Variational approach in statistical mechanics

(Mean field treatment)

Thermodynamics

Recall "thermodynamic variational principle": for fixed \( T, V \), the free energy \( F \) attains its minimal value.

In the statmech, we are simply calculating \( F \)

\[
F = -k_B T \ln Z , \quad Z = Tr e^{-\beta H}
\]

⇒ unique function of \( T, V \); there is no sense to speak about different \( F \)'s at given \( T, V \), since we are summing over all states (all accessible configurations).

Ultimately, we just want to know \( F(T, V) \) as accurately as possible.

Variational approach allows us to strictly bound the exact \( F(T, V) \) from above. If we try to come up with as stringent bound as possible, we can hope to capture the important properties of the system.

Interacting system - \( Z = Tr e^{-\beta H} \) is very difficult to calculate. \( F = -k_B T \ln Z \)

Suppose we have \( H_0 \) defined on the same degrees of freedom (on the same \( p_i, q_i \) in classical statmech or the same Hilbert space in quantum statmech), and suppose we can easily calculate with \( H_0 \)

\[
Z_0 = Tr e^{-\beta H_0} ; \quad F_0 = -k_B T \ln Z_0 .
\]

Then we can provide an upper bound:
\[ F \leq F_0 + \langle H \rangle_0 - \langle H_0 \rangle_0 \]

\[ \langle H \rangle_0 = \text{Tr} \left( H e^{-\beta H_0} \right) / Z_0, \quad \langle H_0 \rangle_0 = \text{Tr} \left( H_0 e^{-\beta H_0} \right) / Z_0 \]

**Proof:** Classical stat mech (so as not to worry about commutators)

\[ Z = \text{Tr} e^{-\beta H} = \text{Tr} \left( e^{-\beta (H-H_0)} e^{-\beta H_0} \right) = Z_0 \text{Tr} \left( e^{-\beta (H-H_0)} e^{-\beta H_0} / Z_0 \right) \]

\[ \ln Z = \ln Z_0 + \ln \left( \sum_{\text{states } S} e^{-\beta (H_0(S)-H(S))} p_S^{(0)} \right) \]

Use the following inequality:

\[ \ln \left( \sum_i x_i p_i \right) \geq \sum_i p_i \ln x_i \quad \text{if } \sum_i p_i = 1 \]

(Actually, "ln" can be replaced with any concave function $f$:)

\[ f(\sum_i x_i p_i) \geq \sum_i f(x_i p_i) \]

with geometric proof:

- $p_i$ - "mass" s.t. total mass = 1
- Center of mass is at $\sum_i p_i (x_i, s_i)$
- Must lie below the curve
- \[ f(x_{cm}) \geq f(s_{cm}) = \sum_i f(x_i p_i) \]

\[ \ln Z \geq \ln Z_0 + \sum_S p_S^{(0)} (-\beta) (H(S)-H_0(S)) \]

\[ F = -\frac{1}{\beta} \ln Z \leq F_0 + \langle H \rangle_0 - \langle H_0 \rangle_0 = F_{\text{trial}} \]
Summary of the derivation of variational mf

\[ Z = \text{Tr} e^{-\beta H} = \text{Tr} \left( e^{-\beta (H + \frac{1}{\beta} \ln g_0)} g_0 \right) \]

\[ \ln Z = \ln \langle e^{-\beta (H + \frac{1}{\beta} \ln g_0)} \rangle \geq \left\langle -\beta (H + \frac{1}{\beta} \ln g_0) \right\rangle \]

\[ F = -\frac{1}{\beta} \ln Z \leq \langle H \rangle_0 + \frac{1}{\beta} \left\langle \ln g_0 \right\rangle_0 \equiv F_{\text{trial}} \]

\[ F_{\text{trial}} = \langle H \rangle_0 - T \cdot \left[ -k_B \text{Tr} (g_0 \ln g_0) \right] \]

(Recall "entropy of mixing": example: binary degree of freedom which can be in state |A with prob \( p_A \)
|B with prob \( p_B \)
\[ \Rightarrow \text{entropy} = -p_A \ln p_A - p_B \ln p_B \])

Choice of \( g_0 \) is often guided by trial or meanfield \( H_0 \)

\[ g_0 = \frac{e^{-\beta H_0}}{Z_0} \Rightarrow \left\langle \ln g_0 \right\rangle = -\beta \langle H_0 \rangle_0 - \ln Z_0 \]

\[ F_{\text{trial}} = \langle H \rangle_0 - \langle H_0 \rangle_0 + F_0 \]
Remarks: In the $T \to 0$ limit, $F = U - TS \to U$, and this reduces to $E_{\text{exact}} \leq \langle H \rangle_0$ - familiar variational principle in quantum mechanics.

- Choice of $H_0 \leftrightarrow$ choice of $g_0 = e^{-\beta H_0} / Z_0$ reflects our idea about the phase of the system. $H_0$ may have variational parameters, which we can vary to minimize $F_{\text{trial}} = F_0 + \langle H \rangle_0 - \langle H_0 \rangle_0$, in the hope that a tighter bound corresponds to a more accurate description.

- Naively, if $H = H_0 + U$ (interactions), $F_{\text{trial}} = F_0 + \langle U \rangle_0$ - similar to the first term in the cumulant expansion. However, in the perturbative treatment we must use $\langle \ldots \rangle_0$ with non-interacting gas, while in the variational (mean field) treatment we can choose non-trivial $\langle \ldots \rangle_0$ inspired by the expected physics - e.g. phase-separated state as in the discussion of vDW gas below.

\[ g_0 = e^{-\beta H_0} / Z_0 \rightarrow H_0 = -1 / \beta \ln(Z_0 g_0) \]

\[ F_0 - \langle H_0 \rangle_0 = -1 / \beta \ln Z_0 - \text{Tr} g_0 \cdot \left[ -1 / \beta \times (\ln Z_0 + \ln g_0) \right] = \]

\[ = +1 / \beta \text{Tr} g_0 \ln g_0 = +T \cdot k_B \text{TR} \ln g_0 = +k_B T \langle \ln g_0 \rangle \]

\[ F_{\text{trial}} = \langle H \rangle_0 + T \cdot k_B \text{TR} \ln g_0 \]

(The term $-k_B T \langle g_0 \ln g_0 \rangle$ is "entropy of mixing".)
Mean field description of the liquid-gas transition

Work in the canonical ensemble to parallel our discussion of the free energy of the van der Waals gas.

\[ H = H_{\text{kin}} + \text{hard-core } + (\text{weak}) \text{long-range repulsion } + \text{attraction} \]

(we need attraction to cause transition to more dense liquid phase and repulsion to prevent collapse)

Consider uniform density "trial states"

\[ H_0 = H_{\text{kin}} + \text{hard-core repulsion} \]

\[ \{ \text{there are no variational parameters here - these will appear shortly when we take phase-separated states} \}. \]

For simplicity, assume that \( F_0(T,s) = F_{\text{ideal}}(T,V,N_b) \)

\[ F_0(T,V) = F_{\text{ideal}}(T,V,N_b) = -Nk_bT \ln(V - N_b) + \]

\[ + NK_bT \left\{ \ln N + \ln \frac{V^3}{T} - 1 \right\} \]

\[ F_{\text{trial}}(T,V) = F_0(T,V) + \left\langle \text{long-range attraction} \right\rangle. \]

\[ \left\langle U(q_i,q_j) \right\rangle = \sum_{i<j} \left\langle V(q_i - q_j) \right\rangle = \frac{N(N-1)}{2} \int_0^a \frac{d^3q_1}{V} \frac{d^3q_2}{V} \]

\[ = \frac{N(N-1)}{2V} \int_0^a d^3q \partial^2 \rho(q) \equiv \frac{\partial}{\partial V} \frac{\partial}{\partial a} \frac{-N^2}{V} \quad \text{attractive} \]
\[ F_{\text{trial}} = -\frac{N^2}{\nu} a - N k_B T \ln (V - Nb) + N k_B T \frac{3}{2} \ln N + \ln \chi_n^3 - 13 \]

This is precisely the free energy of van der Waals gas that we obtained by integrating van der Waals equation of state. Alternatively, we can derive equation of state from \( F_{\text{trial}} \):

\[
P = -\frac{\partial F_{\text{trial}}}{\partial V} = \frac{N k_B T}{V - Nb} - \frac{N^2}{V^2} a \]

\( \text{vdW equation!} \)

We can now interpret what we did after deriving \( F_{\text{vdW}} \) as follows.

- For \( T > T_c \), the above trial state gives sensible results, with \( \frac{\partial^2 F_{\text{trial}}}{\partial V^2} > 0 \) everywhere.

- For \( T < T_c \), upon seeing

we realize that we can actually do better in this region by taking phase-separated trial state.

Calculations with such phase-separated states are precisely like the calculations we did with \( F_{\text{vdW}} \):

Trial state

| \( x_1 \cdot N \) | particles in \((V_1, F_1)\) state |
| \( x_2 \cdot N \) | \( x_2 \cdot N = (1 - x_1) \cdot N \) particles in \((V_2, F_2)\) state |
| \( x_1 \cdot N \) |

\[ x_1 \cdot V_1 + x_2 \cdot V_2 = V = x_1 \cdot V_1 + (1 - x_1) \cdot V_2 \]
Variational parameters are $V_1$ and $V_2$ (which states to phase-separate to) assuming $N$ fixed, $\frac{V_1 + V_2}{N} = \text{vol per particle}$, etc.

$$x_1 = \frac{V-V_2}{V_1-V_2} \quad \text{unique} \quad x_2 = 1-x_1$$

$$F_{\text{trial}} = x_1 F_1 + x_2 F_2 = x_1 F_1 + (1-x_1) F_2$$
$$V = x_1 V_1 + x_2 V_2 = x_1 V_1 + (1-x_1) V_2$$

Given $F_{\text{trial}}(V)$ curve constructed using only uniform trial states,

we can construct more trial energies using phase-separated states $F_1 \leftarrow \text{"phase-separated trial"}$ $F_2$

Minimizing over all possible such constructions $\rightarrow \text{"enveloping curve"}$

Once we have such $F_{\text{trial}}$ satisfying $\frac{\partial F}{\partial V^2} \geq 0$ we cannot do any better with any other attempt at phase separation $\sum x_i N \cdot \psi_i = \sum x_i V_i$, $\sum x_i = 1$

$F_{\text{no}} = \frac{1}{N} \sum x_i N \cdot \phi_i = \frac{1}{N} \sum x_i F_i \geq F(V)$ by convexity.
Kardar's text

Variational approach in the grand canonical ensemble

\[ H = H_{\text{kin}} + (\text{hard-core repulsion}) + (\text{long-distance attraction}) - \mu N \]

\[ \mathcal{Z} = \sum_{N} \frac{\mathcal{Q}^{\text{HN}}}{N!} \left| N \right> \left< N \right| = \text{Tr} e^{-\beta H} \]

\[ \mathcal{Z}(T, V, \mu) = -k_B T \ln \mathcal{Z} \]

Variational approach: \[ \mathcal{Z} \leq \mathcal{Z}_0 + \langle H - H_0 \rangle. \]

\[ H_0 = H_{\text{kin}} + (\text{hard-core repulsion}) - \mu_{\text{var}} N \]

\[ \mu_{\text{var}} \quad \text{\textit{variational parameter.}} \]

Recall calculations for the ideal gas in the absence of hard core interactions

\[ \mathcal{Z}_{\text{ideal gas}} = -k_B T \frac{V e^{\beta \mu_{\text{ideal gas}}}}{\lambda^3} \]

\[ \langle N \rangle_{\text{ideal gas}} = \frac{e^{\beta \mu_{\text{ideal gas}}}}{\lambda^3} \frac{V}{\lambda^3} = -\langle N \rangle k_B T \]

Treat the hard-core repulsion approximately by

\[ V \rightarrow V - \langle N \rangle b \]

\[ \Rightarrow \langle N \rangle = \frac{e^{\beta \mu_{\text{var}}}}{\lambda^3} \left( V - \langle N \rangle b \right) \]

\[ n = \frac{e^{\beta \mu_{\text{var}}}}{\lambda^3} (1 - nb) \Rightarrow n = \frac{x}{1 + xb} \]

\[ n = n(\mu_{\text{var}}) \leftrightarrow \mu_{\text{var}} = \mu_{\text{var}}(n) \quad \text{can view either \mu_{\text{var}} or \mu_{\text{var}}(n) as variational param.} \]
\[ \mathcal{S}_0 = -\langle N \rangle k_B T = -nV \cdot k_B T \quad \text{long-range attraction} \]

\[ \mathcal{S}_{\text{trial}} = \mathcal{S}_0 + \langle H - H_0 \rangle_0 = \mathcal{S}_0 + \langle U \rangle_0 - (\mu - \mu_{\text{var}}) \langle N \rangle \]

\[ = -nk_B T - \frac{\langle N \rangle^2}{V} a - (\mu - \mu_{\text{var}}) \langle N \rangle, \]

\[ \mathcal{S}_{\text{trial}} \frac{1}{V} = -nk_B T - an^2 - (\mu - \frac{1}{\beta} \ln \frac{n\lambda^3}{1-nb}) n = \]

\[ = -k_B T \left\{ n + \beta a n^2 + \beta \mu n - n \ln \frac{\lambda^3}{1-n-b} \right\} \]

\[ \text{same as Kardar's (5.77)} \]

\[ \text{with } a = \frac{1}{2} a, \quad b = \frac{\mu}{2} \]

\[ \mathcal{S}_{\text{trial}}(T, V, \mu) = \min_n \left\{ -k_B T \left\{ \ldots \right\} \right\} \]

\[ \text{Equation for } n : \quad 1 + \beta a \cdot 2n + \beta \mu - \ln \left( \frac{\lambda^3}{1-n-b} \right) + \]

\[ + n \frac{1}{n - b} \cdot \frac{-1}{n^2} = 0. \]

\[ n = n(\mu) \]

\[ \mu = -2an + k_B T \left( \ln \frac{\lambda^3}{1-nb} + \frac{1}{1-nb} - 1 \right) \]

\[ = \text{math. equivalent to earlier treatment in the canonical ensemble!} \]
Indeed, compare with earlier treatment when discussing Maxwell construction.

\[
G = F_{vdW} + pV = -\frac{N^2}{V}a - Nk_BT \ln (V-Nb) + \\
+ Nk_BT \left[ \ln N + \ln \lambda^3 - 1 \right] + \left( \frac{Nk_BT}{V-Nb} - \frac{N^2}{V^2}a \right) V
\]

\[
= -\frac{2}{V} N^2 a + Nk_BT \left[ \ln \frac{N\lambda^3}{V-Nb} - 1 \right] + \frac{Nk_BT}{V-Nb} V
\]

\[
\mu = \frac{G}{N} = -2aN + k_BT \left[ \ln \frac{n\lambda^3}{1-nb} - 1 \right] + \frac{k_BT}{1-nb}
\]

\[\uparrow n\]

\begin{itemize}
  \item Liquid
  \item Discontinuous change in density as we vary \(\mu\)
\end{itemize}

\[\text{gas} \quad \mu_{crit} \quad \mu \text{critical} \quad \mu < \mu_{crit} \quad n = n_{gas}(\mu) \quad \text{all gas}
\]

\[\text{liquid} \quad \mu > \mu_{crit} \quad n = n_{liq}(\mu) \quad \text{all liquid}
\]

\[\mu = \mu_{crit} \quad \text{critical point}
\]

\[\text{1st-order transition}\]