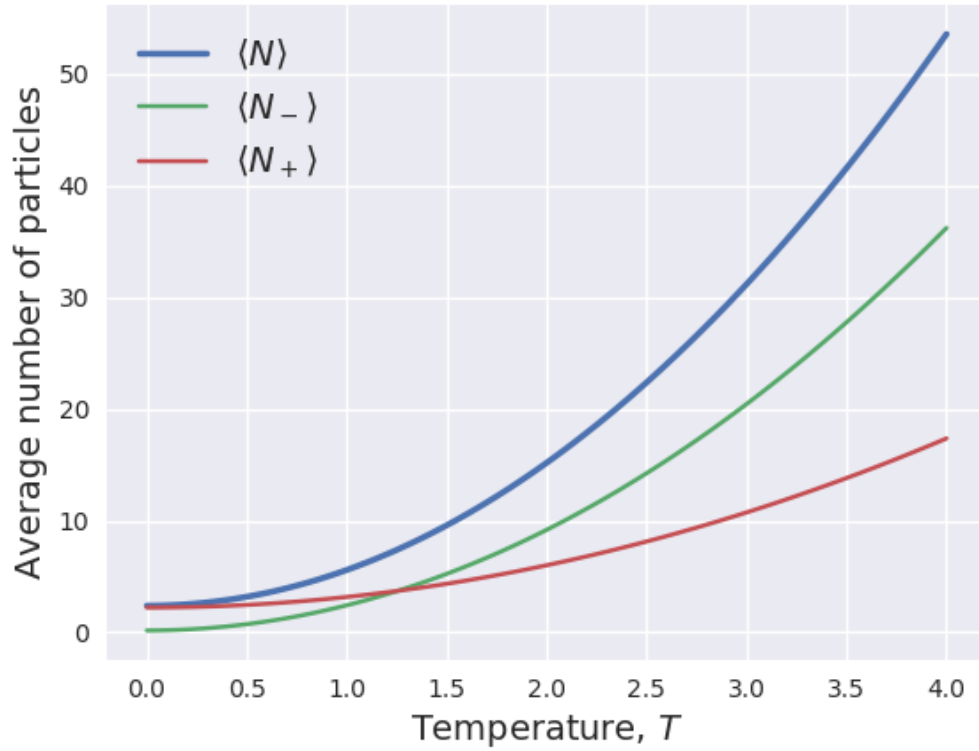


Figure 2: Average number of electrons as a function of the temperature. In this plot the temperature ranges from 0 to 4, while $\mu = 1.0$ and $\mu_B H = 0.7$ just to get a illustration. The numbers should but be taken too serious, but the shape of the graphs should be correct.

2.2 Average magnetization

The average magnetization is given by

$$\langle M \rangle = \mu_B (\langle N_+ \rangle - \langle N_- \rangle) \quad (23)$$

$$= AV\mu_B \left(\frac{2}{3} \left((\mu + \mu_B H)^{3/2} - (\mu - \mu_B H)^{3/2} \right) + \frac{\pi^2}{12} (kT)^2 \left((\mu + \mu_B H)^{-1/2} - (\mu - \mu_B H)^{-1/2} \right) \right) \quad (24)$$

inserted the expression from previous exercise. For low temperatures the second line vanishes, and we get

$$\langle M \rangle = \frac{2}{3} AV\mu_B \left((\mu + \mu_B H)^{3/2} - (\mu - \mu_B H)^{3/2} \right) \quad (25)$$

$$\cong 2AV\mu^{1/2}\mu_B^2 H \quad (26)$$

where the expression is expanded with $\mu \ll \mu_B H$ as the condition.

2.3 Susceptibility

The susceptibility is defined as

$$\chi = \lim_{H \rightarrow 0} \frac{\partial \langle M \rangle}{\partial H} \quad (27)$$

which can now be found. The partial derivative is

$$\frac{\partial \langle M \rangle}{\partial H} = AV\mu_B^2 \left((\mu + \mu_B H)^{1/2} + (\mu - \mu_B H)^{1/2} - \frac{\pi^2}{24} (kT)^2 \left((\mu + \mu_B H)^{-3/2} + (\mu - \mu_B H)^{-3/2} \right) \right) \quad (28)$$

and when we take the limit $\lim_{H \rightarrow 0}$ we get

$$\chi = AV\mu_B^2 \left(2\mu^{1/2} - \frac{\pi^2}{24} (kT)^2 \cdot 2\mu^{-3/2} \right) \quad (29)$$

$$= 2AV\mu_B^2 \mu^{1/2} \left(1 - \frac{\pi^2}{24} \left(\frac{kT}{\mu} \right)^2 \right). \quad (30)$$

For low temperatures the second term vanishes, and we get

$$\lim_{T \rightarrow 0} \chi = 2AV\mu^{1/2}\mu_B^2. \quad (31)$$

Firstly we observe that its positive, which is expected when we deal with a paramagnet. From figure [2] we know that we should only have electrons with magnetic moment antiparallel to the field at zero temperature, which causes this positive sign. Secondly this result is consistent with the expansion in the previous exercise.

3 QUANTUM COULOMB GAS

We will look at a relativistic quantum gas of positive and negative charged particles with coulomb interaction. In addition we have an external box potential with walls of length L . Although we cannot solve this analytically, we will get some results. The given Hamiltonian is as follows

$$\mathcal{H} = \sum_{i=1}^{2N} c|\mathbf{p}_i| + \sum_{i < j}^{2N} \frac{e_i e_j}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (32)$$

3.1 Schrödinger equation

The general Schrödinger equation for a many-body system is

$$H\Psi_n = E_n\Psi_n \quad (33)$$

where Ψ_n is the total wave function of state n and E_n is the corresponding total energy. The wave functions are position dependent and the energies are dependent on the box size,

$$\mathcal{H}\Psi_n(\{\mathbf{r}_i\}) = \epsilon_n(L)\Psi_n(\{\mathbf{r}_i\}). \quad (34)$$

The exact wave functions are unknown, but there are some constraints that apply in general. Since the particles are indistinguishable, all observable should be the same although we swap two coordinates. This results in

$$P(a, b) = P(b, a) \quad \Rightarrow \quad |\Psi(a, b)|^2 = |\Psi(b, a)|^2 \quad (35)$$

and

$$\Psi(a, b) = e^{i\phi}\Psi(b, a). \quad (36)$$

Swapping twice must give back the initial wave function, so there are two possible choices of ϕ : 0 and 2π . The first one gives a symmetric total wave-function under exchange of two particles, which is antisymmetric for the second choice. We denote the former as bosons and the latter as fermions, where the Pauli principle is a consequence of antisymmetry.

3.2 Scaling

We will now scale the coordinates with respect to L such that the box gets size 1^d where d is the number of dimensions, mathematically written as

$$\mathbf{r}'_i = \mathbf{r}_i/L. \quad (37)$$

Using that the operator of linear momentum is given by

$$\mathbf{p}_i = -i\hbar\nabla_i = -i\hbar\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_d}\right), \quad (38)$$

the corresponding scaled momentum operator is

$$\mathbf{p}'_i = -i\hbar\left(\frac{\partial}{\partial x'_1}, \frac{\partial}{\partial x'_2}, \dots, \frac{\partial}{\partial x'_d}\right) = \frac{\mathbf{p}_i}{L} \quad (39)$$

which is the scaling of the first term of the Hamiltonian. The second term is scaled similarly, and by denoting the scaled Hamiltonian as \mathcal{H}' , we get $\mathcal{H} = \mathcal{H}'/L$. We are now set to scale the Schrödinger equation, which becomes

$$\frac{\mathcal{H}'}{L}\Psi_n(\{\mathbf{r}_i\}) = \frac{\epsilon_n(1)}{L}\Psi_n(\{\mathbf{r}_i\}) \quad (40)$$

where the scaled energy is a function of 1 since the length scale now is 1, such that $\epsilon_n(L) = \epsilon_n(1)/L$.

3.3 Connection between V and T^3

We have a fixed number of particles and the energy can fluctuate, so we are in the canonical ensemble. The one particle partition function is given by

$$Z_1 = \sum_{n=0}^{\infty} e^{-\beta\epsilon_n(L)} \quad (41)$$

which in three dimensions is given by

$$Z_1(V, T) = \sum_{n_x, n_y, n_z=0}^{\infty} \exp\left(-\beta^3\epsilon_{n_x}(L)\epsilon_{n_y}(L)\epsilon_{n_z}(L)\right) \quad (42)$$

$$= \sum_{n_x, n_y, n_z=0}^{\infty} \exp\left(-\frac{\epsilon_{n_x}(1)\epsilon_{n_y}(1)\epsilon_{n_z}(1)}{k^3 T^3 L^3}\right). \quad (43)$$

Since $L^3 = V$, the one particle partition function is a function of VT^3 . The total partition function is given by

$$Z(N, V, T) = \frac{Z_1(V, T)^N}{N!} = \frac{1}{N!} \sum_{n_{\mathbf{k}}=0}^{\infty} e^{-N\frac{\epsilon_{n_{\mathbf{k}}}(1)}{k^3 VT^3}} = \mathcal{Z}(N, VT^3) \quad (44)$$

since the particles are indistinguishable.

3.4 Energy and pressure relation

The internal energy in this system is a sum over all energy levels times the probability of that level. Further we observe that we can differentiate the exponential with respect to β to obtain the same expression,

$$E = \frac{1}{Z} \sum_n \epsilon_n e^{-\beta \epsilon_n} = -\frac{1}{Z} \sum_n \frac{\partial}{\partial \beta} e^{-\beta \epsilon_n} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta}. \quad (45)$$

The pressure is given by the differential

$$P(N, V) = -\left(\frac{\partial F}{\partial V}\right)_N \quad (46)$$

where F is Helmholtz free energy,

$$F = -kT \ln Z. \quad (47)$$

We therefore end up with the expressions

$$E = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \quad (48)$$

$$P = kT \frac{1}{Z} \frac{\partial Z}{\partial V} \quad (49)$$

such that they are related through the partition function. We use the same partition function as we found in the previous exercise, and obtain

$$\frac{\partial Z}{\partial \beta} = -3 \frac{\beta^2 \epsilon_{n_{\mathbf{k}}}(1)}{V} Z \quad (50)$$

$$\frac{\partial Z}{\partial V} = \frac{\beta^3 \epsilon_{n_{\mathbf{k}}}(1)}{V^2} Z. \quad (51)$$

where $\epsilon_{n_{\mathbf{k}}}(1) \equiv \epsilon_{n_x}(1)\epsilon_{n_y}(1)\epsilon_{n_z}(1)$. The energy then reads

$$E = 3 \frac{\beta^2 \epsilon_{n_{\mathbf{k}}}(1)}{V} = 3kT \frac{\beta^3 \epsilon_{n_{\mathbf{k}}}(1)}{V^2} V = 3PV. \quad (52)$$

Notice that the calculations are done for the one particle partition function, but it is easy to imagine that the full partition function will give the same result.

3.5 Non-relativistic gas

We now switch from a relativistic gas to a non-relativistic one and redo the exercises in four dimensions. The Hamiltonian is then given by

$$\mathcal{H} = \sum_{i=1}^{2N} \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j}^{2N} \frac{e_i e_j}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (53)$$

We then do a change of scale similarly as in exercise 3.2, and get the Schrödinger equation

$$\frac{\mathcal{H}'}{L^2} \Psi_n(\{\mathbf{r}_i\}) = \frac{\epsilon_n(1)}{L^2} \Psi_n(\{\mathbf{r}_i\}) \quad (54)$$

where \mathcal{H}' is the scaled Hamiltonian. Further the partition function is

$$Z_1 = \sum_{n_x, n_y, n_z, n_u=0}^{\infty} \exp\left(-\beta^4 \epsilon_{n_x}(L) \epsilon_{n_y}(L) \epsilon_{n_z}(L) \epsilon_{n_u}(L)\right) \quad (55)$$

$$= \sum_{n_x, n_y, n_z, n_u=0}^{\infty} \exp\left(-\frac{\beta^4 \epsilon_{n_x}(1) \epsilon_{n_y}(1) \epsilon_{n_z}(1) \epsilon_{n_u}(1)}{L^8}\right) \quad (56)$$

and since we are in 4D, the 4-volume is $V_4 = L^4$. We can reuse the expressions for energy and pressure from above, and find a relation between the energy and pressure as we did in exercise 3.4. Again we calculate the differentials

$$\frac{\partial Z}{\partial \beta} = -4 \frac{\beta^3 \epsilon_{n_{\mathbf{k}}}(1)}{V_4^2} Z \quad (57)$$

$$\frac{\partial Z}{\partial V_4} = 2 \frac{\beta^4 \epsilon_{n_{\mathbf{k}}}(1)}{V_4^3} Z. \quad (58)$$

and find the relation

$$E = 4 \frac{\beta^3 \epsilon_{n_{\mathbf{k}}}(1)}{V_4^2} = 4kT \frac{\beta^4 \epsilon_{n_{\mathbf{k}}}(1)}{V_4^3} = 2PV \quad (59)$$

Again this result is found from the one particle partition function, but it is general for a non-relativistic interacting quantum gas in 4D.

4 NON-INTERACTING FERMIONS

In this problem we study a non-interacting fermi gas with chemical potential μ . The particle states are given by the vector \vec{k} , which is d -dimensional where d is number of spatial dimensions, and the corresponding energy is $\epsilon(\vec{k})$.

4.1 Joint probability

The joint probability of finding a particle in a state given by the occupation numbers $\{n_{\vec{k}}\}$ is given by a product of the general probability statement in a grand canonical ensemble (hence joint),

$$P(\{n_k\}) = \frac{\prod_k e^{-\beta n_k (\epsilon_k - \mu)}}{\Xi}, \quad (60)$$

where Ξ is the grand canonical partition function, given by

$$\Xi = \sum_{\{n_k\}} \prod_k \exp(-\beta(\epsilon_k - \mu)n_k) \quad (61)$$

$$= \prod_k \sum_{n_k=0}^1 \exp(-\beta(\epsilon_k - \mu)n_k) \quad (62)$$

$$= \prod_k \left(1 + \exp(-\beta(\epsilon_k - \mu))\right) \quad (63)$$

since we only can have 0 or 1 fermion in each energy state according to the Pauli principle.

$$P(\{n_k\}) = \frac{\prod_k e^{-\beta n_k (\epsilon_k - \mu)}}{\prod_k (1 + e^{-\beta(\epsilon_k - \mu)})} = \prod_k \frac{e^{\beta n_k (\mu - \epsilon_k)}}{1 + e^{\beta(\mu - \epsilon_k)}} \quad (64)$$

4.2 Average occupation number

The average occupation number of a state is given by Fermi-Dirac distribution

$$\langle n_k \rangle = \frac{1}{e^{\beta(\epsilon_k - \mu)} + 1}. \quad (65)$$

We can easily write the probability as a function of $\langle n_k \rangle$,

$$P(\{n_k\}) = \prod_k \frac{e^{\beta(\epsilon_k - \mu)(1 - n_k)}}{e^{\beta(\epsilon_k - \mu)}} \quad (66)$$

$$= \prod_k \langle n_k \rangle e^{\beta(\epsilon_k - \mu)(1 - n_k)} \quad (67)$$

$$= \prod_k \langle n_k \rangle \frac{e^{\beta(\epsilon_k - \mu)}}{e^{\beta(\epsilon_k - \mu)n_k}} \quad (68)$$

4.3 Maximum entropy of random variable

A random variable has l discrete outcomes with their own respective probabilities. Since the probabilities are independent, we can find the information entropy

$$S = H(p_1, \dots, p_n) = - \sum_{n=1}^l p_n \log p_n \quad (69)$$

where p_i is the probability of outcome i . We cannot find the exact entropy without knowing the probabilities, but we can find which probabilities that give maximum entropy. For this we use Lagrange multipliers with the normalization condition as the constraint,

$$\sum_{n=1}^L p_n = 1 \quad \Rightarrow \quad G(p_1, \dots, p_l) = \sum_{n=1}^L p_n - 1 = 0. \quad (70)$$

The Lagrangian is thus given by

$$\mathcal{L}(p_1, \dots, p_l, \lambda) = H(p_1, \dots, p_l) + \lambda \cdot G(p_1, \dots, p_l) \quad (71)$$

$$= - \sum_{n=1}^l p_n \log p_n + \lambda \left(\sum_{n=1}^l p_n - 1 \right) \quad (72)$$

which we maximize to find the maximum entropy. The gradient with respect to all variables is

$$\nabla_{p_1, \dots, p_l, \lambda} \mathcal{L}(p_1, \dots, p_l, \lambda) = \left(\frac{\partial \mathcal{L}}{\partial p_1}, \dots, \frac{\partial \mathcal{L}}{\partial p_l}, \frac{\partial \mathcal{L}}{\partial \lambda} \right) \quad (73)$$

$$\frac{\partial \mathcal{L}}{\partial p_j} = -\log p_j - 1 + \lambda, \quad \frac{\partial \mathcal{L}}{\partial \lambda} = \sum_{n=1}^l p_n - 1 \quad (74)$$

which is maximized when $\nabla \mathcal{L} = \vec{0}$. This occurs when

$$p_j = e^{\lambda-1} = \frac{1}{l} = P \quad (75)$$

such that all probabilities are equal and $P = 1/l$ due to the normalization. The information entropy is therefore maximized as

$$S = -lP \log P = \log l \quad (76)$$

This also matches to the principle of maximum entropy, which states that

”If nothing is known about a distribution except that it belongs to a certain class, then the distribution with the largest entropy should be chosen as the least-informative default”.

The least-informative default is clearly when all probabilities are equal.

4.4 Entropy of fermi gas

We will now find the entropy of the gas discussed in exercises 4.1 and 4.2. Recall the average occupation number

$$\langle n_k \rangle = \frac{1}{e^{\beta(\epsilon_k - \mu)} + 1} = \frac{n_k}{g_k} \quad (77)$$

where n_k is the total number of particles and g_k is the density function. A fermi gas has in general a number of available states given by

$$W = \prod_k \frac{g_k!}{n_k!(g_k - n_k)!}. \quad (78)$$

We can then find the (Boltzmann) entropy by

$$S = k \log W \quad (79)$$

$$= k \sum_k \log \left(\frac{g_k!}{n_k!(g_k - n_k)!} \right) \quad (80)$$

$$= k \sum_k g_k \left(\log g_k - \frac{n_k}{g_k} \log n_k - \left(1 - \frac{n_k}{g_k} \right) \left(\log g_k + \log \left(1 - \frac{n_k}{g_k} \right) \right) \right) \quad (81)$$

$$= k \sum_k g_k \left(-\langle n_k \rangle \log \langle n_k \rangle + (1 - \langle n_k \rangle) \log (1 - \langle n_k \rangle) \right) \quad (82)$$

where Sterling's approximation is applied. The average occupation number from equation [77] can now be inserted, and we get

$$S = k \sum_k g_k \left(\log(e^{-\beta(\epsilon_k - \mu)} + 1) + \frac{\log(e^{\beta(\epsilon_k - \mu)} + 1) - \log(e^{-\beta(\epsilon_k - \mu)} + 1)}{e^{\beta(\epsilon_k - \mu)} + 1} \right)$$

For the low-temperature limit some terms vanish and we are left with

$$\lim_{T \rightarrow 0} S = \frac{\log(e^{\beta(\epsilon_k - \mu)} + 1)}{e^{\beta(\epsilon_k - \mu)} + 1} \quad (83)$$

which is a " ∞/∞ -expression". We solve it by L'Hôpital's rule:

$$\lim_{T \rightarrow 0} S = k \sum_k g_k \frac{\frac{d}{dT} \log(e^{\beta(\epsilon_k - \mu)} + 1)}{\frac{d}{dT} e^{\beta(\epsilon_k - \mu)}} \quad (84)$$

$$= k \sum_k g_k \frac{(e^{\beta(\epsilon_k - \mu)} + 1)^{-1} \cdot \left(-\frac{1}{kT^2} (\epsilon_k - \mu) \right) e^{\beta(\epsilon_k - \mu)}}{\left(-\frac{1}{kT^2} (\epsilon_k - \mu) \right) e^{\beta(\epsilon_k - \mu)}} \quad (85)$$

$$= k \sum_k \frac{g_k}{e^{\beta(\epsilon_k - \mu)} + 1} \quad (86)$$

which is the average number of fermions times Boltzmann's constant. At zero temperature the energy levels are fully occupied up to the Fermi energy, and the distribution becomes a step function.

5 Appendix

In this appendix we will study how to express the integral

$$\int_0^\infty d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon-x)} + 1} \quad (87)$$

in terms of the Sommerfeld expansion. This is what we used in section 2. We start rewriting the Fermi-Dirac distribution as

$$\bar{n}(\varepsilon) = \frac{1}{2} - \frac{1}{2} \tanh(\beta(\varepsilon - x)/2) \quad (88)$$

which turns the integral into

$$\int_0^\infty d\varepsilon \varepsilon^{1/2} \left(\frac{1}{2} - \frac{1}{2} \tanh(\beta(\varepsilon - x)/2) \right). \quad (89)$$

Thereafter we integrate by parts, and since

$$\bar{n}'(\varepsilon) = -\frac{\beta}{4} \frac{1}{\cosh^2(\beta(\varepsilon - x)/2)}, \quad (90)$$

the integral now goes like

$$\int_0^\infty d\varepsilon \left(\frac{2}{3} \varepsilon^{3/2} \frac{\beta}{4} \frac{1}{\cosh^2(\beta(\varepsilon - x)/2)} \right) \quad (91)$$

where the constant term from integration by parts disappeared due to symmetry. The next step is to evaluate this integral around $\varepsilon = x$, getting

$$\int_0^\infty d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon-x)} + 1} \cong x^{3/2} \left[1 + \frac{\pi^2}{8} \left(\frac{kT}{x} \right)^2 + \dots \right] \quad (92)$$

where the Sommerfeld expansion is applied, including the two largest terms. For further reading see compendium p.118.