Notes for Lecture 21

Interacting particles (SCSM)

In this out-of-sequence lecture, we go back to the semi-classical statistical mechanics, and consider interactions. This is a good topic to treat out of sequence, but it is also a good topic to learn, as we are going into the topic of phase transition, in the regular lecture.

Generally, interactions make problems extremely difficult, if not impossible. One way to say this is that we do not have a general, reliable way of calculating the properties of a liquid or a solid from first principles, i.e. starting from the N body Hamiltonian, even for the semi-classical case. For a low-density gas, we do have a systematic way of doing it, namely the Mayer cluster expansion, as we will see here. The difference between gas and liquid/solid is that in a gas there is a very small chance that a gas molecule belongs in a small "cluster" of molecules. Essentially, the gas molecule spends most of the time as a lone molecule (cluster of size 1), while occasionally colliding with other molecules forming a cluster of size 2 or 3 or more, with an increasingly smaller probability to do so as the cluster size increases. In a liquid or a solid, this idea fails completely, since the specimen is so dense that the "cluster" that a molecule belongs is always the total group of *all* the molecules.

21.1 Nature of the problem

Let us start by reminding ourselves of the result of the non-interacting problem, Eq 6.44, which we now refer to as Z_0.

$$Z_0(N,T,V) = \frac{1}{N!}\left(\frac{V}{\lambda(T)^3}\right)^N$$

where $\lambda(T) = h/\sqrt{2\pi m k_B T}$ is the thermal de Broglie wavelength. This partition function resulted from the non-interacting Hamiltonian, $\sum_i p_i^2/(2m)$. Now, let us add the interaction term

$$H_{int}(\vec{q}) = \sum_{i>j} U(\vec{q}_i - \vec{q}_j) \equiv \sum_{i>j} U_{ij} = \frac{1}{2}\sum_{i\neq j} U_{ij} \tag{21.1}$$

where $\vec{q}_i$'s are generalized coordinates. What would be the change to the partition function? Since this interaction term does not depend on the momentum variables at all, we get

$$Z(N,T,V) = Z_0(N,T,V)\,\frac{R(N,T,V)}{V^N} \tag{21.2}$$

$$R(N,T,V) = \int \prod_{i=1}^{N} d^3\vec{q}_i \,\exp\left(-\beta H_{int}\right) \tag{21.3}$$

The integral of H_{int} is impossible to do, in general.

One way to think about R is that R/V^N is the ensemble average value of $\exp(-\beta H_{int})$ in the *non-interacting* state. That is,

$$\frac{R}{V^N} = \langle \exp(-\beta H_{int})\rangle^0 \tag{21.4}$$

where the superscript 0 is used for the expectation value in the non-interacting state. Written this way, it is clear that R/V^N is the characteristic function of the probability distribution of the canonical ensemble for the non-interacting system, with a complex "wave number," $k = i\beta$. Thus, it follows that

$$\log R - N\log V = \sum_{n=1}^{\infty} \frac{(-\beta)^n}{n!}\,\langle H_{int}^n\rangle_c^0 \tag{21.5}$$

Taking this approach, the cumulants of H_{int} can be analyzed in terms of the cumulants of U_{ij}.

While formally attractive, this cumulant approach is not good when applied to intermolecular interaction problem, as we here consider for the interacting semi-classical gas. The reason is that the intermolecular potential diverges as the intermolecular distance vanishes, making the cumulants diverge in general. Instead, it is much better to stay with $\langle \exp(-\beta H_{int})\rangle^0$.

21.2 Perturbation treatment of the grand partition function

The grand partition function can be written as

$$\mathcal{Q}(\mu,T,V) = \sum_{\mathcal{N}=0}^{\infty} \exp(\beta\mu\mathcal{N}) Z(\mathcal{N},T,V) \tag{21.6}$$

where, for $\mathcal{N}=0$, $Z=1$[1].

Also, note that

$$\mathcal{G}(\mu,T,V) \equiv -k_B T \log \mathcal{Q} = E - TS - \mu N = -PV \tag{21.7}$$

where the assumption of the extensivity is used in the last step (Section 2.4.1: $E = TS + Jx + \mu N$).

We will also use the following very useful relation (from Eq. 7.13)

$$N = \frac{1}{\beta} \left. \frac{\partial \log \mathcal{Q}}{\partial \mu} \right|_{T,V}$$

Note that, in terms of the notation, N is used as the ensemble average for $\mathcal{N}$.

21.2.1 First order

Let us see what the $\mathcal{N}=1$ term looks like. Using Eqs. 21.2,21.3 and the first equation of this note, we get

$$Z(\mathcal{N}=1,T,V) = \frac{1}{\lambda^3}\int d^3\vec{q} \qquad \text{since } H_{int}=0 \text{ for } N \le 1 \tag{21.8}$$

$$= \frac{V}{\lambda^3} \qquad \text{same as } Z_0(N=1,T,V) \tag{21.9}$$

So, it follows that, if we keep terms up to $\mathcal{N}=1$, we get

$$\mathcal{Q}^{(1)} = 1 + \frac{e^{\beta\mu}V}{\lambda^3} \tag{21.10}$$

$$PV = k_B T \log \mathcal{Q}^{(1)} = k_B T \log\left(1 + \frac{e^{\beta\mu}V}{\lambda^3}\right)$$

$$\approx k_B T \frac{e^{\beta\mu}V}{\lambda^3} = N k_B T \tag{21.11}$$

---

[1] You must absolutely convince yourself why this must be!

where the last step is from

$$N = \frac{1}{\beta} \frac{\partial \log \mathcal{Q}^{(1)}}{\partial \mu}\bigg|_{T,V} = \frac{e^{\beta\mu} V}{\lambda^3}$$

Two important things need to be noted. First, we just recovered the ideal gas equation of state. Second, we did this by assuming that a perturbation theory is possible, with

$$N \approx \frac{e^{\beta\mu} V}{\lambda^3} \ll 1 \tag{21.12}$$

Here, I use the $\approx$ sign rather than $=$, since the equality is valid only at this order of the perturbation.

That is, we are using the "hardly any particles" approximation as explained in page 6 of LN 8.

21.2.2 Second order

$$Z(\mathcal{N} = 2, T, V) = \frac{1}{2}\left(\frac{1}{\lambda^3}\right)^2 \int d^3\vec{q}_1 d^3\vec{q}_2 \exp\left(-\beta U(\vec{q}_1, \vec{q}_2)\right) \tag{21.13}$$

The integral here is defined as $V^2 + b_2$. On physical grounds, we expect that the intermolecular interaction show an exponentially diverging repulsion at small distance and a weak van der Waals attraction at large distance. There are various schemes for approximating the intermolecular potential, e.g., the Lennard-Jones potential, to ensure the computational convenience and the retention of the basic physics. Here, we will make a simple model in which the potential is approximated as a hard core repulsion at $r < r_0$ and van der Waals attraction at $r > r_0$.

$$U(\vec{q}_1, \vec{q}_2) = \begin{cases} \infty & r < r_0 \\ -u_0 \frac{r_0^6}{r^6} & r > r_0 \end{cases} \qquad u_0 > 0, r \equiv |\vec{q}_2 - \vec{q}_1| \tag{21.14}$$

The u_0 parameter characterizes the strength of the van der Waals attraction. It is very convenient to define

$$f_{12} \equiv e^{-\beta U(\vec{q}_1, \vec{q}_2)} - 1 = \begin{cases} -1 & r < r_0 \\ -1 + \exp\left(\beta u_0 \frac{r_0^6}{r^6}\right) & r > r_0 \end{cases} \tag{21.15}$$

and note that

$$f_{12} \to \begin{cases} -1 & r < r_0 \\ e^{\beta u_0} - 1 > 0 & r \to r_0 + 0^+ \\ 0^+ & r \to \infty \end{cases} \tag{21.16}$$

So, f_{12} goes from -1 to a positive value at at $r = r_0$ (r_0 is on the order of the molecular radius) and then it decreases steadily to 0^+ as r increases. In particular, we are interested in the high-temperature limit, $\beta u_0 \ll 1$, and in this temperature limit, we get

$$f_{12} \;\equiv\; e^{-\beta U(\vec{q}_1,\vec{q}_2)} - 1 \;\approx\; \begin{cases} -1 & r < r_0 \\ \beta u_0 \frac{r_0^6}{r^6} & r > r_0 \end{cases} \qquad \beta u_0 \ll 1 \tag{21.17}$$

So, we get

$$V^2 + b_2 \;\equiv\; \int d^3\vec{q}_1 d^3\vec{q}_2 \exp\left(-\beta U(\vec{q}_1,\vec{q}_2)\right) \tag{21.18}$$

$$= \int d^3\vec{q}_1 d^3\vec{q}_2 \left(1 + f_{12}\right) \tag{21.19}$$

$$= V^2 + V \int d^3\vec{r} f_{12}. \tag{21.20}$$

Using $\int d^3\vec{r} f_{12} = 4\pi \int dr\, r^2 f_{12} = 4\pi \left(-\frac{r_0^3}{3} + \beta u_0 r_0^6 \int_{r_0}^{\infty} dr \frac{1}{r^4}\right) = \frac{4\pi r_0^3}{3}\left(-1 + \beta u_0\right)$, and defining

$$\Omega \;\equiv\; \frac{4\pi r_0^3}{3} \qquad \text{excluded volume} \tag{21.21}$$

we get

$$V^2 + b_2 \;=\; V^2 + V\,\Omega(\beta u_0 - 1). \tag{21.22}$$

Now, it is clear why we defined the integral as V^2 plus b_2. In any case, we have a second order perturbation solution.

$$\mathcal{Q}^{(2)} \;=\; 1 + V\frac{e^{\beta\mu}}{\lambda^3} + \frac{V}{2}\left(\frac{e^{\beta\mu}}{\lambda^3}\right)^2 (V + \bar{b}_2) \qquad \bar{b}_2 \equiv b_2/V \tag{21.23}$$

$$PV \;=\; k_B T \log \mathcal{Q}^{(2)} \tag{21.24}$$

$$\approx\; k_B T\left[V\frac{e^{\beta\mu}}{\lambda^3} + \frac{V}{2}\left(\frac{e^{\beta\mu}}{\lambda^3}\right)^2 (V + \bar{b}_2) - \frac{1}{2}V^2\left(\frac{e^{\beta\mu}}{\lambda^3}\right)^2\right] \tag{21.25}$$

$$=\; k_B T V\left[\frac{e^{\beta\mu}}{\lambda^3} + \frac{\bar{b}_2}{2}\left(\frac{e^{\beta\mu}}{\lambda^3}\right)^2\right] \tag{21.26}$$

$$N \;=\; \frac{1}{\beta}\left.\frac{\partial \log \mathcal{Q}^{(2)}}{\partial \mu}\right|_{T,V} \;=\; V\left[\frac{e^{\beta\mu}}{\lambda^3} + \bar{b}_2\left(\frac{e^{\beta\mu}}{\lambda^3}\right)^2\right] \tag{21.27}$$

Note here that all expressions PV and N are written up to order "$(e^{\beta\mu})^2$" only since we are doing the perturbation theory up to second order here. Now the last equation can be inverted so that[2]

$$z \equiv \frac{e^{\beta\mu}}{\lambda^3}. \qquad\qquad \text{"fugacity-2"} \tag{21.28}$$

can be expressed in terms of the density. Note from the previous section that N is in fact our perturbation parameter. This means that n is proportional to the perturbation parameter and so, we must keep z up to n^2 also. This is easy to do (by re-arranging the above equation as $z - n = -\bar{b}_2 z^2$ and using the method of iteration on it),

$$z = n - \bar{b}_2 n^2 \tag{21.29}$$

Plugging this into the PV equation above, we get

$$P\beta = n - \frac{\bar{b}_2}{2} n^2. \tag{21.30}$$

Here, $-\frac{\bar{b}_2}{2}$ is the so-called second virial coefficient. Since $\bar{b}_2 = \Omega(\beta u_0 - 1)$ (Eq. 21.22), we can re-arrange this equation as

$$\begin{aligned}
P\beta &= n - \frac{\Omega}{2} u_0 \beta n^2 + \frac{\Omega}{2} n^2 && (21.31)\\
\left(P + \frac{\Omega}{2} u_0 n^2\right) &= n k_B T \left(1 + \frac{n\Omega}{2}\right) && (21.32)\\
\left(P + \frac{\Omega}{2} u_0 n^2\right)\left(1 - \frac{n\Omega}{2}\right) &\approx n k_B T && (21.33)\\
\left(P + \frac{u_0 \Omega}{2} n^2\right)\left(V - \frac{N\Omega}{2}\right) &\approx N k_B T && (21.34)\\
\left(P + a n^2\right)(V - bN) &\approx N k_B T
\end{aligned}$$

So, we get the famous van der Waals equation of state

$$\boxed{\left(P + a n^2\right)(V - bN) = N k_B T \qquad \text{van der Waals equation of state} \qquad (21.35)}$$

where

$$a \equiv \frac{u_0 \Omega}{2}, \tag{21.36}$$

$$b \equiv \frac{\Omega}{2}. \tag{21.37}$$

---

[2] For the classical problem at hand here, it is more convenient to define z as the "fugacity." Note the different symbol used for it from $z \equiv e^{\beta\mu}$ (Eq. 12.23).

The van der Waals equation has the form of the ideal gas equation, except that V is replaced by $V - bN$ and P is replaced by $P + an^2$.

The fact that $V - bN$ replaces V can be interpreted as the decrease of the available volume due to the hard core repulsion. Note that bN is the reduced volume per particle, since the excluded volume per a pair of particles is Ω, and since there are $N(N-1)/2 \approx N^2/2$ pairs.

Next, the fact that $P + an^2$ replaces P means that the pressure decreases, due to the attraction, with respect to the ideal gas with the same temperature and the same effective volume ($V - bN$ not V). Why this decrease? From $dE = -PdV + TdS$, note that P is the rate at which the energy increases in an adiabatic reversible process in which the volume is decreased by a purely mechanical process that does work on the system. With the attractive interaction, it is quite clear that this rate will not be as high as the ideal gas case. The energy will increase due to kinetic energy, for sure, but also the attractive interaction will make it decrease, when the volume is reduced. This reasoning makes it clear that an^2 is precisely the amount by which the energy drops due to attraction, following a unit volume decrease. This can be seen as follows. Note that for a single particle, the attraction energy that it feels is

$$\begin{aligned} \varepsilon_{attr} &= (N-1)\int d^3\vec{r}\,\frac{1}{V}U \quad \text{attractive energy for a single particle, due to all other particles} && (21.38)\\ &= -u_0 n\Omega \quad \text{Eq. 21.14 (only the attractive part, } r > r_0\text{); also } N-1 \to N && (21.39)\\ &= -2na. && (21.40)\end{aligned}$$

Therefore the total attractive energy is given by, multiplying N and then dividing by 2, correcting for double counting,

$$E_{attr} = -Nna. \quad (21.41)$$

Indeed the pressure due to the attractive interaction is given by,

$$P_{attr} = -\frac{\partial E_{attr}}{\partial V} = -an^2, \quad (21.42)$$

which proves our assertion above that an^2 is the amount by which the energy drops due to attraction following a unit volume decrease.

A similar reasoning can be provided using a kinetic argument as well. The pressure on a test surface presented in a fluid is the result of molecules bumping into the test surface in an elastic collision. Assuming that such a test surface has the surface normal in the z direction, the pressure is given by

$$\begin{aligned} P &= \langle n v_z \cdot 2 m v_z \rangle_{v_z > 0} && (21.43)\\ &= nm\left\langle v_z^2 \right\rangle = \frac{2}{3} n \langle \kappa \rangle \qquad \kappa \equiv \frac{mv^2}{2} = \frac{m(v_x^2 + v_y^2 + v_z^2)}{2} && (21.44)\end{aligned}$$

where nv_z is the flux of particles along the z direction, the $2mv_z$ is the momentum transfer to the test surface, $\kappa = \frac{mv^2}{2}$ is the kinetic energy per particle, and the test surface is assumed to be above the gas (so $v_z > 0$ for gas molecules that hit the surface). Note that this result is in precise agreement with Eq. 12.32, in the case of the ideal gas, as it should. However, Eqs. 21.43, 21.44 present a more general form from the kinematic point of view, within the semi-classical statistical mechanics that we are doing here.

In particular, if there is an attractive interaction, then the kinetic energy κ will become less at the surface of the fluid. Why? Since we are considering a low density limit, the mean free path of the intermolecular collision is very large compared to the molecular dimension (or the intermolecular distance)[3] Consider that a molecule hitting the test surface originates from the last collision at distance $\sim l$ from the surface. So, the attraction energy that this particle feels changes from ε_{attr} (Eq. 21.40) to $\frac{1}{2}\varepsilon_{attr}$, since at the surface the integral of U is over only half the solid angle. Following this *increase* of the potential energy, the kinetic energy of the particle goes down by exactly $|\frac{\varepsilon_{attr}}{2}| = na$ per particle hitting the test surface. Now, we can make connection to Eqs. 21.43 and 21.44. Note the following first. Since the local equilibrium is established on the length scale of l, we *cannot* apply the ensemble averages ($\langle v_z^2 \rangle$ etc.) expressed in these equations, without checking their validity, first! We must consider the change of v_z^2 from a purely Newtonian dynamics point of view, for those particles that actually hit the surface. Namely, we must apply the argument that we just considered—the interplay between the kinetic energy and the potential energy—and apply it to Eq. 21.43 (with v_z taken at the value at the moment that particles hit the surface), but we can *not* apply it to Eq. 21.44. Note that the decrease of the kinetic energy na occurs only in $mv_z^2/2$ due to the symmetry of the problem. So, we see that $\delta(v_z^2) = -2na/m$, when the particle arrives at the surface to hit it. Then, by Eq. 21.43, we note that $\delta P = -an^2$, which agrees perfectly with the result of Eq. 21.42. These two points of views, used to derive the same result, are equivalent.

21.3 Mayer cluster expansion

This is a general result using which we can carry out the perturbation theory described in the previous section in a more systematic manner. While its derivation is not complicated, and it is actually interesting from the point of view of the field theory ("Wick's theorem"), we shall not spend any time to derive this result. Readers are

---

[3] One collision occurs when a volume swept by a molecule is equal to $1/n$. So, the mean free path l is given by $l\pi a^2 \sim 1/n$. And so, $l \sim 1/(\pi a^2 n)$. Here, a is the molecular radius ($\sim r_0$), $n \sim 1/r_i^3$, r_i is the intermolecular distance.

referred to Section 4.3 of Feynman or Section 5.2 of Kardar.

Here we simply summarize the goal and the results. The goal is to obtain virial coefficients $B_2(T)$, $B_3(T)$, etc. in

$$\frac{P}{k_B T} = n + B_2(T) n^2 + B_3(T) n^3 + \ldots \tag{21.45}$$

where $n = N/V$. Note that, by definition, virial coefficients can be functions of only T, since they must be intensive, and they cannot depend on n.

The Mayer cluster cluster expansion is given by

$$\log \mathcal{Q} = \sum_{k=1}^{\infty} z^k \frac{b_k}{k!} \tag{21.46}$$

where $z = \frac{e^{\beta\mu}}{\lambda^3}$ as defined above (Eq. 21.28).

Using Eq. 7.13 (with $n = 1$) on the above, we get

$$N = \sum_{k=1}^{\infty} z^k \frac{b_k}{(k-1)!}. \tag{21.47}$$

As $\log \mathcal{Q} = -\mathcal{G}\beta = PV\beta$, and as b_k is proportional to the volume V (see near the end of this section), we can re-write the above two results as

$$\frac{P}{k_B T} = \sum_{k=1}^{\infty} z^k \frac{\text{ƀ}_k}{k!} \qquad\qquad \text{ƀ}_k \equiv b_k/V \tag{21.48}$$

$$n = \sum_{k=1}^{\infty} z^k \frac{\text{ƀ}_k}{(k-1)!}. \tag{21.49}$$

These two equations can be combined by eliminating z, to arrive at the virial expansion

$$\frac{P}{k_B T} = n + \sum_{k=2}^{\infty} B_k n^k, \tag{21.50}$$

$$B_k = \frac{(1-k)}{k!} \text{đ}_k. \qquad\qquad B_k = \text{virial coefficient, } k \geq 2 \tag{21.51}$$

For instance, you are urged to work out directly that

$$B_2(T) = -\frac{\text{ƀ}_2}{2}, \tag{21.52}$$

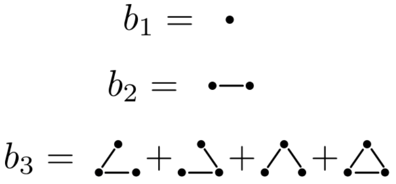
$$B_3(T) = \text{ƀ}_2^2 - \frac{\text{ƀ}_3}{3} \equiv -\frac{1}{3} \text{đ}_3. \tag{21.53}$$

Finally, of course, the symbols b_k and $đ_k$ must be explained.

The symbol b_k stands for the sum of integrals, each of which is diagrammatically represented by a **linked cluster** of k dots. The i-th dot represents the integration variable $\vec{q}_i$. Each pair of dots ("bond") may be connected or not connected. A linked cluster means that all dots are connected to one another. The integrand is 1 for an unconnected bond, and $f_{ij} = \exp(-\beta U_{ij}) - 1$ for a connected bond, where $U_{ij} = U(\vec{q}_i, \vec{q}_j)$ is the pair potential. As all dots are connected in any diagram of b_k, it follows that there is one and only one free variable that can be integrated out to give a factor of V for any diagram of b_k. And so, any diagram contributing to b_k is of order V. This is why we defined above $ƀ_k(T) \equiv b_k/V$: $ƀ_k(T)$ is now independent of V and so is a function of T only. (Note that we used symbols b_2 and $ƀ_2$ just the same way in Section 21.2, anticipating more general results given here.)

The symbol $đ_k$ stands for the sum of of integrals that correspond to a special subset—the **one particle irreducible diagrams**—of those diagrams for $ƀ_k$. The one particle irreducible diagrams are those diagrams that cannot be separated into two disjoint diagrams by removing one dot.

Here is a diagram for the fist three b_k's.



To re-cap, each dot can be labeled as 1, 2, etc, corresponding to integration variables $\vec{q}_1$, $\vec{q}_2$, and $\vec{q}_3$, and each connecting line corresponds to f_{ij}. For example, the three dots in each diagram of b_3 can be labeled as 1, 2, and 3, in any order. Clearly the first three diagrams are identical. So, the mathematical formula for these three terms are

$$b_1 = \int d^3\vec{q} = V, \tag{21.54}$$

$$b_2 = \int d^3\vec{q}_1 d^3\vec{q}_2 f_{12}, \tag{21.55}$$

$$b_3 = 3\int d^3\vec{q}_1 d^3\vec{q}_2 d^3\vec{q}_3 f_{12} + \int d^3\vec{q}_1 d^3\vec{q}_2 d^3\vec{q}_3 f_{12} f_{23}. \tag{21.56}$$

Note that the last diagram for b_3 is $ƀ_3$, since that is the only diagram that is one-particle irreducible.

21.3.1 Perturbation parameter

In Section 21.2, it was pointed out that N itself is a perturbation parameter. In this section, we see that $B_2 n$ can be taken as a perturbation parameter, since $P\beta \approx n(1 + B_2 n)$. We have evaluated B_2 already for a simplified potential (Eqs. 21.22 and 21.30):

$$B_2 n = -\frac{\bar{b}_2}{2} n = \frac{1 - \beta u_0}{2} \Omega n \qquad (21.57)$$

Note that the expression for $\bar{b}_2$ was obtained assuming high temperature $\beta u_0 \ll 1$, and so we see that the perturbation parameter is Ωn, which is about 1/1000 for a gas, since $\Omega \sim 1/n_{liquid}$. We can see that the perturbation theory will not be good if the density is high or the temperature is low. If the temperature is low, then the above expression cannot be used, and $\bar{b}_2$ must be re-evaluated.

In any case, the information contained in the previous paragraph may be good starting information with which to figure out the answer for the "cluster expansion paradox" as posed on-line at the forum site.

21.4 Van der Waals theory as a mean field theory

Note that the total volume excluded by the hard core repulsion is given by $N(N-1)\Omega/2 \approx N^2\Omega/2$. So, it is $N\Omega/2$ per particle.

Note that from, Eq. 21.2,

$$Z(N, T, V) = \frac{1}{N!\lambda^{3N}} \int d\vec{\mathbf{q}} \exp\left(-\sum_{i<j} \beta U_{ij}\right) \qquad (21.58)$$

A mean field approach can be taken to replace the potential energy with an average potential energy for the attractive part. Then, the integrand is simply a number, and the integral becomes $V(V-\Omega)...(V-(N-1)\Omega)$ due to the excluded volume. By taking the log and approximating (using $\Omega/V \ll 1$) and then re-exponentiating, we get

$$V(V-\Omega)...(V-(N-1)\Omega) \approx (V - N\Omega/2)^N \qquad (21.59)$$

which basically reflects our observation above that the excluded volume is $N\Omega/2$ per particle.

Now, let us look at the mean-field approximation for the potential energy

$$\begin{aligned} \overline{\sum_{i<j} U_{ij}} &= \frac{1}{2} n^2 \int_{\text{attractive part}} d^3\vec{x}_1 d^3\vec{x}_2 U(\vec{x}_1, \vec{x}_2) && (21.60) \\ &= \frac{1}{2} n^2 V \, 4\pi \int_{r_0}^{\infty} dr r^2 \left(-u_0 \frac{r_0^6}{r^6} \right) && (21.61) \end{aligned}$$

using the same approximate model as in Section 21.2. The integral is elementary, and we get

$$\overline{\sum_{i<j} U_{ij}} = -\frac{1}{2} nN \, \Omega u_0. \qquad (21.62)$$

Thus, the partition function is given by

$$Z(N,T,V) = \frac{1}{N! \lambda^{3N}} \left[\left(V - N\frac{\Omega}{2} \right) \exp\left(\frac{n\beta\Omega u_0}{2} \right) \right]^N . \qquad (21.63)$$

We can compare this equation with the partition function for an ideal gas (Eq. 6.44)

$$Z_0(N,T,V) = \frac{1}{N!} \left(\frac{V}{\lambda^3} \right)^N .$$

Clearly, it indicates that the reduction of volume by $N\Omega/2$. Also, note that

$$\begin{aligned} P &= -\frac{\partial F}{\partial V} = k_B T \frac{\partial \log Z}{\partial V} && (21.64) \\ &= \left[\frac{N k_B T}{V - N\frac{\Omega}{2}} - \frac{u_0 \Omega n^2}{2} \right] && (21.65) \\ &= \left[\frac{N k_B T}{V - bN} - an^2 \right]. \qquad a \equiv u_0\Omega/2,\ b \equiv \Omega/2 && (21.66) \\ (P + an^2)(V - bN) &= N k_B T. && (21.67) \end{aligned}$$

So, the van der Waals equation of state can be derived as a mean field theory where the average attraction energy per particle $-n\Omega u_0/2 = -an$ reduces the pressure and the core repulsion reduces the volume by bN.

From the partition function, the Helmholtz free energy can be written down as (using Sterling's formula)

$$\begin{aligned} \frac{F(N,T,V)}{-k_B T} &= N \log\left(V - N\frac{\Omega}{2} \right) + \frac{Nn\beta\Omega u_0}{2} - N\log N + N - 3N\log\lambda && (21.68) \\ &= N \log\left(\frac{V}{N} - \frac{\Omega}{2} \right) + \frac{Nn\beta\Omega u_0}{2} + N(1 - 3\log\lambda). && (21.69) \end{aligned}$$

21.5 Van der Waals theory and phase transition

As derived using the cluster expansion, the van der Waals equation is not expected to be quantitatively valid at all near the gas-liquid phase transition where the temperature is low and the density is high. Nevertheless, this equation is historically the first equation to explain the gas-liquid phase transition, as it has the qualitative features to do so.

Note that, within the grand canonical ensemble, the number N is determined as that value of $\mathcal{N}$ that minimizes the "would-be"[4] grand potential $\mathcal{G}_{\mathcal{N}}(\mu, T, V) = F(\mathcal{N}, T, V) - \mu\mathcal{N}$, a central subject explained in LN 7 and lectures prior to it. Since V is fixed in this minimization, we might as well divide by V. Also, diving by $-k_B T$ as above for F, we get

$$\begin{aligned}\frac{\mathcal{G}_{\mathcal{N}}(\mu, T, V)}{-k_B T V} &= x \log\left(x^{-1} - \frac{\Omega}{2}\right) + \frac{\beta\Omega u_0}{2} x^2 + (1 - 3\log\lambda + \beta\mu)x && (21.70)\\ &= x \log\left(x^{-1} - b\right) + \beta a x^2 + (1 - 3\log\lambda + \beta\mu)x. && (21.71)\end{aligned}$$

Here, $x = \mathcal{N}/V$. Due to the division by a negative number, we would need to maximize this function of x, and $x_{max} = n$, the average particle density.

The physical range of x is from 0 to $1/b$, since b is the excluded volume per particle. As $x \to 1/b$, the above function goes to $-\infty$ and as $x \to 0$, it behaves as $-x \log x$. Thus, we expect that this function to reach a positive value and then drop to $-\infty$ as $x \to 1/b$, ensuring at least one maximum (i.e., a minimum of $\mathcal{G}$), as it naturally should.

It can be further shown that the above function has either (1) one maximum, or (2) two maxima and one minimum. When the latter happens (at low temperature), the two maxima correspond to the density of the gas phase and the density of the liquid phase. These behaviors gave the first prediction of a phase transition between a gas phase and a liquid phase. This is a great discovery of the van der Waals theory! While the theory is not quantitatively correct (most mean field theories aren't!), the theory is able to qualitatively predict (most mean field theories do!) that there *is* a phase transition.

As a function of T, the thermodynamic potential develops a non-analytic behavior—a singularity. Just above T^*, the transition temperature, the grand potential is lower in the gas phase, so the density n_{gas} prevails. Just below T^*, the liquid density n_{liquid} prevails, since the liquid phase is lower in the grand potential. So, the density is discontinuous! Or, by the van der Waals equation of state, the pressure is discontinuous!

---

[4] This is not the grand potential just yet. It becomes the grand potential when we have found $\mathcal{N}$ that minimizes it. So, the subscript $\mathcal{N}$. For this reason, $\mathcal{G}_{\mathcal{N}} \neq -PV$. In other words, μ here is not a function of $\mathcal{N}$ at all, but is a function of N.

Note that $\mathcal{G} = -PV$, and so it means that G is discontinuous. The reason for the singularity is not to be found in the non-analyticity of the expression Eq. 21.71, which is analytical in the physical domain. In fact, since all thermodynamic potentials are nice integrals and they seem to have no reason to be non-analytic. If you see the argument above for the non-analyticity, you realize that the non-analyticity comes from the rule of large numbers. **So, a phase transition is possible only in the thermodynamic limit ($N, V \to \infty$).** We have no problem accepting it—we surely know that phase transitions occur predictably when we boil water, for example, and also know that the water in a pot has a lot of molecules. Here, we find that these two facts are inseparable.

Other typical discussion of the gas-liquid phase transition involves the Maxwell construction of the Gibbs free energy along an isotherm, and the Clapeyron equation at phase coexistence ($dP/dT = \Delta S/\Delta V$; $\Delta S = L/T$ is the entropy change (L being the latent heat), ΔV = the volume change across the transition). These are elementary topics, which can be made using the van der Waals equation of state, and they are usually covered in lower-level courses. So, this topic is left for your reading (and possibly some homework problem).

The gas-liquid transition is a first order transition, as the latent heat is involved. However, as T is raised, the latent heat vanishes just when the gas-liquid phase transition disappears. The T_c, P_c values at which this transition just disappears define the critical point. At the critical point, the transition is a second order transition, since the latent heat vanishes (or more generally, since the first derivative of the Gibbs free energy is continuous).

A more aggressive application of van der Waals equation is to use it to quantitatively describe the behavior near this critical point, and generally to apply it to describe the phase transition quantitatively. However, this naive hope was not met with success. Instead, a proper way of dealing with critical phenomena is the renormalization group, which we will consider in a later part of this course. The description of the critical behaviors and the universality classes associated with them is one of the most important developments in the statistical physics in the latter half of the last century.

21.6 Scattering of wave

It is important that all of what we discussed so far is related to measurements. Of course, the equation of state is something that one can measure. Indeed, measurements and characterizations of the equation of state came first, and then our theoretical understanding of it came next. The measured data of the specific heat of solids gives access to the entropy and other fundamental thermodynamic functions.

Here, we describe another important class of experiments—wave scattering. One should always keep in mind that experimental data carry all information regarding the interactions present in the system. Just because the theory cannot handle all the interactions should not deter a phenomenological understanding which can be developed from a critical set of data and good theoretical ideas. Scattering experiments give very detailed information for many thermodynamic quantities. In particular, within the semi-classical statistical mechanics, a complete measurement of pair density function from a scattering experiment gives a complete access to all equilibrium thermodynamic functions, as shall see in the next section.

The starting point of any scattering discussion can be taken as

$$\psi(\vec{x}) = \exp(i\vec{k}\cdot\vec{x}) + f(k,\theta)\frac{\exp(ikr)}{r} \tag{21.72}$$

Here, we shall confine ourselves to an elastic scattering, so $k = |\vec{k}|$ is used in the second term. In other words, we are considering a diffraction. While this form of wave function should be familiar from the elementary non-relativistic quantum mechanics, this wave function can be viewed as generally applicable for any wave that gets scattered. That is, whether the incoming wave consists of light particles, electrons, or even atoms/molecules, the above formula says that it comes in in a plane wave form (by an experimental setup) and then scatters off with the quantum amplitude $f(k,\theta)$[5] in a form of a spherical wave with angle dependence included by the θ dependence of θ.

In the above expression, the implicit assumption is that the sample by which the wave is scattered is small in comparison to the experimental dimensions. Importantly, the sample to the detector distance (r) is large compared to the sample size so that it makes sense to say that the sample is placed at the origin and acts as a point source, as the above equation implies.

So the above equation satisfactorily describes the experiment. But, here we are interested in the theory of it.

For doing the theory, let us suppose that our sample consists of N fixed scattering centers.[6] We can then apply the above formula to each scattering center. With the

---

[5] You can, and should try to, show that $d\sigma/d\Omega = |f(k,\theta)|^2$.

[6] So, here we are assuming that during the scattering event, the movement of atoms is much less than the wavelength of the probing wave.

origin fixed somewhere inside the sample, the individual scatterer will cause

$$\psi_j(\vec{x}) = \exp(i\vec{k}\cdot\vec{x}) + \exp(i\vec{k}\cdot\vec{x}_j)\, f_j(k,\theta_j)\,\frac{\exp\left(ik|\vec{x}-\vec{x}_j|\right)}{|\vec{x}-\vec{x}_j|} \tag{21.73}$$

In the second term, the $\exp(i\vec{k}\cdot\vec{x}_j)$ term accounts for the extra phase that is accumulated before the light scatters off of the scattering center at $\vec{x}_j$ as opposed to the origin.[7] The angle θ_j means the angle of the outgoing vector $\vec{x}-\vec{x}_j$. Now, let us assume that each scattering center is identical. So, we put $f_m(\theta_j) \equiv f_j(\theta_j)$ with the subscript m is used to mean "molecule."[8] In addition, using the same assumption as described above, we can put $\theta_j \approx \theta$ and $|\vec{x}-\vec{x}_j| \approx r$ where r, θ as defined in Eq. 21.73. Also, note[9] that $\exp(ik|\vec{x}-\vec{x}_j|) \approx \exp(ikr)\exp(-i\vec{k}'\cdot\vec{x}_j)$. Thus, summing up[10] the diffracted part of the wave we get the total scattering amplitude $f(k,\theta)$ of Eq. 21.73 in terms of $f_m(k,\theta)$'s as

$$f(k,\theta) = f_m(k,\theta)\sum_j e^{-i(\vec{k}'-\vec{k})\cdot\vec{x}_j} = f_m(k,\theta)\sum_j e^{-i\vec{q}\cdot\vec{x}_j} \tag{21.74}$$

$$\vec{q} \equiv \vec{k}'-\vec{k}. \tag{21.75}$$

Note that, here, we use $\vec{q}$ as the change of the wave vector[11], not as the generalized coordinate. For a unit incoming flux of the wave, the intensity of the diffracted wave measured by the detector at distance r and angle θ is given by

$$I = \frac{|f(k,\theta)|^2}{r^2} = \frac{|f_m(k,\theta)|^2}{r^2}\sum_{j,l} e^{-i\vec{q}\cdot(\vec{x}_j-\vec{x}_l)} \tag{21.76}$$

Finally, the ensemble average must be taken to take into account of the statistical distribution of microstates

$$I = \frac{|f(k,\theta)|^2}{r^2} = \frac{|f_m(k,\theta)|^2}{r^2}\left\langle\sum_{j,l} e^{-i\vec{q}\cdot(\vec{x}_j-\vec{x}_l)}\right\rangle \tag{21.77}$$

Analogous to the single particle density function (Eq. 5.23), **the two particle density function** is defined as

$$f_2(\vec{p},\vec{x},\vec{p}\,',\vec{x}\,') \equiv \left\langle\sum_{j\neq l}\delta(\vec{p}-\vec{p}_j)\delta(\vec{x}-\vec{x}_j)\delta(\vec{p}\,'-\vec{p}_l)\delta(\vec{x}\,'-\vec{x}_l)\right\rangle \tag{21.78}$$

---

[7] Some more discussions of this problem can be found at this URL (homework 3).

[8] If a molecule consists of an array of atoms, then one can extend this current consideration to see that $f_m = \sum_l \exp(i\vec{k}\cdot(\vec{x}_{a,l,j}-\vec{x}_j))f_{a,l,j}$, where $f_{a,l,j}$ is an *atomic form factors* for the l-th atom, and f_m is the so-called a *structure factor*.

[9] This was shown in class, or, it is left for your exercise. Also, see footnote 7.

[10] You may wonder whether this simple summing up is good enough, worrying that there may be multiple scattering events. Indeed such multiple scattering is known to be important for electron diffraction. For X-ray scattering, it is less of a concern.

[11] This is a very common notation in scattering physics.

Integrating over momentum variables, we get **the two particle distribution function** or **the pair distribution function**:

$$n_2(\vec{x},\vec{x}\,') = \left\langle \sum_{j\neq l} \delta(\vec{x}-\vec{x}_j)\delta(\vec{x}\,'-\vec{x}_l) \right\rangle \tag{21.79}$$

From this, it follows that

$$\left\langle \sum_{j\neq l} e^{-i\vec{q}\cdot(\vec{x}_j-\vec{x}_l)} \right\rangle = \int d^3\vec{x}\, d^3\vec{x}\,'\, e^{-i\vec{q}\cdot(\vec{x}-\vec{x}\,')}\, n_2(\vec{x},\vec{x}\,') \tag{21.80}$$

So, the diffraction intensity is proportional to the Fourier transform of the two particle distribution function. Note that the two particle distribution function is of order N^2. So, we can drop the diagonal sum $(j = l)$ in Eq. 21.76, which gives N. Assuming the translation invariance[12],

$$n_2(\vec{x},\vec{x}\,') = n_2(\vec{x}-\vec{x}\,'). \tag{21.81}$$

So, the diffraction intensity is finally expressed as

$$I(k,r,\theta) = \frac{|f_m(k,\theta)|^2\, V}{r^2}\, \widetilde{n}_2(\vec{q}) \tag{21.82}$$

where, to recapitulate, k is the magnitude of the wave number of the probe wave, r,θ define the position of the outgoing wave detector relative to the sample, V is the sample volume, and $\widetilde{n}_2(\vec{q})$ is the Fourier transform of $n_2(\vec{x},\vec{x}\,')$ with respect to $\vec{x}-\vec{x}\,'$.

So, assuming that one has access to the f_m quantity, it is possible to measure directly the two particle density function just by doing a scattering experiment!

21.7 Thermodynamic functions from pair distribution

In semi-classical statistical mechanics, if $n_2(\vec{x},\vec{x}\,')$ is known completely, then all thermodynamic functions can be derived. Here, we will show that E can be derived. In a homework problem, you will show that P can de derived from $n_2(\vec{x},\vec{x}\,')$. Recall from the discussion of LN 7, we have

$$E = \left.\frac{\partial \log Z}{\partial(-\beta)}\right|_{N,V}.$$

---

[12] This should hold in normal situations. However, if considering a system of liquid residing in a crystalline environment, this assumption will be broken. In that case, the crystal translation invariance must be incorporated instead.

The partition function is given by, in the presence of the interaction (from Eq. 21.2),

$$\log Z = -3N \log \lambda(T) - \log N! + \log \left[\int d\vec{\mathbf{q}} \exp(-\beta H_{int}) \right] \tag{21.83}$$

where $H_{int} = \sum_{i<j} U_{ij}$. The derivative of the $\lambda(T)$-containing term gives the equipartition energy for the kinetic energy, which is $\frac{3}{2} N k_B T$. Effecting the differentiation of the last term, we get

$$\begin{aligned}
E &= \frac{3}{2} N k_B T + \frac{\int d\vec{\mathbf{q}}\, H_{int} \exp(-\beta H_{int})}{\int d\vec{\mathbf{q}} \exp(-\beta H_{int})} && (21.84)\\
&= \frac{3}{2} N k_B T + \frac{1}{2} \sum_{i \neq j} \frac{\int d\vec{\mathbf{q}}\, U_{ij} \exp(-\beta H_{int})}{\int d\vec{\mathbf{q}} \exp(-\beta H_{int})} && (21.85)\\
&= \frac{3}{2} N k_B T + \frac{1}{2} \sum_{i \neq j} \frac{\int d\vec{\mathbf{q}}\, U(\vec{q}_i, \vec{q}_j) \exp(-\beta H_{int})}{\int d\vec{\mathbf{q}} \exp(-\beta H_{int})} && (21.86)\\
&= \frac{3}{2} N k_B T + \frac{1}{2} \sum_{i \neq j} \int d^3\vec{x} d^3\vec{x}' U(\vec{x}, \vec{x}') \frac{\int d\vec{\mathbf{q}}\, \delta(\vec{x} - \vec{q}_i) \delta(\vec{x}' - \vec{q}_j) \exp(-\beta H_{int})}{\int d\vec{\mathbf{q}} \exp(-\beta H_{int})} && (21.87)\\
&= \frac{3}{2} N k_B T + \frac{1}{2} \int d^3\vec{x} d^3\vec{x}' U(\vec{x}, \vec{x}') \sum_{i \neq j} \frac{\int d\vec{\mathbf{q}}\, \delta(\vec{x} - \vec{q}_i) \delta(\vec{x}' - \vec{q}_j) \exp(-\beta H_{int})}{\int d\vec{\mathbf{q}} \exp(-\beta H_{int})} && (21.88)\\
&= \frac{3}{2} N k_B T + \frac{1}{2} \int d^3\vec{x} d^3\vec{x}' U(\vec{x}, \vec{x}') n_2(\vec{x}, \vec{x}') && (21.89)
\end{aligned}$$

where, in the last step, we simply used the fact that the pair distribution function is the ensemble average of the two particle density in the real space, $\sum_{i \neq j} \delta(\vec{x} - \vec{q}_i) \delta(\vec{x}' - \vec{q}_j)$. Thus, assuming that we know the form of the two particle potential, $U(\vec{x}, \vec{x}')$, we can simply derive the energy from the two particle distribution function!