

Overview of Thermodynamics and Statistical Mechanics

Consider an isolated system with total energy E , total number of particles N and volume V . Isolated means that the system cannot exchange energy or particles with the environment.

The central assumption in Stat Mech is ergodicity, which means that, classically, the system explores all parts of the allowed phase space equally. Quantum mechanically, each microscopic state consistent with E, V, N has equal probability. Let the number of such microscopic states be

$$\Omega(E, V, N) \quad (1)$$

The entropy is defined as

$$S(E, V, N) = k \ln(\Omega) \quad (2)$$

k is Boltzmann's constant.

S is clearly extensive. Imagine putting together two identical systems with E, V, N , each to make a single system with $2E, 2V, 2N$. The total number of microstates is

$$(\Omega(E, V, N))^2 = \Omega(2E, 2V, 2N) \quad (3)$$

and thus

$$S(2E, 2V, 2N) = 2S(E, V, N) \quad (4)$$

More generally S is a homogeneous function of E, V, N

$$S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N) \quad (5)$$

Setting $\lambda = 1 + \epsilon$ and taking the $\epsilon \rightarrow 0^+$ limit one obtains

$$S(E + \epsilon E, V + \epsilon V, N + \epsilon N) = S(E, V, N) + \epsilon S(E, V, N) \quad (6)$$

expand the LHS

$$\begin{aligned} S(E, V, N) + \epsilon \left[E \left(\frac{\partial S}{\partial E} \right)_{V, N} + V \left(\frac{\partial S}{\partial V} \right)_{E, N} + N \left(\frac{\partial S}{\partial N} \right)_{E, V} \right] \\ = S(E, V, N) + \epsilon S(E, V, N) \end{aligned}$$

or

$$S(E, V, N) = E \left(\frac{\partial S}{\partial E} \right)_{V, N} + V \left(\frac{\partial S}{\partial V} \right)_{E, N} + N \left(\frac{\partial S}{\partial N} \right)_{E, V} \quad (7)$$

One defines

$$\left(\frac{\partial S}{\partial E} \right)_{V, N} = \frac{1}{T} \quad (8)$$

$$\left(\frac{\partial S}{\partial V} \right)_{E, N} = \frac{P}{T} \quad (9)$$

and

$$\left(\frac{\partial S}{\partial N} \right)_{E, V} = -\frac{\mu}{T} \quad (10)$$

where, as we will soon see T is the temperature, P the pressure and μ the chemical potential.

Thus

$$S = \frac{E}{T} + \frac{PV}{T} - \frac{\mu N}{T} \quad (11)$$

or

$$E = TS - PV + \mu N \quad (12)$$

Take a total derivative of $S(E, V, N)$

$$dS = dE \left(\frac{\partial S}{\partial E} \right)_{V, N} + dV \left(\frac{\partial S}{\partial V} \right)_{E, N} + dN \left(\frac{\partial S}{\partial N} \right)_{E, V}$$

$$dS = \frac{dE}{T} + \frac{PdV}{T} - \frac{\mu dN}{T} \quad (13)$$

$$dE = TdS - PdV + \mu dN \quad (14)$$

Consistency between (12), (14) leads to the Gibbs-Duhem relation

$$SdT - VdP + Nd\mu = 0 \quad (15)$$

Now consider a system in thermal contact with a reservoir. The total energy is $E_R + E$, the total volume $V_R + V$, and the total number of particles $N_R + N$

V and N are assumed to be constant but E and E_R can change while keeping $E + E_R = E_{\text{tot}}$ fixed.

Now comes the II Law of Thermodynamics: In equilibrium the entropy of an isolated system is maximized.

In Statistical Mechanical terms, this is natural, because if one starts the ensemble in a small region of the allowed phase space, over time it samples the entirety of the allowed phase space (ergodicity)

So we want

$$S(E, V, N) + S_R(E_{\text{tot}} - E, V_R, N_R) \quad (17)$$

to be a maximum. But V, V_R, N, N_R are fixed

$$\Rightarrow \frac{\partial S}{\partial E}(\bar{E}, V, N) - \frac{\partial S_R}{\partial E_R}(E_{\text{tot}} - \bar{E}, V_R, N_R) = 0 \quad (18)$$

Where \bar{E} is the value of E where this occurs.

$$\frac{1}{T} = \frac{1}{T_R} \quad (19)$$

which tells us that the temperature of the system matches that of the reservoir in equilibrium.

Let us take this a bit further. Since the reservoir is large

$$E \ll E_R \quad (20)$$

$$S_R(E_{\text{tot}} - E, V_R, N_R) = S_R(E_{\text{tot}}, V_R, N_R) - E \frac{\partial S_R(E_{\text{tot}}, V_R, N_R)}{\partial E_R} + \frac{1}{2} E^2 \frac{\partial^2 S_R(E_{\text{tot}}, V_R, N_R)}{\partial E_R^2} + \dots \quad (21)$$

The 1st term is a constant (independent of E)

$$\frac{\partial S_R}{\partial E_R} = \frac{1}{T_R} \quad (22)$$

The third term is small. To see this note that S_R is extensive (proportional to E_R)

$$\text{so } \frac{\partial^2 S_R}{\partial E_R^2} \sim \frac{1}{E_R} \quad (23)$$

$$\text{The 3rd term is } \sim E \left(\frac{E}{E_R} \right) \ll \frac{E}{T_R} \quad (24)$$

Similar logic allows us to ignore all higher order terms

$$S_{\text{tot}} = S - \frac{E}{T_R} + \text{constant} \quad (25)$$

\Rightarrow maximizing S_{tot} at constant T_R
means

minimizing $F = E - T_R S$

$F = E - TS$ = Helmholtz Free energy. (26)

Here T is a control parameter which is controlled externally (27)

$$dF = dE - TdS - SdT = -SdT - PdV + \mu dN$$

F is a function of T, V, N

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \quad P = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \quad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \quad (28)$$

One can go further. If V, V_R can also change keeping $V_{\text{tot}} = V + V_R$ fixed (think of a piston and cylinder) then one must maximize

$$S - \frac{E}{T} - \frac{PV}{T} \Rightarrow \text{minimize} \quad (29)$$

$G = E - TS + PV$ = Gibbs Free Energy.

Here P and T are external parameters while E, V fluctuate.

$$dG = dE - TdS - SdT + PdV + VdP = -SdT + VdP + \mu dN \quad (30)$$

$$G = G(T, P, N)$$

$$S = -\left(\frac{\partial G}{\partial T}\right)_{P, N} \quad V = \left(\frac{\partial G}{\partial P}\right)_{T, N} \quad \mu = \left(\frac{\partial G}{\partial N}\right)_{T, P} \quad (31)$$

Finally, if particles and energy can be exchanged between the system and reservoir but V is held fixed one maximizes

$$S - \frac{E}{T} + \frac{\mu N}{T} \Rightarrow \text{minimize}$$

$$\Omega = E - TS - \mu N = \text{Grand potential.} \quad (32)$$

$$d\Omega = -SdT - PdV - Nd\mu \quad (33)$$

$$\Rightarrow \Omega = \Omega(T, V, \mu) \quad (34)$$

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

From the homogeneity relation (12) $E = TS - PV + \mu N$ we also know

$$F = -PV + \mu N \quad (35)$$

$$G = \mu N \quad (36)$$

$$\Omega = -PV \quad (37)$$

Now let us find out how to compute F or Ω .

Consider a particular microstate α of the system with energy E_α . Since the total energy is fixed, the reservoir must have energy

$$E_R = E_{\text{tot}} - E_\alpha$$

The probability of the system being in α is the same as the probability of the reservoir having $E_R = E_{\text{tot}} - E_\alpha$, and is proportional to the number of microstates of the reservoir with energy $E_{\text{tot}} - E_\alpha$

$$P_\alpha = \text{constant} e^{k \Omega_R(E_{\text{tot}} - E_\alpha)}$$

$$P_\alpha = \text{constant} e^{-\frac{E_\alpha}{kT}} \quad (38)$$

This is the **Gibbs distribution**. To normalize we divide by the partition f^n

$$Z = \sum_\alpha e^{-E_\alpha/kT} \quad (39)$$

The average energy is

$$\langle E \rangle = \frac{1}{Z} \sum_\alpha E_\alpha e^{-E_\alpha/kT} = kT^2 \frac{\partial}{\partial T} \ln Z \quad (40)$$

On the other hand, from (26), (28)

$$F = E - TS \quad ; \quad S = - \left(\frac{\partial F}{\partial T} \right)_{V, N}$$

$$E = F + TS = F - T \frac{\partial F}{\partial T} = -T^2 \frac{\partial}{\partial T} \left(\frac{F}{T} \right) \quad (41)$$

Comparing (40), (41) we conclude

$$F = -kT \ln Z, \quad (42)$$

This is the key, because it allows us to express a thermodynamic function as a sum over microscopic states.

Similarly, one can express the grand potential via the grand canonical partition function

$$\mathcal{Z}_{GC} = \sum_{\alpha} e^{-\frac{1}{kT} (E_{\alpha} - \mu N_{\alpha})} \quad (43)$$

where the sum is now over all possible numbers of particles in the system as well.

$$\Omega = -kT \ln \mathcal{Z}_{GC} \quad (44)$$

As an example of a simple many-body interacting system consider the 1D Ising Model. This model has "spins" $= \pm 1$ at each site. One can imagine this as arising microscopically from a crystal with spin- $\frac{1}{2}$ electrons localized at each site, with spin-orbit coupling producing a $-D S_z^2$ term leading to Ising anisotropy.

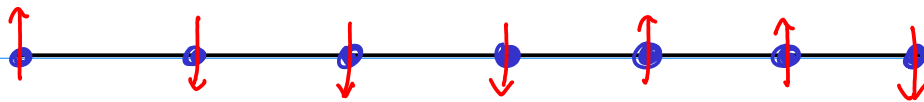
So, let $S_i = \text{spin at site } i = \pm 1$

There is a ferromagnetic interaction between nearest neighbors

$$E(\{S_i\}) = -J \sum_i S_i S_{i+1}$$

(45)

Assume periodic b.c on an N-site ring



$$S_i S_{i+1} = \quad -1 \quad 1 \quad 1 \quad -1 \quad 1 \quad -1$$

The partition f^n is

$$Z = \sum_{\{S_i = \pm 1\}} e^{\frac{J}{kT} \sum_j S_j S_{j+1}}$$

(46)

one is supposed to sum over all possible configurations of S_i . Usually one adds a term coupling each spin to an external field (the "source" term)

$$E(\{S_i\}) = -J \sum_i S_i S_{i+1} - \sum_i h S_i \quad (47)$$

To compute Z , note that this can be written as a product of transfer matrices,

$$\Pi_{ss'} = e^{\frac{J}{T} ss' + \frac{h}{T} s} \quad (48)$$

$$Z = \Pi_{s_1 s_2} \Pi_{s_2 s_3} \Pi_{s_3 s_4} \cdots \Pi_{s_N s_1} = \text{Tr}(\Pi^N) \quad (49)$$

Periodic b.c

Define

$$\tilde{J} = J/T \quad (50)$$

$$\tilde{h} = h/T \quad (51)$$

implicit sum over repeated indices.

$$\Pi = \begin{bmatrix} e^{\tilde{J} + \tilde{h}} & e^{-\tilde{J} + \tilde{h}} \\ e^{-\tilde{J} - \tilde{h}} & e^{\tilde{J} - \tilde{h}} \end{bmatrix}$$

$$\Pi = e^{\tilde{J}} \begin{bmatrix} e^{\tilde{h}} & e^{-2\tilde{J} + \tilde{h}} \\ e^{-2\tilde{J} - \tilde{h}} & e^{-\tilde{h}} \end{bmatrix} \quad (52)$$

$$\text{Now } \det \Pi = e^{2\tilde{J}} - e^{-2\tilde{J}} = 2\text{Sh}(2\tilde{J}) > 0$$

$$\text{Tr } \Pi = 2e^{\tilde{J}} \text{ch}(\tilde{h}) > 0 \quad (53)$$

Both eigenvalues must be positive as long as $\tilde{J} > 0$

The eigenvalue eqⁿ is

$$\lambda^2 - \lambda \text{Tr} \Pi + \det \Pi = 0$$

$$\lambda_{\pm} = \frac{\text{Tr} \Pi}{2} \pm \sqrt{\frac{(\text{Tr} \Pi)^2}{4} - \det \Pi}$$

$$= e^{\tilde{J}} \text{ch} \tilde{h} \pm \sqrt{e^{2\tilde{J}} \text{ch}^2 \tilde{h} - e^{2\tilde{J}} + e^{-2\tilde{J}}}$$

$$\lambda_{\pm} = e^{\tilde{J}} \text{ch} \tilde{h} \pm \sqrt{e^{2\tilde{J}} \text{sh}^2 \tilde{h} + e^{-2\tilde{J}}} \quad (55)$$

Now express Π in its spectral representation

If

$$\Pi |+\rangle = \lambda_+ |+\rangle \quad (56)$$

$$\Pi |-\rangle = \lambda_- |-\rangle \quad (57)$$

$$\Pi = |+\rangle \lambda_+ \langle +| + |-\rangle \lambda_- \langle -| \quad (58)$$

$$\Pi^n = |+\rangle \lambda_+^n \langle +| + |-\rangle \lambda_-^n \langle -| \quad (59)$$

$$\mathcal{Z} = \text{Tr} \Pi^N = \lambda_+^N + \lambda_-^N = \lambda_+^N \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right] \quad (60)$$

The Free energy is

$$F = -kT \ln \mathcal{Z} = -N kT \ln \lambda_+ \quad (61)$$

$$- kT \ln \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right]$$

Let us now take the **thermodynamic limit** $N \rightarrow \infty$. The second term vanishes

because $\frac{\lambda_-}{\lambda_+} < 1$ and $\lim_{N \rightarrow \infty} \left(\frac{\lambda_-}{\lambda_+}\right)^N = 0$

$$\Rightarrow \boxed{F = -NkT \ln(\lambda_+)} \quad (62)$$

What is the average value of S_i ?
Clearly it cannot depend on the site because of translation invariance

$$\langle S_i \rangle = \frac{1}{N} \left\langle \sum_i S_i \right\rangle = \frac{1}{N} \sum_{\{S_i\}} T \frac{\partial}{\partial h} e^{\frac{1}{kT} E(\{S_j\})}$$

$$= \frac{T}{N} \frac{\partial}{\partial h} \ln Z = T \frac{\partial}{\partial h} \ln(\lambda_+)$$

$$\boxed{\langle S_i \rangle = \frac{T}{\lambda_+} \frac{\partial \lambda_+}{\partial h} = \frac{1}{\lambda_+} \frac{\partial \lambda_+}{\partial \tilde{h}}} \quad (63)$$

Since λ_+ is an even, analytic, f^h of \tilde{h} , for small \tilde{h} it should behave like

$$\lambda_+(\tilde{h}) = \lambda_+(0) + \frac{\tilde{h}^2}{2} \frac{\partial^2 \lambda_+(0)}{\partial \tilde{h}^2} + \dots$$

\Rightarrow the magnetization per site m

$$\boxed{m = \langle S_i \rangle = \frac{\tilde{h}}{\lambda_+} \frac{\partial^2 \lambda_+(0)}{\partial \tilde{h}^2}} \quad (64)$$

For any paramagnet (unmagnetized at $\tilde{h}=0$) this is generically true, and the constant of proportionality is the magnetic susceptibility χ

$$m = \chi h \quad (65) \quad \text{for } h \rightarrow 0$$

Now consider the correlations between spins at sites i and j . The correlation function is defined as

$$g_{ij} = \langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle \quad (66)$$

and captures the fact that spins close together are more likely to point in the same direction. (67)

$$g_{ij} = \frac{1}{Z} \text{Tr} \left\{ \prod_{s_1 s_2} \prod_{s_2 s_3} \dots \prod_{s_{i-1} s_i} S_i \prod_{s_i s_{i+1}} \dots \prod_{s_{j-1} s_j} S_j \prod_{s_j s_{j+1}} \dots \right\}$$

S_i and S_j can be replaced by the matrix

$$\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (68)$$

$$g_{ij} = \frac{1}{Z} \text{Tr} \left\{ \prod_{s_1 s_2} \dots \prod_{s_{i-1} s_i} \sigma_3 \prod_{s_i s_{i+1}} \dots \prod_{s_{j-1} s_j} \sigma_3 \prod_{s_j s_{j+1}} \dots \right\}$$

The trace is the same in any basis so let us go to the eigenbasis of \prod

$$\sigma_3 = s_{++} |+\rangle\langle +| + s_{--} |-\rangle\langle -| + s_{+-} |+\rangle\langle -| + s_{-+} |-\rangle\langle +| \quad (69)$$

Claim $s_{++} = m$. To see this take $\langle S_i \rangle$ in this notation. (70)

$$\langle S_i \rangle = \frac{\lambda_+^N s_{++} + \lambda_-^N s_{--}}{\lambda_+^N \left(1 + \left(\frac{\lambda_-}{\lambda_+}\right)^N\right)} \rightarrow s_{++} \quad \text{as } N \rightarrow \infty \quad (71)$$

$\left(\frac{\lambda_-}{\lambda_+}\right)^N \rightarrow 0$

Now we let $N \rightarrow \infty$ while keeping the lattice distance $j-i$ fixed. Then for

$l < i$ and $l > j$ we need keep only the $|+\rangle$ in π . The correlator then becomes

$$g_{ij} = \frac{\lambda_+^{N-(j-i)}}{\lambda_+^N} \left\{ s_{++}^2 \lambda_+^{j-i} + s_{+-} s_{-+} \lambda_-^{j-i} \right\} - s_{++}^2 \quad (72)$$

↑
from $-\langle S_i \rangle^2$
in g_{ij}

$$g_{ij} = \left(\frac{\lambda_-}{\lambda_+}\right)^{j-i} s_{+-} s_{-+} = e^{(j-i) \ln\left(\frac{\lambda_-}{\lambda_+}\right)} s_{+-} s_{-+} \quad (73)$$

Since $\lambda_- < \lambda_+$ define

$$\ln\left(\frac{\lambda_-}{\lambda_+}\right) = -\frac{1}{\xi} \quad (74)$$

$$g_{ij} = e^{-|j-i|/\xi} \text{ constant} \quad (75)$$

The correlations decay exponentially with distance, and are characterized by the **correlation length**

$$\xi = \frac{-1}{\ln(\lambda_-/\lambda_+)} \quad (76)$$

A few comments

- ① The 1D Ising Model has a free energy that is analytic as a f^n of J, h which implies that there is no phase transition.
- ② Generically, if a Π can be defined, a phase transition would imply a crossing of the eigenvalues of Π .
- ③ If a local operator connects the two crossing eigenvectors ξ will diverge at the transition.