1　熱力学の復習

物理の学問を学んでいく上で大前提となるのはその問題設定 (仮定) である。新しく学問を学ぶ上でその学問の適用範囲を知っておくことは何よりも大事である。

1.1　熱力学の問題設定

(1) マクロな系に注目（熱力学が生まれた当時、系のミクロな構造は未発見）
(2) 始状態と終状態は熱平衡状態 (系が均一であり、マクロなカレントが存在しない状態)
(3) エントロピーの存在 (と諸々の性質) およびエントロピー最大原理（孤立系ではエントロピーが最大となるように熱平衡状態が決まる）

コメントになるが、統計力学では、よりミクロなモデルから出発することが出来、熱力学よりも解析できる物理量自体は広い。統計力学における仮定として等重率の原理が出てくるが、これはエントロピー最大の原理の言い換えになっている事を後で学ぶ。

1.2　エントロピーに関する仮定と帰結

1.2.1　熱力学第１法則

$$\Delta U = \Delta W_{in} + \Delta Q_{in}, \tag{1}$$
$$\Delta Q_{in} = T\Delta S \tag{2}$$

によってエントロピー S の存在を認め、定義する。ここで U は内部エネルギー, W_{in} は (外界から受けた) 仕事量, Q_{in} は熱 (吸熱量) である。以下では簡単のために $W = W_{in}, Q = Q_{in}$ として表記する。U と W は計測できる量として、両者を測った際に差異が生じるので、その差分を「熱」として定義し、それを温度で割ったものをエントロピー S と定義する。

1.2.2　エントロピーの相加性

「X が相加変数である」とは、ある独立な２つの系が存在する時に、それぞれの系の変数 X_1, X_2 と２つを合わせた系の変数 X_{12} の間に $X_{12} = X_1 + X_2$ が成立するような変数

である。熱力学では特に内部エネルギー U, 体積 V, 粒子数 N を考える。熱力学ではエントロピーに関して、

1. エントロピーは相加変数の関数として表せる ($S = S(U,V,N)$)
2. エントロピーの相加性 (独立な系 1,2 に対して $S_{12} = S_1(U_1,V_1,N_1) + S_2(U_2,V_2,N_2)$)

均一系を考えると、同一な系を n 個合わせて

$$S(U,V,N) = n\,S\big(U/n,\,V/n,\,N/n\big). \tag{3}$$

が成立

1.2.3 エントロピーの連続性

エントロピーが U,V,N について連続な変数であると仮定すると、式 (3) を $\lambda = 1/n$ を正の実数に拡張でき、任意の実数 $\lambda \geq 0$ に対して

$$S(\lambda U, \lambda V, \lambda N) = \lambda\,S(U,V,N). \tag{4}$$

1.2.4 エントロピー最大原理と凸性

ある一つの系に全てを遮断する仕切りを入れ二つに分割する事を考える。分割する前は分割後の 2 つの系は必ずしも独立ではないとすると任意の $0 \leq \lambda \leq 1$ について

$$\begin{aligned} &S\big(\lambda U_1 + (1-\lambda)U_2,\, \lambda V_1 + (1-\lambda)V_2,\, \lambda N_1 + (1-\lambda)N_2\big) \\ &\qquad \geq \lambda\,S(U_1,V_1,N_1) + (1-\lambda)\,S(U_2,V_2,N_2). \end{aligned} \tag{5}$$

が成立する。仕切りはある種の制約条件として捉えることが出来るが、制約なしでのエントロピーの最大化と制約付きでの最大化を行った際に、制約なしの最大化の方が大きくなる (等号含む) 事から上記の不等式を得ることが出来る。

凸性: 関数 $f(\vec{x})$ が以下の性質を満たす時、$f(\vec{x})$ を凸関数 (下に凸な関数) と呼ぶ

$$f(\lambda\vec{x_1} + (1-\lambda)\vec{x_2}) \geq \lambda f(\vec{x_1}) + (1-\lambda)f(\vec{x_2}), \quad (0 \leq \lambda \leq 1, \forall\vec{x_1},\vec{x_2})$$

上の式と大小関係が逆の不等式を満たす場合は、上に凸な関数と呼ぶ。また $f(\vec{x})$ が (少なくとも 2 回微分可能な程度に) 滑らかな関数である場合、上の条件は以下と等

価である。
f の Hessian Matrix(ヘッセ行列)

$$\mathrm{H}[f(\vec{x})] = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}$$

が半正定値 (全ての固有値が非負) である。(上に凸な関数な場合は、「全ての固有値が非正」が必要十分条件)

1.2.5 (強) 増加性

$X > X'$ ならば $f(x) \geq f(X')$ を満たす時、f は X に対して増加性を持つという。(等号が外れた場合は「強増加性」と呼ぶ) $S(U,V,N)$ は U,V,N について増加性を持ち、特に U について強増加性を持つと仮定すると、S の関数形 $S = S(U,V,N)$ がわかっていれば、U の関数形 $U = U(S,V,N)$ を一意に決定できる。

示量変数 U,V,N に対して示強変数 (相加性を持つ関数を相加変数で微分したもの) を

$$T = \left(\frac{\partial U}{\partial S}\right)_{V,N}, \quad p = -\left(\frac{\partial U}{\partial V}\right)_{S,N}, \quad \mu = \left(\frac{\partial U}{\partial N}\right)_{S,V} \tag{6}$$

と定義する．これにより $S(U,V,N)$ の関数形が分かっていれば $U(S,V,N)$ を決定でき、そこから示強変数：(T,p,μ) も同定できる。このように、関数形を知っていれば熱力学に必要な変数と量を全て同定することが出来る関数を「完全な熱力学量」と呼ぶ。($U(S,V,N)$ についても同様)

1.2.6 熱力学第二法則

Thomson の原理いわゆる熱力学第二法則
任意の温度における等温サイクルにおいて

$$W_{\mathrm{cyc}}^{\mathrm{out}} = 0 \quad (\text{total で取り出す仕事の和は 0 以下})$$
$$(\Rightarrow \text{ 第二種永久機関の禁止})$$

いくつか等価な statement が存在する (最大仕事の原理, カラテオドリの定理など)。そのうちの一つ、Clausius の不等式は、

$$\text{等温過程において } \Delta S \geq \frac{\Delta Q}{T} \tag{7}$$

1.3 Legendre 変換

Legendre(ルジャンドル) 変換は凸性を保存する変換である。関数が凸性を持つ (凸関数である) 事が保証されているといくつか嬉しい事がある。(例えば $f(x)$ 下に凸である事が保証されていれば、$\partial f/\partial x = 0$ を解けば、$f(x)$ の最大値を与える x が直ちにわかる。他にものちに確率・統計についての章でやる事になるが、凸関数についての期待値には Jensen(イェンセン) の不等式が成立し、これも使い勝手が良い。)

下に凸な滑らかな関数 $f(x)$ に対する以下の変換をルジャンドル変換と呼ぶ

$$g(w) = \max_x \Big[wx - f(x) \Big] \tag{8}$$

また以下のように定義することも出来る。(x_w は w に陽に依存する w の関数として扱う)

$$g(w) = wx_w - f(x_w), \left(w = \frac{\partial f}{\partial x}\bigg|_{x_w} \right) \tag{9}$$

物理では後者の方の定義を使う事が多い。この変換は図のように視覚的に理解する事が出来る。つまりある変数 x について下に凸の関数 $f(x)$ を x における接線の傾き $w = \frac{\partial f}{\partial x}$ とその y 切片 $-g(w)$ によって表現しなおすような変換になっている。

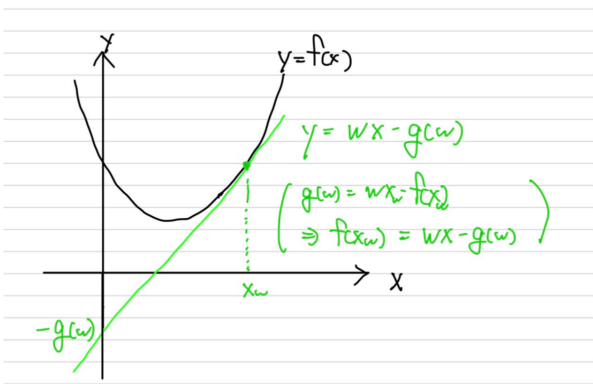

図１　ルジャンドル変換のイメージ

さて変換後の新たな関数 $g(w)$ が w について凸性が維持されている事を確認しよう。

$$\frac{\partial g}{\partial w} = x_w + w * \frac{\partial x_w}{\partial w} - \left.\frac{\partial f}{\partial x}\right|_{x_w} * \frac{\partial x_w}{\partial w} = x_w + \left(w - \left.\frac{\partial f}{\partial x}\right|_{x_w}\right) = x_w \tag{10}$$

$$w = \left.\frac{\partial f}{\partial x}\right|_{x_w} \Rightarrow 1 = \frac{\partial}{\partial w}\left(\left.\frac{\partial f}{\partial x}\right|_{x_w}\right) = \frac{\partial x_w}{\partial w} * \left.\frac{\partial^2 f}{\partial x^2}\right|_{x_w} = \frac{\partial^2 g}{\partial w^2} * \left.\frac{\partial^2 f}{\partial x^2}\right|_{x_w} \tag{11}$$

$f(x)$ は下に凸より $\frac{\partial^2 f}{\partial x^2} > 0(\forall x)$。よって式 (11) より $\frac{\partial^2 g}{\partial w^2} > 0(\forall w)$。よって変換後の $g(w)$ も下に凸の関数である。

1.4 内部エネルギー U の下凸性

以下簡略化のために、V, N をまとめて X と書く。(つまり本当は X はベクトル変数)。S は U について強増加 ($\frac{\partial S}{\partial U}$) であり、$S(U, X)$ は上に凸な関数であるので、

$$H[S] = \begin{pmatrix} \frac{\partial^2 S}{\partial U^2} & \frac{\partial^2 S}{\partial U \partial X}, & \frac{\partial^2 S}{\partial X \partial U} & \frac{\partial^2 S}{\partial X^2} \end{pmatrix} \leq 0 \ \ , \frac{\partial S}{\partial U} > 0 \tag{12}$$

行列 $\hat{M}$ の半正定値性 ($\hat{M} \geq 0$) は以下と等価である。$\forall \vec{v}$ に対して

$$\vec{v}^T \hat{M} \vec{v} \geq 0 \tag{13}$$

例えば $\vec{v} = [1, 0]$ とすれば、$H[S] \leq 0$ より $\frac{\partial^2 S}{\partial U^2} \leq 0$, $\vec{v} = [0, 1]$ とすれば、$\frac{\partial^2 S}{\partial X^2} \leq 0$ である事が分かる。さて式 (12) の条件を使って、内部エネルギー $U(S, X)$ が下に凸な関数である事を示してみよう。すなわち

$$H[U] \geq 0 \Leftrightarrow \vec{v}^T H[U] \vec{v} \geq 0 \text{ を満たす線型独立なベクトルを 2 つ存在する}$$

を示せば良い。

証明：$U = U(S, X) = U(S(U, X), X)$ に対して、両辺をそれぞれ偏微分すると以下の関係式が得られる。(左辺は変数 U であり、右辺は関数 U である事に注意。)

$$\frac{\partial}{\partial U} \Rightarrow 1 = \frac{\partial U}{\partial S}\frac{\partial S}{\partial U}\left(\Rightarrow \frac{\partial U}{\partial S} > 0\right) \tag{14}$$

$$\frac{\partial^2}{\partial U^2} \Rightarrow 0 = \frac{\partial^2 U}{\partial S^2}\left(\frac{\partial S}{\partial U}\right)^2 + \frac{\partial U}{\partial S}\frac{\partial^2 S}{\partial U^2} \tag{15}$$

$$\frac{\partial^2}{\partial U \partial X} \Rightarrow 0 = \frac{\partial^2 U}{\partial S^2}\frac{\partial S}{\partial U}\frac{\partial S}{\partial X} + \frac{\partial U}{\partial S}\frac{\partial^2 S}{\partial U \partial X} + \frac{\partial^2 U}{\partial S \partial X}\frac{\partial S}{\partial U} \tag{16}$$

$$\frac{\partial^2}{\partial X^2} \Rightarrow 0 = \frac{\partial^2 U}{\partial S^2}\left(\frac{\partial S}{\partial X}\right)^2 + 2\frac{\partial^2 U}{\partial S \partial X}\frac{\partial S}{\partial X} + \frac{\partial^2 U}{\partial X^2} \tag{17}$$

ここで $\vec{v_1}=[\frac{\partial S}{\partial U},0]$, $\vec{v_2}=[\frac{\partial S}{\partial X},1]$ とすれば

$$\vec{v_1}^T H[\hat{U}]\vec{v_1}=\left(\frac{\partial S}{\partial U}\right)^2\frac{\partial^2 U}{\partial S}=-\frac{\partial U}{\partial S}\frac{\partial^2 S}{\partial U^2}\geq 0\quad\left(\frac{\partial U}{\partial S}>0,\frac{\partial^2 S}{\partial U^2}\leq 0\text{ より}\right)\tag{18}$$

$$\vec{v_2}^T H[\hat{U}]\vec{v_2}=-\frac{\partial U}{\partial S}\frac{\partial^2 S}{\partial X^2}\geq 0\quad\left(\text{途中式変形は省略した}\right)\tag{19}$$

$\vec{v_1}$ は $\frac{\partial S}{\partial U}>0$ より $\vec{v_1},\vec{v_2}$ は線型独立 ($\vec{v_1}=0=0\vec{v_2}$ になる事がない) なので、内部エネルギー $U(S,X)$ は下に凸な関数である事が示された。

1.5 エネルギー最小の原理

孤立した系においては ((V,N) を固定したとき)U が最小になるように (各部分系の) 状態が決まる

$$U(\{S_1,S_2\},\{V_1,V_2\},\{N_1,N_2\})=U_1(S_1,V_1,N_1)+U_2(S_2,V_2,N_2)\tag{20}$$

平衡状態においては系は均一であり、よって部分系の示強変数は一致するはずである。

$$\text{示強変数 } T_i=\frac{\partial U_i}{\partial S_i},\ p_i=-\frac{\partial U_i}{\partial V_i},\ \mu_i=\frac{\partial U_i}{\partial N_i}\tag{21}$$

$$T=\frac{\partial U_1}{\partial S_1}=\frac{\partial U_2}{\partial S_2},\ p=-\frac{\partial U_1}{\partial V_1}=-\frac{\partial U_2}{\partial V_2},\ \mu=\frac{\partial U_1}{\partial N_1}=\frac{\partial U_2}{\partial N_2}\tag{22}$$

今全系のエントロピー S や体積 V, 粒子数 N は一定であり、仮想的な系の分割に依存しないとすれば

$$S_2=S-S_1,\quad V_2=V-V_1,\quad N_2=N-N_1\tag{23}$$

であるから、$\frac{\partial U_2(S_2)}{\partial S_2}=\frac{\partial U_2(S-S_1)}{\partial (S-S_1)}=-\frac{\partial U_2(S-S_1)}{\partial S_1}$

$$\frac{\partial U}{\partial S_1}=\frac{\partial U_1}{\partial S_1}+\frac{\partial U_2}{\partial S_1}=\frac{\partial U_1}{\partial S_1}-\frac{\partial U_2}{\partial S_2}=0\tag{24}$$

同様にして $\frac{\partial U}{\partial V_1}=0,\ \frac{\partial U}{\partial N_1}=0$ が成立。 U は下に凸より U は最小値を取る。

1.6 (Helmholtz の) 自由エネルギー最小の原理

$$\begin{aligned}U_{\mathrm{tot}}&=U_1+U_2+U_B\\&=U_1(S_1,V_1,N_1)+U_2(S_2,V_2,N_2)+U_B(S-S_1-S_2,\ V-V_1-V_2,\ N-N_1-N_2)\end{aligned}\tag{25}$$

$S_{\mathrm{tot}} = S_1 + S_2 + S_B$ であり V, N についても同様、ただし、$S_B \gg S_1, S_2,\ V_B \gg V_1, V_2$

$$U_B(S - S_1 - S_2,\ V - V_1 - V_2,\ N - N_1 - N_2) \simeq U_B(S) + (S_1 + S_2)\underbrace{\left(\frac{\partial U}{\partial S}\right)}_{T} \tag{26}$$

$$\begin{aligned} U_{\mathrm{tot}} &= U_1 + U_2 - (S_1 + S_2)T + U_B(S) \\ &= \underbrace{(U_1 - S_1 T)}_{=F_1} + \underbrace{(U_2 - S_2 T)}_{=F_2} + U_B(S) \end{aligned}$$

F : 自由エネルギー

$$U_{\mathrm{tot}} - U_B = F_1 + F_2 \tag{27}$$

U_B は固定して考えると U_{tot}「(孤立系) のエネルギー最小の原理」は、「(熱浴に浸っている) 注目系の自由エネルギー最小の原理」に書き換えられる。ちなみにこの Helmholtz の自由エネルギーは U を変数 S についてルジャンドル変換したものである。(つまり S 依存性は消えており、T について上に凸な関数になっている)

1.7 統計力学へ

1.7.1 問題設定

(1) (大数の法則が成立するような) マクロな系を考える

(2) 等重率の原理を仮定

(3) 熱平衡 ($\oplus$ 線形非平衡) 状態を考える

1.7.2 立ち位置

1. (2)$\Rightarrow$ 系の状態が確率的に実現するように記述
2. 1$\Rightarrow$ 確率統計のテクニックを使って物理量・物性を計算可
3. 1$\Rightarrow$ 確率分布を系のミクロな模型から決定

2 確率・統計の基礎

2.1 諸々の定義

確率変数: ある確率分布から生成される変数
$\hat{X} \sim p(X) \cdots$ 確率変数 $\hat{X}$ は確率分布 $p(X)$ によって生成される。($p(X)$ は $\hat{X} = X$ が生成される確率)

確率分布: ある確率変数 $\hat{X}$ がある値 X を取る確率 $p(X)$
確率分布は以下の条件を満たす。

$$\sum_n p_n = 1, \quad p_n \geq 0 \ (\forall n) \qquad \text{(離散確率変数の場合)}$$

$$\int dX p(X) = 1, \quad p(X) \geq 0 \ (\forall X) \qquad \text{(連続確率変数の場合)}$$

確率変数が離散値を取る場合、離散 $\sim$ と書く。(例：サイコロの目など)
確率変数が離散値を取る場合、連続 $\sim$ と書く。(例：紙飛行機を適当に飛ばして飛んだ距離など)

また今後のために以下も導入しておく

同時確率分布: ある 2 つの確率変数 $\hat{X}, \hat{Y}$ を同時に生成 (サンプル) し、その値がそれぞれ $\hat{X} = X, \hat{Y} = Y$ である確率分布。$p(X, Y)$ と書く。同時確率分布は以下を満たす。

$$\int dX \int dY \ p(X, Y) = 1 \ \left(\sum_{x,y} p_{xy} = 1 \ \text{(離散の場合。以降は省略する)} \right)$$

条件付き確率: ある条件のもとで確率変数 $\hat{X}$ について $\hat{X} = X$ が生成 (サンプル) される確率。$p(X|[\text{条件}])$ と書く。(条件がある確率変数 $\hat{Y}$ について $\hat{Y} = Y$ である場合 $p(X|Y)$ と書く。)

$$\int dX \ p(X|Y) = 1 \ \left(\text{逆の} \int dY \ p(X|Y) = 1 \ \text{は成立しないので注意} \right)$$

同時確率と条件付き確率の間に以下の関係が成り立つ。

$$p(X, Y) = p(X|Y)\ p(Y) = p(Y|X)\ p(X) \tag{28}$$

$$\Leftrightarrow p(Y|X) = p(X|Y)\ p(Y)/p(X)\ \ \ldots(\text{ベイズの定理}) \tag{29}$$

以後、視覚的に判別しやすくするため、離散確率変数の場合は $\hat{n}, p_n$ と書き、連続確率変数の場合は $\hat{X}, p(X)$ と書く事にする。

2.2 期待値・分散

期待値: 確率変数 $\hat{X}/\hat{n}$ の期待値は以下のように書き、定義する

$$\boldsymbol{E}[\hat{n}] = \sum_n n p_n \qquad (\text{離散確率変数の場合})$$

$$\boldsymbol{E}[\hat{X}] = \int dX\ X p(X) \qquad (\text{連続確率変数の場合})$$

分散: 確率変数 $\hat{X}/\hat{n}$ の分散は以下のように書き、定義する

$$\boldsymbol{V}[\hat{n}] = \sum_n \Big(n - \boldsymbol{E}[\hat{n}]\Big)^2 p_n \qquad (\text{離散確率変数の場合})$$

$$\boldsymbol{V}[\hat{X}] = \int dX\ \Big(X - \boldsymbol{E}[\hat{X}]\Big)^2 p(X) \qquad (\text{連続確率変数の場合})$$

2.3 大数の (弱) 法則

2.3.1 標本平均

i.i.d.(independent identically distributed) によって N 回確率変数 $\hat{X}$ の標本を取ることを考える。

$$\hat{m} = \frac{1}{N}\sum_{i=1}^{N} \hat{X}_i$$

ただし、$\hat{X}_i \sim p(X)$ で i 番目の標本を $\hat{X}_i = X_i$ とする。

元の確率変数の期待値が

$$\mathbb{E}[\hat{X}] = \sum_j X_j p_j = \mu \quad (\hat{X}_i \sim \{p_j\})$$

であるとすると、標本平均の期待値は以下のように、元の確率変数の期待値と同じ値になる。

$$\begin{aligned}\mathbb{E}[\hat{m}] &= \mathbb{E}\Big[\frac{1}{N}\sum_{i=1}^{N}\hat{X}_i\Big] \\ &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[\hat{X}_i] = \frac{1}{N}\sum_{i=1}^{N}\mu = \mu\end{aligned}$$

分散はどうであろうか？元の確率変数の分散が

$$\mathbb{V}[\hat{X}] = \sum_j (X_j - \mu)^2 p_j = \sigma^2$$

であるとする、標本平均の分散は

$$\begin{aligned}\mathbb{V}[\hat{m}] &= \mathbb{E}\Big[(\frac{1}{N}\sum_{i=1}^{N}\hat{X}_i - \mu)^2\Big] \\ &= \mathbb{E}\Big[\frac{1}{N^2}(\sum_{i=1}^{N}\hat{X}_i)^2 - 2\mu\underbrace{\frac{1}{N}\sum_{i=1}^{N}\hat{X}_i}_{=\mu} + \mu^2\Big] \\ &= \frac{1}{N^2}\mathbb{E}\Big[(\sum_{i=i}^{N}\hat{X}_i)^2\Big] - \mu^2\end{aligned}$$

2.3.2　Chebyckev の不等式

今後の便利のために以下の記法を導入する。

条件を満たす確率: 確率変数 $\hat{X}$ が条件を満たす確率を以下のように書き、定義する

$$\Pr[\text{条件式}] = \sum_j \theta(x_j)p_j\Big(= \int dx\ \theta(x)\ p(x)\Big)$$

$$\theta(x_j) = \begin{cases}1 & (x_j\text{が条件式を満たす}) \\ 0 & (\text{そうでない場合})\end{cases}$$

Chebyshev の不等式は以下で表される。

$$\Pr\Big[|\hat{x}-\mu| \geq \varepsilon\Big] \leq \frac{\sigma^2}{\varepsilon^2}$$

証明：

$$\Pr\Big[|\hat{x}-\mu| \geq \varepsilon\Big] = \int dx\ \theta(x)\ p(x), \qquad \theta(x) = \begin{cases} 1 & (|\hat{x}-\mu| > \varepsilon) \\ 0 & (\times) \end{cases} \tag{30}$$

ここで、$\theta(x)$ は以下の定義を満たす。

$$\theta(x) = \theta^2(x) \tag{31}$$

$$\theta(\hat{x}) \leq \frac{|\hat{x}-\mu|}{\varepsilon}\Big(\forall x\Big) \tag{32}$$

$$\theta(x) = \theta^2(x) \leq \frac{|x-\mu|^2}{\varepsilon^2} \tag{33}$$

よって

$$\Pr\Big[|\hat{x}-\mu| \geq \varepsilon\Big] = \int dx\ \theta(x)\ p(x) \leq \int dx \frac{|x-\mu|^2}{\varepsilon^2} p(x) = \frac{1}{\varepsilon^2} \underbrace{\int dx (x-\mu)^2 p(x)}_{\sigma^2} \tag{34}$$

まとめて

$$\Pr\Big[|\hat{x}-\mu| \geq \varepsilon\Big] \leq \frac{\sigma^2}{\varepsilon^2}$$

2.3.3 大数の (弱) 法則

Chebyshev の不等式と標本平均の分散の結果から

$$\Pr\Big[|\hat{m}-\mu| \geq \varepsilon\Big] \leq \frac{\sigma^2}{N\varepsilon^2} \tag{35}$$

より、標本平均が真の期待値から ε 以上ズレる確率は $N \to \infty$ でゼロ ($\Leftrightarrow$ 右辺の値を保ったままεを $N^{-1/2}$ で小さく出来る)

$\Rightarrow$ 有限の分散を持つ確率分布から i.i.d. で十分なサンプルを集めれば標本平均の期待値は真の期待値 μ に (平均二乗) 集束する。

2.4 生成関数

何かしらの期待値を計算する際、分布関数について積分する必要があるが、欲しい期待値ごとに積分を実行するのは、しばしば苦労を要する。そこで一回だけ積分を実行しておいて、のちに気になる物理量の n 乗の期待値を計算するときには微分するだけで良いようにする手法がある。これが生成関数を用いる手法である。(薄々勘づいているかもしれないが、積分よりも微分の方がはるかに簡単である。)

モーメント生成関数: $\hat{X} \sim p(X),\ f(\hat{X})$ 物理量 (確率変数の関数) としてモーメント生成関数を次で定義する。

$$G_f(\theta) \equiv \int dX\, p(X)\ \exp\Big[-i\theta f(X)\Big] \quad \Big(\text{離散確率変数なら} \sum_j p_j \exp\Big[-i\theta f(X_j)\Big]\Big)$$

また n 次のモーメントを以下で定義する。

$$G_f^{(n)} = \mathbb{E}\Big[f(\hat{X})^n\Big] = \left(\frac{\partial}{-i\partial\theta}\right)^2 G_f(\theta)\Big|_{\theta=0} \tag{36}$$

キュムラント生成関数: $\hat{X} \sim p(X),\ f(\hat{X})$ 物理量 (確率変数の関数) としてキュムラント生成関数を次で定義する。

$$C_f(\theta) \equiv \log \int dX\ p(X)\ \exp\Big[-i\theta f(X)\Big] = \log\, G_f(\theta)$$

また n 次のキュムラントを以下で定義する。

$$C_f^{(n)} = \left(\frac{\partial}{-i\partial\theta}\right)^2 C_f(\theta)\Big|_{\theta=0} \tag{37}$$

キュムラントの方が、モーメントよりもより分布の形を反映したものになっている。実際 2 次のキュムラントは分散と等価である。(余力があれば確認してみよ)

$$C^{(2)} = V\Big[f(\hat{X})\Big]$$

2.5 中心極限定理

「分散が有限な任意の確率分布の標本平均の確率分布は、標本数 (サンプル数) を増やしていくとガウス分布に収束する」という定理であり、確率統計の分野におけるガウス分布の重要性を示唆する定理である。

中心極限定理: 有限の期待値・分散 (μ, σ^2) を持つ確率分布 $p(X)$ から i.i.d. で集めた N サンプルについて

$$\hat{M} = \frac{\sqrt{N}\hat{m}}{\sigma} = \frac{1}{\sqrt{N}} \sum_i \Big(\frac{X_i - \mu}{\sigma}\Big)$$

の従う分布は $N \to \infty$ でガウス分布 $N(0,1)$ に収束する。

$$\begin{aligned}
G_M(\theta) &= \int dX\ p(X)\ \exp\Big[-i\theta M\Big] \\
&= \int dX\ p(X)\ \exp\Big[-i\theta\Big(\frac{\sqrt{N}}{\sigma}\frac{1}{N}\sum_{i=1}^{N}(x-\mu)\Big)\Big] \\
&= \int dX\ p(X)\ \prod_{i=1}^{N} \exp\Big[-i\theta\frac{1}{\sqrt{N}\sigma}(x-\mu)\Big] \\
&= \Big(\int dX\ p(X)\ \exp\Big[-i\theta\Big(\frac{x-\mu}{\sqrt{N}\sigma}\Big)\Big]\Big)^N = \Big(G_{(x-\mu)/\sqrt{N}\sigma}(\theta)\Big)^N \\
&= \Big(\int dX\ p(X)\ \Big[1 - i\theta\frac{(x-\mu)}{\sqrt{N}\sigma} - \frac{\theta^2}{2}\frac{(x-\mu)^2}{N\sigma^2} + \mathcal{O}[N^{-\frac{3}{2}}]\Big]\Big)^N \\
&= \Big(1 - i\theta\cdot 0 - \frac{\theta^2}{2N} + \mathcal{O}[N^{-\frac{3}{2}}]\Big)^N \\
&\xrightarrow{N\to\infty} \Big(1 - \frac{\theta^2}{2N}\Big)^N = \exp\Big[-\frac{\theta^2}{2}\Big] \\
G_M(\theta) &\xrightarrow{N\to\infty} \exp\Big[-\frac{\theta^2}{2}\Big]
\end{aligned}$$

3　ミクロカノニカル分布

3.1　等重率の原理とミクロカノニカル分布

孤立系の平衡状態の物理量の値は、以下のミクロカノニカル分布による期待値によって与えられる。

$$p_{MC}(x|E) = \begin{cases} \frac{1}{W_E} & (H(X) = E) \\ 0 & (H(X) \neq E) \end{cases}$$

ただし W_E は $H(X) = E$ を満たす状態 X の数とする。として、

$$\underbrace{f^*}_{\text{平衡状態の物理量}} = \mathbb{E}_{MC}[f(\hat{x})] = \sum_x p_{MC}(x|E) f(x)$$

コメントとして、等重率の原理の書き方は教科書によって異なる。(古い教科書だとエルゴード仮説から出発する事が多いかも)。多くの場合、以下の 3 パターン。

(1) 等重率の原理を仮定として認める。

(2) エルゴード仮説から出発して等重率の原理を導く。

(3) (量子系で)ETH(Eigenstate Thermalization Hypothesis(固有状態熱化仮説)) を認めて導出

3.2　エルゴード仮説

系がマルコフかつ決定論的なダイナミクスで記述される時、ある初期状態の粒子が描くトラジェクトリ (相空間上の軌跡) はループ構造もしくは空間全てを覆うような軌跡となる (相空間上の体積が 0or1(相空間全体の体積))。よって (ループ構造上に初期状態がなければ) そのトラジェクトリの時間平均を平衡状態における物理量と考えたときに、相空間上の全体での平均値と等しくなるはずである。

$$f^* = \lim_{t\to\infty} \underbrace{\frac{1}{t}\int_0^t dt' f_{t'}}_{=f_t^*}$$

としたとき

$$\frac{1}{t}\int_0^t dt' f_{t'} = \mathbb{E}_{MC}[f(x)]$$

これが、エルゴード仮説からの等重率の原理の説明である。

3.3 ETH(固有状態熱化仮説)

- (強い)ETH
  $\cdots$ 全ての固有状態 $|\psi_n\rangle$ $(H|\psi_n\rangle = E|\psi_n\rangle)$ において $\langle\psi_n|\hat{f}|\psi_n\rangle \simeq \mathbb{E}_{MC}[f(\hat{x})]$ を満たすという仮説。数値計算によって多くの系 (非可積分系) で確認されている。(ややこしいがここの左辺での $\hat{f}$ は演算子の意味でハットをつけている)
- (弱い)ETH
  $\cdots$ ほとんどすべての固有状態 $|\psi_n\rangle$ において $\langle\psi_n|\hat{f}|\psi_n\rangle \simeq \mathbb{E}_{MC}[f(\hat{x})]$ を満たすという仮説。並進対称性を持ち局所的な相互作用のみ持つ量子多体系においては厳密に示されている。

3.4 状態数、状態密度

注目している系のエネルギーがある範囲に収まっているような (例えば $E \leq H(x) \leq E + \delta E$) を満たす相空間上の x の数を状態数と呼ぶ。また、狭い区間の状態数を考えて、その幅で割ったもの (つまりあるエネルギー面直上の状態の数) を状態密度と呼ぶ事にする。

例えば、$0 \leq H(X) \leq E$ を満たす状態数を Ω_E とすると、$H(X) = E$ の面上の状態の数 (状態密度)W_E は、以下のように計算できる。

$$\Omega_{E+\delta E} - \Omega_E = \underbrace{W_E}_{\text{状態密度}} \delta E$$

例：1 辺の長さ L の立方体に閉じ込められた 1 個の自由粒子 (粒子間に相互作用なし) を考える。1 つの粒子の状態は $n = \{n_x, n_y, n_z\}$ で表現でき、固有エネルギー E_n が $E_n \leq E$ を満たす場合、

$$E_n = \underbrace{\frac{\pi^2\hbar^2}{2mL^2}}_{E_0}\left(n_x^2 + n_y^2 + n_z^2\right) \leq E \tag{38}$$

$$\Leftrightarrow \underbrace{n_x^2 + n_y^2 + n_z^2}_{=n^2} \leq \underbrace{\frac{E}{E_0}}_{=r^2}$$

$\Omega_E =$ (制約条件を満たす領域の体積) + (条件の等号が成立する表面でのズレ) となるが、$r = E/E_0$ が十分大きい場合を考えると、第一項は球の体積 (r^3) に比例するのに対し、

第二項は高々表面積 (r^2) のオーダーである。よって、r が十分大きい時に、最も大きなオーダーのみを考える場合は、第二項は無視できる。3 次元なら半径 r の球の体積は $\frac{3}{4}\pi r^3$ かつ、$n_x, n_y, n_z = 1, 2, \ldots$ であり、第一象限のみを考えるので、$\Omega_E \simeq \frac{1}{8} \times \frac{3}{4}\pi(\frac{E}{E_0})^{3/2}$

$$
\begin{aligned}
W_E \delta E &= \Omega_{E+\delta E} - \Omega_E \\
&= \frac{1}{6}\pi\Big(\frac{E}{E_0}\Big)^{\frac{3}{2}}\Big[\Big(1+\frac{\delta E}{E}\Big)^{\frac{3}{2}} - 1\Big] \\
&\simeq \frac{1}{6}\pi\Big(\frac{E}{E_0}\Big)^{\frac{3}{2}}\frac{3}{2}\frac{\delta E}{E} \\
&= \frac{1}{4}\pi\Big(\frac{E}{E_0}\Big)^{\frac{3}{2}}\frac{\delta E}{E}
\end{aligned}
$$

さらに N 個の粒子が、上記の箱の中にある場合を考えてみよう。の状態は $n = \{n_i\}_{i=1}^N$ で指定。$(n_i = \{n_i^x, n_i^y, n_i^z\})$ $E_n \leq E$ を満たす状態の数 Ω_E を考える。つまり、

$$
E_0 \sum_i \sum_{\alpha=x,y,z} n_i^{\alpha\,2} \leq E \iff \sum_i \sum_{\alpha} n_i^{\alpha\,2} \leq \frac{E}{E_0}
$$

N 次元空間の球の体積が $2\pi^{N/2}/N\Gamma(N/2)r^N$ で与えられる事を用いると、この場合も第一象限の体積を考えるので、

$$
\Omega_E \simeq \frac{1}{2^{3N}}\frac{2\pi^{3N/2}}{3N\Gamma(3N/2)}\Big(\frac{E}{E_0}\Big)^{\frac{3N}{2}} \tag{40}
$$

状態密度は

$$
\begin{aligned}
W_E \delta E &= \Omega_{E+\delta E} - \Omega_E \\
&\simeq \frac{1}{2^{3N}}\frac{2\pi^{3N/2}}{3N\Gamma(3N/2)}\Big(\frac{E}{E_0}\Big)^{\frac{3N}{2}}\Big[\Big(1+\frac{\delta E}{E}\Big)^{\frac{3N}{2}} - 1\Big] \\
&\simeq \frac{1}{2^{3N}}\frac{\pi^{3N/2}}{\Gamma(3N/2)}\Big(\frac{E}{E_0}\Big)^{\frac{3N}{2}}\frac{\delta E}{E}
\end{aligned}
$$