Goal: Relation between thermodynamics & information theory
entropy & Shannon entropy (measure of missing info)
equilibrium ensembles & maximal entropy principle

Information - theoretical derivation of ensembles

Information content of probability distributions

Problem: Suppose you choose a number $0 \le n < M$ (say, $M=1000$). How many "binary questions" must I ask to find out n?

"Binary question" – question with YES/NO answer
$1/0$

Solution: Write n in binary representation

$$n = \sum_{k=0}^{\dots} n_k 2^k \qquad n_k = 0 \text{ or } 1 \qquad 0 \le k < k_{max}$$

$$\downarrow$$

$$0 \le n < 2^{k_{max}}$$

$\dots n_2 n_1 n_0$

binary		decimal
...000	=	∅
...001	=	1
...010	=	2
...011	=	3
...100	=	4
...101	=	5

⇒ Number of questions = (smallest number of bits needed to represent any number $0 \le n < M$) $= \log_2 M$

Disclosure of n has information content $= \log_2 M$ bits $= \log_2 \frac{1}{p}$,

$p = \frac{1}{M}$ – probability of event n.

General case:

* $i = 1, 2, \dots, M$ disjoint outcomes (events), p_i – probability of event i
* message that event i has occured has information content $-\log p_i$
* Information content of a randomly picked outcome (out of the distribution) $= \langle -\log p_i \rangle$

average information we gain from a trial

$$\boxed{H = -\sum_i p_i \log p_i}$$ – can be viewed as missing information about the state of the system (before any trial)

= "Shannon entropy"
= "information – theoretical entropy"

Examples

1) $p_{i_0} = 1$, $p_{i \neq i_0} = 0$ } $\Rightarrow$ any outcome is i_0, so do not need to try anything to know the state

$H = \emptyset$ – consistent with the expectation of no new information learned

2) $p_i = \frac{1}{M}$ $\Rightarrow$ $\boxed{H = \log M}$ – compare w entropy in the microcanonical ensemble

Maximal information gain because all events are equally likely.

Indeed, maximize $H = -\sum_i p_i \log p_i$ subject to only $\sum_i p_i = 1$
Lagrange multiplier λ

$$-\sum_i p_i \log p_i - \lambda\left(\sum_i p_i - 1\right)$$ – maximize over λ & p_i indep.

$$\frac{\delta}{\delta \lambda} = 0 \Rightarrow \sum_i p_i = 1$$

$$\frac{\delta}{\delta p_i} = 0 \qquad -\log p_i - 1 - \lambda = 0, \qquad p_i = e^{-1-\lambda} = \text{const}$$ indep of i

$$\Rightarrow p_i = \frac{1}{M}.$$

$p_i = \frac{1}{M}$ – natural subjective choice when we do not have any other information – least biased choice

$\Rightarrow$ Probability distribution $\{p_i\}$ is "least biased" or

Shannon entropy (Shannon's argument)

M possible outcomes : $i = 1, 2, \ldots, M$

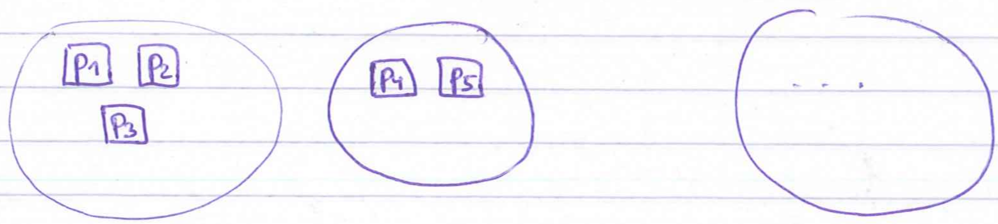
p_i – probability of outcome i

Shannon asks: how to quantify the uncertainty inherent in the distribution $\{p_i\}$

$$H_S(p_1, p_2, \ldots, p_M) \text{ – Shannon entropy ("missing information")}$$

Desirable to satisfy composition rule :

for any breaking into groups



$q_1 = p_1 + p_2 + p_3$, $q_2 = p_4 + p_5$, $q_3 = \ldots$ etc.

$$H_S(p_1, p_2, \ldots, p_M) = \overbrace{H_S(q_1, q_2, \ldots)}^{\text{measure of information from knowing which group}}$$

$$+ \underbrace{q_1}_{\text{prob. of group 1}} \underbrace{H_S\left(\frac{p_1}{q_1}, \frac{p_2}{q_1}, \ldots, \frac{p_a}{q_1}\right)}_{\text{measure of info from knowing within group}} + q_2 H_S\left(\frac{p_{a+1}}{q_2}, \ldots, \frac{p_{a+b}}{q_2}\right) + \ldots$$

Can prove that, up to overal constant factor,

$$\boxed{H_S(\{p_i\}) = -\sum_i p_i \ln p_i}$$

Check: $-q_1 \ln q_1 + q_1 \left(-\sum_{i=1}^{a} \frac{p_i}{q_1} \ln\left(\frac{p_i}{q_1}\right)\right) =$

$$= -q_1 \ln q_1 + q_1 \left(-\frac{1}{q_1} \sum_{i=1}^{a} p_i (\ln p_i - \ln q_i)\right) = -\sum_{i=1}^{a} p_i \ln p_i$$

Schematic proof:

⊛ Consider special case $p_1 = p_2 = \ldots = p_M = \frac{1}{M}$ and denote

$$H_s\left(\frac{1}{M}, \frac{1}{M}, \ldots, \frac{1}{M}\right) = J(M)$$

Consider B groups, with $\frac{M}{B}$ in each

$$q_i = \frac{1}{B} \quad ; \quad \frac{p_j}{q_i} = \frac{1/M}{1/B} = \frac{B}{M}$$

$$\underbrace{J(M)}_{= X \cdot Y} = H_s\left(\frac{1}{M}, \frac{1}{M}, \ldots\right) = H_s\left(\frac{1}{B}, \frac{1}{B}, \ldots\right) + B \times \frac{1}{B} H_s\left(\frac{B}{M}, \frac{B}{M}, \ldots\right)$$

$$= \underbrace{J(B)}_{= X} + \underbrace{J\left(\frac{M}{B}\right)}_{= Y}$$

$$J(X \cdot Y) = J(X) + J(Y) \Rightarrow J(X) = \ln X$$

⊛ $$H_s(q_1, q_2, \ldots) = H_s(p_1, p_2, \ldots, p_M) - \sum_i q_i H_s\left(\left\{\frac{p_\#}{q_i}\right\}\right)$$

with $p_i = \frac{1}{M}$

$$= J(M) - \sum_i q_i J(M \cdot q_i) = \ln M - \sum_i q_i \ln(M \cdot q_i)$$

$$= -\sum_i q_i \ln q_i \quad . \quad \text{— Shannon entropy.}$$

Imagine starting with all $p_i = \frac{1}{M}$; grouping into groups $\frac{q_i}{(1/M)} = q_i \cdot M$ in group i

– this is "engineering" of $\{q_i\}$ by grouping equally-distributed events.

"most likely" (among all distributions consistent with given data) when missing information H about the system is maximal

We are led to consider

$\Rightarrow$ **Maximal Entropy Principle**

Suppose we know only the expectation values $\langle A^{(\nu)} \rangle$ of the observables $A^{(\nu)}$ ($\nu = 1 \dots m$; $m \ll N$).
Then the "least biased" distribution $\{p_i\}$ consistent with the given expectation values of the observables is given by

maximizing $\boxed{S(p_1, \dots, p_N) = -k_B \sum_i p_i \ln p_i}$

subject to conditions

$$\boxed{\begin{aligned} & 0 \le p_i \le 1, \\ & \sum_i p_i = 1 \\ & \sum_i p_i A_i^{(\nu)} = \langle A^{(\nu)} \rangle, \quad \nu = 1, \dots, m \end{aligned}}$$

Theorem 1. The solution of maximal entropy principle is given by

$$\boxed{\begin{aligned} p_i &= \frac{1}{Z} e^{-\sum_\nu \lambda^{(\nu)} A_i^{(\nu)}} \\ Z &= \sum_i e^{-\sum_\nu \lambda^{(\nu)} A_i^{(\nu)}}, \end{aligned}}$$

– generalized Boltzman distribution

where the $\lambda^{(\nu)}$ are determined by the equations

$$\boxed{-\frac{\partial \ln Z}{\partial \lambda^{(\nu)}} = \langle A^{(\nu)} \rangle}$$

– given data

The covariances of $A^{(1)}, \dots, A^{(\nu)}, \dots, A^{(m)}$ are given by

$$\boxed{\left\langle \left[A^{(\nu)} - \langle A^{(\nu)} \rangle\right]\left[A^{(\nu')} - \langle A^{(\nu')} \rangle\right] \right\rangle \equiv \sum_i p_i \left(A_i^{(\nu)} - \langle A^{(\nu)} \rangle\right)\left(A_i^{(\nu')} - \langle A^{(\nu')} \rangle\right) = \frac{\partial^2 \ln Z}{\partial \lambda^{(\nu)} \partial \lambda^{(\nu')}}.}$$

Proof: Use method of Lagrange multipliers for the constrained maximization problem:

$$S(\{p_i\}) - k_B \lambda^{(0)}\left(\sum_i p_i - 1\right) - k_B \sum_\nu \lambda^{(\nu)} \left(\sum_i p_i A_i^{(\nu)} - \langle A^{(\nu)} \rangle\right)$$

– unconstrained minimization over $p_i, \lambda^{(\nu)}, \nu = 0, 1, \dots, m$

$$\frac{\delta}{\delta \lambda^{(0)}} = 0 \Rightarrow \sum_i p_i = 1$$

$$\frac{\delta}{\delta \lambda^{(\nu)}} = 0 \Rightarrow \sum_i p_i A_i^{(\nu)} = \langle A^{(\nu)} \rangle$$

$$\frac{\delta}{\delta p_i} = 0 \Rightarrow -\ln p_i - 1 - \lambda^{(0)} - \sum_\nu \lambda^{(\nu)} A_i^{(\nu)} = 0$$

$$p_i = \exp\left(-1 - \lambda^{(0)} - \sum_\nu \lambda^{(\nu)} A_i^{(\nu)}\right) =$$

$$= \underbrace{\text{const}}_{= 1/Z} \cdot e^{-\sum_\nu \lambda^{(\nu)} A_i^{(\nu)}}$$

$$\sum_i p_i = 1 \Rightarrow \boxed{Z = \sum_i e^{-\sum_\nu \lambda^{(\nu)} A_i^{(\nu)}}}$$

$$\boxed{p_i = \frac{1}{Z} e^{-\sum_\nu \lambda^{(\nu)} A_i^{(\nu)}}}$$

specified values

Require $\displaystyle \underbrace{\sum_i p_i A_i^{(\nu)}}_{\frac{1}{Z}\sum_i A_i^{(\nu)} e^{-\sum_{\nu'} \lambda^{(\nu')} A_i^{(\nu')}}} = \langle A^{\nu} \rangle$

$$\frac{1}{Z}\sum_i A_i^{(\nu)} e^{-\sum_{\nu'} \lambda^{(\nu')} A_i^{(\nu')}} = \frac{1}{Z}\cdot\frac{-\partial}{\partial \lambda^{(\nu)}} Z = -\frac{\partial \ln Z}{\partial \lambda^{(\nu)}}.$$

$\lambda^{(\nu)}$ are determined from $\boxed{-\frac{\partial \ln Z}{\partial \lambda^{(\nu)}}(\lambda^{(1)}, \lambda^{(2)}, \dots) = \langle A^{(\nu)} \rangle_{data}}$

Consider

$$\frac{\partial^2 \ln Z}{\partial \lambda^{(\nu)} \partial \lambda^{(\nu')}} = \frac{\partial}{\partial \lambda^{(\nu)}}\left(\frac{1}{Z}\frac{\partial Z}{\partial \lambda^{(\nu')}}\right) = \frac{1}{Z}\frac{\partial^2 Z}{\partial \lambda^{(\nu)} \partial \lambda^{(\nu')}} -$$

$$-\frac{1}{Z^2}\frac{\partial Z}{\partial \lambda^{(\nu)}}\frac{\partial Z}{\partial \lambda^{(\nu')}} = \frac{1}{Z}\sum_i A_i^{(\nu)} A_i^{(\nu')} e^{-\sum_\mu \lambda^{(\mu)} A_i^{(\mu)}} - \langle A^{(\nu)} \rangle_p \langle A^{(\nu')} \rangle$$

$$= \langle A^{(\nu)} A^{(\nu')} \rangle_p - \langle A^{(\nu)} \rangle_p \langle A^{(\nu')} \rangle_p$$

$$= \underbrace{\left\langle \left(A^{(\nu)} - \langle A^{(\nu)} \rangle\right)\left(A^{(\nu')} - \langle A^{(\nu')} \rangle\right)\right\rangle}_{= \langle A^{(\nu)} A^{(\nu')} \rangle - \underbrace{\langle A^{(\nu)} \cdot \langle A^{(\nu')} \rangle \rangle}_{\langle A^{(\nu)} \rangle \langle A^{(\nu')} \rangle} - \underbrace{\langle \langle A^{(\nu)} \rangle A^{(\nu')} \rangle}_{\langle A^{(\nu)} \rangle \langle A^{(\nu')} \rangle} + \underbrace{\langle \langle A^{(\nu)} \rangle \langle A^{(\nu')} \rangle \rangle}_{+\, \langle A^{(\nu)} \rangle \langle A^{(\nu')} \rangle}}.$$

— covariance of $A^{(\nu)}$ and $A^{(\nu')}$.

Theorem 2

Maximum of the entropy is given by

$$S = k_B \ln Z - k_B \sum_\nu \lambda^{(\nu)} \frac{\partial \ln Z}{\partial \lambda^{(\nu)}} \qquad (S \text{ as a function of } \lambda\text{'s})$$

$$= k_B \ln Z + k_B \sum_\nu \lambda^{(\nu)} \langle A^{(\nu)} \rangle \qquad (S \text{ as a function of } \langle A \rangle\text{'s})$$

functions of $\langle A \rangle$'s

Proof :

$$\boxed{S = -k_B \sum_i p_i \ln p_i = -k_B \sum_i p_i \ln\left(\frac{1}{Z} e^{-\sum_\nu \lambda^{(\nu)} A_i^{(\nu)}}\right)}$$

$$= k_B \ln Z + k_B \sum_i p_i \left(\sum_\nu \lambda^{(\nu)} A_i^{(\nu)}\right) =$$

$$= \boxed{k_B \ln Z + k_B \sum_\nu \lambda^{(\nu)} \langle A^{(\nu)} \rangle} \quad \blacktriangle$$

Remark: meaning of λ-s

Suppose we solve for $\lambda^{(\nu)}$ from $-\frac{\partial \ln Z}{\partial \lambda^{(\nu)}} = \langle A^{(\nu)} \rangle$ and thus express S as a function of $\langle A^{(\nu)} \rangle$.

$$\Rightarrow \boxed{\frac{\partial S}{\partial \langle A^{(\nu)} \rangle}} = k_B \sum_\mu \frac{\partial \ln Z}{\partial \lambda^{(\mu)}} \cdot \frac{\partial \lambda^{(\mu)}}{\partial \langle A^{(\nu)} \rangle} +$$

$$+ k_B \sum_\mu \frac{\partial \lambda^{(\mu)}}{\partial \langle A^{(\nu)} \rangle} \cdot \langle A^{(\mu)} \rangle + k_B \lambda^{(\nu)} = \boxed{k_B \lambda^{(\nu)}}$$

⇒ meaning of λ's

Remark: connection with Legendre transforms; meaning of ln Z

Consider function $f(\lambda^{(1)}, \dots) = k_B \ln Z(\lambda^{(1)}, \lambda^{(2)}, \dots)$ and consider Legendre transformation

$$\phi(\underset{\substack{\uparrow \\ \text{new variables}}}{x^{(1)}, x^{(2)}, \dots}) = \min_{\lambda^{(1)}, \lambda^{(2)}, \dots} \Big\{ f(\lambda^{(1)}, \lambda^{(2)}, \dots) - \sum_\nu \underset{\substack{\uparrow \\ \text{old vars}}}{\lambda^{(\nu)}} \cdot (\underbrace{-k_B \cdot x^{(\nu)}}_{\text{new vars}}) \Big\}$$

$$\frac{\delta}{\delta \lambda^{(\nu)}} = 0 \Rightarrow \left.\frac{\partial \ln Z}{\partial \lambda^{(\nu)}} + x^{(\nu)} = 0 \right] \text{- equations for } \lambda^{(1)}, \lambda^{(2)}, \dots$$

$$-k_B x^{(\nu)} = \frac{\partial}{\partial \lambda^{(\nu)}} (k_B \ln Z)$$

Inverse Legendre transform: $\lambda^{(\nu)} = -\frac{\partial \phi(x)}{\partial(-k_B x^{(\nu)})}$

Compare $x^{(\nu)} = -\frac{\partial}{\partial \lambda^{(\nu)}} \ln Z$

with $\langle A^{(\nu)} \rangle = -\frac{\partial}{\partial \lambda^{(\nu)}} \ln Z$

$\Rightarrow x^{(\nu)} = \langle A^{(\nu)} \rangle$

Compare $\phi(x^{(1)}, x^{(2)}, \dots) = k_B \ln Z + k_B \sum_\nu \lambda^{(\nu)} x^{(\nu)}$

with $S = k_B \ln Z + k_B \sum_\nu \lambda^{(\nu)} \langle A^{(\nu)} \rangle$

$$\Rightarrow \phi(x^{(1)}, x^{(2)}, \dots) = S(\langle A^{(1)} \rangle, \langle A^{(2)} \rangle, \dots).$$

In particular,

$$\Rightarrow \boxed{k_B \lambda^{(\nu)} = \frac{\partial S}{\partial \langle A^{(\nu)} \rangle}}$$

Earlier in thermodynamics we were Legendre-transforming energy rather than entropy. The connection with the above is as follows:

Thermodynamics: $U(S, V, N)$, $T = \left(\frac{\partial U}{\partial S}\right)_{V,N}$

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

$$-\frac{F}{\left(\frac{\partial U}{\partial S}\right)_{V,N}} = S - \frac{U}{\left(\frac{\partial U}{\partial S}\right)_{V,N}} = S - \left(\frac{\partial S}{\partial U}\right)_{V,N} \cdot U$$

Thus, $\underbrace{-\frac{F}{\left(\frac{\partial U}{\partial S}\right)_{V,N}}}_{k_B \ln Z}$ is Legendre transform of S with respect to U

$$k_B \ln Z = S - S \cdot \frac{\partial S}{\partial \langle E \rangle}$$

precisely as above in the maxEnt derivation of ensembles

Application to Statistical Mechanics

Relation between information theory and thermodynamics:

1) Thermodynamic quantities = averages with respect to probability distribution $\{p_i\}$

2) Thermodynamic entropy
$\sim$ Shannon entropy = missing information about the system (relative to what we know from macroscopic measurements)

3) Equilibrium $\Leftrightarrow$ maximum missing information

Derivation of ensembles

Microcanonical: fixed $E \le E_i \le E + \delta E$

$$N_i = N$$

No macro-condition of type $\sum_i p_i A_i^{(\nu)} = \langle A^{(\nu)} \rangle$
(only $\sum_i p_i = 1$) $\Rightarrow$ no $\lambda^{(\nu)}$ for $\nu = 1, 2, \ldots$; only $\lambda^{(0)}$

$\Rightarrow$
- $p_i = \frac{1}{Z} \cdot e^0 = \frac{1}{Z}$
- $Z = \sum_i e^0 = \Omega$
- No lagrange params $\lambda^{(1)}, \ldots$
- entropy $S = k_B \ln Z$
- entropy as a function of macroscopic data

$$S(E, V, N)$$

- meaning of Z: $k_B \ln Z(E, V, N) = S(E, V, N)$

Canonical ensemble

Macroscopic conditions of parametric type: $\overset{\text{"fixed"}}{N_i = N}$
(fixed N)

Macroscopic conditions of "on average" type:

$$\sum_i p_i E_i = \langle E \rangle$$

- probability $p_i = \frac{1}{Z} e^{-\beta E_i}$
- partition function $Z = \sum_i e^{-\beta E_i}$
- Lagrange parameter β for $\langle E \rangle$
- Entropy $S = k_B \ln Z + k_B \cdot \beta \cdot \langle E \rangle$
- S as a function of macroscopic data $S(\langle E \rangle, V, N)$
- meaning of Lagrange parameter:

$$k_B \beta = \frac{\partial S}{\partial \langle E \rangle} = \frac{1}{T}$$

from $dS = \frac{1}{T} dU + \frac{p}{T} dV - \frac{\mu}{T} dN$

- $S = k_B \ln Z + \frac{\langle E \rangle}{T}$
- compare with thermodynamic relations

$$S = -\frac{F}{T} + \frac{U}{T} \qquad (F = U - TS)$$

$\Rightarrow$ gives meaning of Z:

$$k_B \ln Z(T, V, N) = -\frac{F(T, V, N)}{T}$$

	microcanonical ensemble	canonical ensemble	grand-canonical ensemble	pressure ensemble
macroscopic conditions of parametric type	$E \le E_i \le E+\delta E$ $V_i = V$ $N_i = N$	$V_i = V$ $N_i = N$	$V_i = V$	$N_i = N$
macroscopic conditions of "on-average" type	—	$\sum_i p_i E_i = \langle E \rangle$	$\sum_i p_i E_i = \langle E \rangle$ $\sum_i p_i N_i = \langle N \rangle$	$\sum_i p_i E_i = \langle E \rangle$ $\sum_i p_i V_i = \langle V \rangle$
probability p_i	$\frac{1}{Z} e^{\emptyset}$	$\frac{1}{Z} e^{-\beta E_i}$	$\frac{1}{Z} e^{-\beta E_i - \gamma N_i}$	$\frac{1}{Z} e^{-\beta E_i - \lambda V_i}$
partition fnctn Z	$\sum_i e^{\emptyset} = \Omega$	$\sum_i e^{-\beta E_i}$	$\sum_i e^{-\beta E_i - \gamma N_i}$	$\sum_i e^{-\beta E_i - \lambda V_i}$
Lagrange parameters	—	β (for $\langle E \rangle$)	β (for $\langle E \rangle$), γ (for $\langle N \rangle$)	β (for $\langle E \rangle$) λ (for $\langle V \rangle$)
entropy S	$k_B \ln Z$	$k_B \ln Z + k_B \beta \langle E \rangle$	$k_B \ln Z + k_B \beta \langle E \rangle +$ $+ k_B \gamma \langle N \rangle$	$k_B \ln Z + k_B \beta \langle E \rangle$ $+ k_B \gamma \langle N \rangle$
S as a fnctn of macr. data	$S(E, V, N)$	$S(\langle E \rangle, V, N)$	$S(\langle E \rangle, V, \langle N \rangle)$	$S(\langle E \rangle, \langle V \rangle, N)$
meaning of Lagrange params	—	$k_B \beta = \frac{\partial S}{\partial \langle E \rangle} = \frac{1}{T}$	$k_B \gamma = \frac{\partial S}{\partial \langle N \rangle} = -\frac{\mu}{T}$	$k_B \lambda = \frac{\partial S}{\partial \langle V \rangle} = \frac{p}{T}$

$$dS(E,V,N) = \frac{1}{T} dE + \frac{p}{T} dV - \frac{\mu}{T} dN$$

	microcanonical ensemble	canonical ensemble	grand-canonical ensemble	pressure ensemble
entropy	$S = k_B \ln Z$	$S = k_B \ln Z + \frac{\langle E \rangle}{T}$	$S = k_B \ln Z + \frac{\langle E \rangle}{T} - \frac{\mu \langle N \rangle}{T}$	$S = k_B \ln Z + \frac{\langle E \rangle}{T} + \frac{p \langle V \rangle}{T}$
Compare with thermodynamic relations	—	$S = -\frac{F}{T} + \frac{U}{T}$	$S = \frac{-\Omega}{T} + \frac{U}{T} - \frac{\mu N}{T}$ "grand potential" $\Omega = -pV$ (Gibbs-Duheim)	$S = \frac{pV}{T} + \frac{U}{T} - \underbrace{\frac{\mu N}{T}}_{\to G/T}$ "Gibbs potential"
gives meaning of Z	$k_B \ln Z(E,V,N) = S(E,V,N)$	$k_B \ln Z(T,V,N) = -\frac{F(T,V,N)}{T}$	$k_B \ln Z(T,V,\mu) = -\frac{\Omega(T,V,\mu)}{T} = \frac{pV}{T}$	$k_B \ln Z(T,p,N) = -\frac{G(T,p,N)}{T}$
Z yields all other thdyn. quantities	$\left(\frac{\partial S}{\partial E}\right)_{V,N} = \frac{1}{T}$ $\left(\frac{\partial S}{\partial V}\right)_{E,N} = \frac{p}{T}$ $\left(\frac{\partial S}{\partial N}\right)_{E,V} = -\frac{\mu}{T}$	$\left(\frac{\partial F}{\partial T}\right)_{V,N} = -S$ $\left(\frac{\partial F}{\partial V}\right)_{T,N} = -p$ $\left(\frac{\partial F}{\partial N}\right)_{T,V} = +\mu$ $\left(\frac{\partial \ln Z}{\partial \left(\frac{1}{k_BT}\right)}\right)_{V,N} = -U$	$\left(\frac{\partial \ln Z}{\partial \left(\frac{1}{k_BT}\right)}\right)_{V,\mu} = -U + \mu N$ $\left(\frac{\partial \ln Z}{\partial \left(\frac{\mu}{k_B}\right)}\right)_{T,V} = \frac{N}{T}$ $\left(\frac{\partial \ln Z}{\partial \beta}\right)_{\gamma} = -U$	$\left(\frac{\partial G}{\partial T}\right)_{p,N} = -S$ $\left(\frac{\partial G}{\partial p}\right)_{T,N} = V$ $\left(\frac{\partial G}{\partial N}\right)_{T,V} = \mu$ $\left(\frac{\partial \ln Z}{\partial \left(\frac{1}{k_BT}\right)}\right)_{p,N} = -U - pV$ $\left(\frac{\partial \ln Z}{\partial \left(\frac{p}{T}\right)}\right)_{T,N} = -\frac{V}{T}$
ensemble convenient	when Ω (# of micro-states) is easy to compute [rare]; when molecular dynamics simulations are performed	for interacting systems when fixed N poses no problem	when fixed N yields unwieldy condition; when phase equilibria at issue (common μ)	when fixed V yields unwieldy constraints (rare); when Monte Carlo simulations at const. p are performed.

Fluctuations. Equivalences of Ensembles.
3rd law of thermodynamics

Variance

$$\left\langle \left(A^{(\nu)} - \langle A^{(\nu)}\rangle\right)^2 \right\rangle = \frac{\partial^2 \ln Z}{\partial (\lambda^{(\nu)})^2} = -\frac{\partial \langle A^{(\nu)}\rangle}{\partial \lambda^{(\nu)}}$$

→ change of $\langle A^{(\nu)}\rangle$ due to change in external parameter $\lambda^{(\nu)}$

≡ "response function" : gain/loss of $A^{(\nu)}$ from/to reservoir, i.e "dissipation"

→ fluctuation – dissipation theorem

$$\sqrt{\frac{\left\langle \left(A^{(\nu)} - \langle A^{(\nu)}\rangle\right)^2 \right\rangle}{\langle A^{(\nu)}\rangle^2}} = \frac{1}{\langle A^{(\nu)}\rangle}\sqrt{-\frac{\partial \langle A^{(\nu)}\rangle}{\partial \lambda^{(\nu)}}} \sim \frac{\sqrt{N}}{N} \sim \frac{1}{\sqrt{N}}$$

$\to 0$ as $N \to \infty$

Consequences:

* real fluctuations $\sim 10^{-11}$ for $N \sim 10^{22}$
* almost all systems in GCE (grand canonical ensemble) have the same N ⇒ GCE ⇔ CE
* almost all states in CE (canonical ensemble) have the same E ⇒ CE ⇔ MCE (micro-canonical ensemble)

Examples : * $A^{(\nu)} = E$ in CE or GCE :

$$\left\langle (E - \langle E\rangle)^2 \right\rangle = -\frac{\partial \langle E\rangle}{\partial \frac{1}{k_B T}} = k_B T^2 \frac{\partial \langle E\rangle}{\partial T} = k_B T^2 C_V$$

* $A^{(\nu)} = N$ in GCE

$$\langle (N - \langle N \rangle)^2 \rangle = -\left(\frac{\partial \langle N \rangle}{\partial \left(-\frac{\mu}{k_B T} \right)} \right)_{T,V} = k_B T \left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{T,V}$$

Third law of thermodynamics

Suppose we work in the canonical ensemble and

$-\infty < E_1 < E_2 < \dots$

$\uparrow \quad \uparrow$ energy eigenvalues

Then

$$\lim_{T\to 0} p_i = \lim_{\beta\to\infty} \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}} = \lim_{\beta\to\infty} \frac{e^{-\beta(E_i - E_1)}}{1+\sum_{j=2}^{\infty} e^{-\beta(E_j-E_1)}}$$

$$= \begin{cases} 1 & \text{if } i=1 \\ \emptyset & \text{if } i\neq 1 \end{cases}$$

$\Rightarrow$ Fluctuations for any observable go to $\emptyset$ as $T\to\emptyset$

$$\left.\begin{aligned} \lim_{T\to 0} U &= E_1 \\ \lim_{T\to 0} S &= 0 \\ \lim_{T\to 0} F &= E_1 \end{aligned}\right\}$$ – 3rd law of thermodynamics, assuming non-degenerate ground state of a quantum system.

$$\left. S(V,T) = \int_0^T c_V(V,T') \frac{dT'}{T'} \longrightarrow 0 \text{ for } T\to 0 \right]$$ – this implies that $C_V(V,T) \to 0$ as $T\to 0$.