

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

$$\Delta U = \Delta W_{in} + \Delta Q_{in}, \quad (1)$$

$$\Delta Q_{in} = T\Delta S \quad (2)$$

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

$$S(U, V, N) = n S(U/n, V/n, N/n). \quad (3)$$

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

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N). \quad (4)$$

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

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

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)

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

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

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

と定義する。これにより $S(U, V, N)$ の関数形が分かっているならば $U(S, V, N)$ 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

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

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、

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

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。他にもものに確率・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

$$g(w) = \max_x [wx - f(x)] \quad (8)$$

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)

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

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。

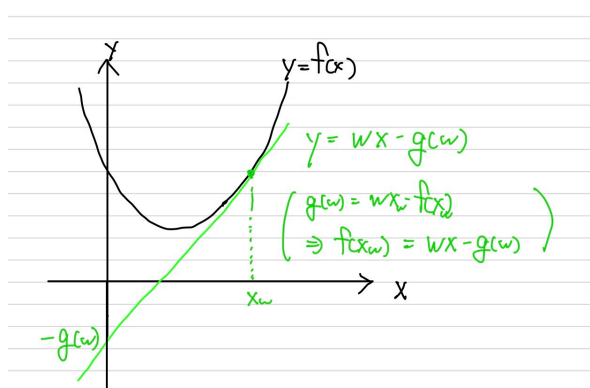


図1 ルジャンドル変換のイメージ

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。

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

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

$f(x)$ 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、

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

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

$$\vec{v}^T \hat{M} \vec{v} \geq 0 \quad (13)$$

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。)

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

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

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

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

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

$$\vec{v}_1^T H[U] \vec{v}_1 = \left(\frac{\partial S}{\partial U} \right)^2 \frac{\partial^2 U}{\partial S^2} = -\frac{\partial U}{\partial S} \frac{\partial^2 S}{\partial U^2} \geq 0 \quad \left(\frac{\partial U}{\partial S} > 0, \frac{\partial^2 S}{\partial U^2} \leq 0 \text{ より} \right) \quad (18)$$

$$\vec{v}_2^T H[U] \vec{v}_2 = -\frac{\partial U}{\partial S} \frac{\partial^2 S}{\partial X^2} \geq 0 \quad \left(\text{途中式変形は省略した} \right) \quad (19)$$

\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

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

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。

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

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

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

$$S_2 = S - S_1, \quad V_2 = V - V_1, \quad N_2 = N - N_1 \quad (23)$$

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

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

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

$$U_{\text{tot}} = U_1 + U_2 + U_B$$

$$= U_1(S_1, V_1, N_1) + U_2(S_2, V_2, N_2) + U_B(S - S_1 - S_2, V - V_1 - V_2, N - N_1 - N_2) \quad (25)$$

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

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

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

F : 自由エネルギー

$$U_{\text{tot}} - U_B = F_1 + F_2 \quad (27)$$

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。

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

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

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。

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

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。)

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

同時確率と条件付き確率の間に以下の関係が成り立つ。

$$p(X, Y) = p(X|Y) p(Y) = p(Y|X) p(X) \quad (28)$$

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

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

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

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

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

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

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

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、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。

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

分散はどうだろうか？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

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

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

$$\begin{aligned}\Pr[\text{条件式}] &= \sum_j \theta(x_j) p_j \left(= \int dx \theta(x) p(x) \right) \\ \theta(x_j) &= \begin{cases} 1 & (x_j \text{ が条件式を満たす}) \\ 0 & (\text{そうでない場合}) \end{cases}\end{aligned}$$

Chebyshev の不等式は以下で表される。

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

証明：

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

ここで、 $\theta(x)$ は以下の定義を満たす。

$$\theta(x) = \theta^2(x) \quad (31)$$

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

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

よって

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

まとめて

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

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

$$\Pr\left[|\hat{m} - \mu| \geq \varepsilon\right] \leq \frac{\sigma^2}{N\varepsilon^2} \quad (35)$$

より、標本平均が真の期待値から ε 以上ズレる確率は $N \rightarrow \infty$ 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、しばしば苦勞を要する。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。

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

また n 次のモーメントを以下で定義する。

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

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。

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

また n 次のキュムラントを以下で定義する。

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

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)

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

また物理においては $-i\theta \rightarrow -\theta$ としてモーメント/キュムラント生成関数を定義する事が多い。(一般の場合には、積分の収束性を担保するために \exp の肩を虚数にするべきだが、物理の問題においては収束性が担保されている事が多いためであろう。)

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

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

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

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

2.6 Kullback–Leibler divergence と Sanov の定理

2.6.1 Kullback–Leibler divergence

ここで言う Divergence とは、擬距離の事である。(発散ではない。) A と B の距離と言われると想像できると思うが、では擬距離とはなんなのか？統計の分野では主に以下の条

件を満たす $D(p|q)$ を擬距離と呼ぶ。

1. $D(p|q) \geq 0$
2. $D(p|q) = 0 \Leftrightarrow p = q$
3. 一般化ピタゴラスの定理を満たす

上の3つの条件は通常の距離も満たすが、擬距離の特徴的な点は、必ずしも $D(p|q) = D(q|p)$ でなくても良い点である。(距離の場合はこれを満たさなければならない。) つまり、点 P から測る点 Q までの擬距離と、点 Q から測る点 P への擬距離が異なっている点も良いのである。

さて確率・統計の分野で (おそらく) 最も用いられる擬距離は、以下で定義される Kullback–Leibler divergence であろう。確率分布 $p(X), q(X)$ (もしくは $\{p_n\}, \{q_n\}$) の間の擬距離として、

$$D_{\text{KL}}(p|q) = \int dX \, p(X) \log[p(X)/q(X)] \quad \left(= \sum_n p_n \log[p_n/q_n] \right) \quad (38)$$

で定義される。これはもちろん上で挙げた擬距離が満たすべき性質を満たす。(練習問題として、離散確率分布の場合について 1, 2 を証明してみよ。(3 は情報幾何学の領域であるのでここでは深く立ち入らない。) また余裕があれば $\{p_n\}$ について上に凸な関数である事も示せ。)

さて、ただいま導入した Kullback–Leibler divergence (以降 KL-divergence と呼ぶ) だが、確率統計の分野で自然に出てくる量であることを次の Sanov の定理によって見てみよう。

2.6.2 Sanov の定理

Sanov の定理 (の簡略版): 真の離散確率分布 $q = \{q_i\}_{i=1}^l$ から i.i.d で十分な量の N サンプルを集め、 i の事象がそれぞれ n_i 回サンプルされ ($\sum_i n_i = N$ を満たす。)、そのサンプルから推定される経験分布 $p = \{p_i\}_{i=1}^l = \{n_i/N\}_{i=1}^l$ が得られる確率 $P(p \sim q)$ は

$$P(p \sim q) \simeq \exp \left[-N D_{\text{KL}}(p|q) \right]$$

で与えられる。

証明：真の確率分布 q から経験分布 p が得られる確率 $P(p \sim q)$ は

$$\log P(p \sim q) = \log \frac{N!}{n_1! n_2! \dots n_l!} q_1^{n_1} q_2^{n_2} \dots q_l^{n_l} \quad (39)$$

$$\simeq N(\log N - 1) + \sum_i \left[-n_i(\log n_i - 1) + n_i \log q_i \right] \quad (40)$$

$$= N \sum_i \left[-\frac{n_i}{N} \log \frac{n_i}{N} + \frac{n_i}{N} \log q_i \right] \quad (41)$$

$$= -N \sum_i p_i \log[p_i/q_i] = -N D_{\text{KL}}(p|q) \quad (42)$$

二行目の式変形は Stirling の公式 ($\log N! = N(\log N - 1)$) を用いた。(Stirling の公式の証明は 3.5 章を参照。) よって Sanov の定理が導かれた。

この Sanov の定理の意味するところは、十分な回数サンプルを取った後にそのサンプルから得られた経験分布が実は真の確率とは異なる確率は、真の確率との KL-divergence によって決まると言うことである。真の確率分布から擬距離の離れた経験分布が生成される確率はその擬距離とサンプル回数の \exp で減衰するため、経験分布の信頼度の保証を与えてくれる。

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、

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

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 0 or 1 (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。

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

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
 ... 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
 ... 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、

$$\begin{aligned} E_n &= \underbrace{\frac{\pi^2 \hbar^2}{2mL^2}}_{E_0} (n_x^2 + n_y^2 + n_z^2) \leq E \\ \Leftrightarrow \frac{n_x^2 + n_y^2 + n_z^2}{=n^2} &\leq \frac{E}{\underbrace{E_0}_{=r^2}} \end{aligned} \quad (43)$$

$\Omega_E = (\text{制約条件を満たす領域の体積}) + (\text{条件の等号が成立する表面でのズレ})$ 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、 $n_x, n_y, n_z = 1, 2, \dots$ であり、第一象限のみを考えるので、 $\Omega_E \simeq \frac{1}{8} \times \frac{3}{4}\pi (\frac{E}{E_0})^{3/2}$

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

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、

$$\Omega_E \simeq \frac{1}{2^{3N}N!} \frac{2\pi^{3N/2}}{3N\Gamma(3N/2)} \left(\frac{E}{E_0}\right)^{\frac{3N}{2}} \quad (45)$$

と書ける。ただし N 個の粒子はそれぞれ区別しないとして $1/N!$ のファクターをかけた。状態密度は

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

3.5 統計力学におけるエントロピー

統計力学におけるエントロピーの表式を考えてみよう。簡単のためエントロピーが確率分布関数 p の汎関数として書ける場合を考える。すなわち $S(p) = \int dX p(X) f(p(X))$ と書ける場合に、 f の関数系を同定したい。エントロピーの持つべき性質として、相加性を課すと、独立した系 1,2 に対して

$$S_{12}(p_{12}) = S_1(p_1) + S_2(p_2) \quad (46)$$

$$p_{12}(X, Y) = p_1(X) p_2(Y) \quad (47)$$

が満たされるべきである。

$$S_{12}(p_{12}) = \int dX \int dY p_{12}(X, Y) f(p_{12}(X, Y)) \quad (48)$$

$$= \int dX \int dY p_1(X) p_2(Y) f(p_1(X) * p_2(Y)) \quad (49)$$

また

$$S_1(p_1) + S_2(p_2) = \int dX p_1(X) f(p_1(X)) + \int dY p_2(Y) f(p_2(Y)) \quad (50)$$

$$= \int dX \int dY p_1(X) p_2(Y) \left(f(p_1(X)) + f(p_2(Y)) \right) \quad (51)$$

よって、式 (46) が任意の確率分布 $p_1(X), p_2(Y)$ について成立する条件は、

$$f(p_1(X) * p_2(Y)) = f(p_1(X)) + f(p_2(Y)) \quad (52)$$

となる。これが任意の確率分布 $p_1(X), p_2(Y)$ について成立するので、両辺を $p_1(X)$ で微分し、 $p_1(X) = 1$ を代入すると

$$p_2(Y) \frac{\partial f}{\partial p}(p_2(Y)) = \frac{\partial f}{\partial p}(1) = a \quad (53)$$

ただし $a = \frac{\partial f}{\partial p}(1)$ を満たす定数である。よって $p_2(Y)$ を右辺に移して両辺積分して

$$f(p(X)) = a \log[p(X)] + C \quad (54)$$

さらに式 (46) に代入すると $C = 0$ となることがわかる。ちなみに熱力学の経験的な知識と極限で一致させる事を考えると、 $a = -k_B$ (k_B はボルツマン定数) となることが分かり、統計力学のエントロピーの表式は、

$$S = -k_B \int dX p(X) \log[p(X)] \quad (55)$$

となる。また $-\log[p(X)]$ の事を確率・統計の分野では「情報量」と呼び、 $-\int dX p(X) \log[p(X)]$ の事を「Shannon エントロピー」と呼ぶ。

さらに、全系のエネルギーの値がわかっている場合に、ミクロカノニカル分布を式 (55) に代入すると

$$S = k_B \log W_E \quad (56)$$

が得られる。これがボルツマンの原理である。また式 (55) で表されるエントロピーを、 $H(X) = E$ の条件下の下で最大化すると、ミクロカノニカル分布が得られる。すなわち等重率の原理とは、熱力学における「エントロピー最大の原理」の言い換えになっている。

さて式 (55) の表式を用いて、熱力学での結果を再現してみよう。先ほど考えた 1 辺の長さ L の立方体に閉じ込められた N 個の粒子を考える。ボルツマンの原理より、

$$\frac{S}{k_B} = \log W_E \quad (57)$$

$$\simeq \frac{3N}{2} \left\{ \log \left[\frac{m}{2\pi\hbar^2} E \right] - \left(\log \frac{3N}{2} - 1 \right) \right\} - N(\log N - 1) + N \log V \quad (58)$$

$$\simeq \frac{3N}{2} \log \left[\frac{m}{2\pi\hbar^2} \frac{E}{N} \right] + N \log \frac{V}{N} \quad (59)$$

と書ける。ここで $N \gg 1$ としてスターリング (Stirling) の公式 $\log N! \simeq N(\log N - 1)$ を用いた。熱力学における定義に従えば $\partial S / \partial E = 1/T$ によって温度が定義される。よって

$$\frac{1}{k_B} \frac{\partial S}{\partial E} = \frac{1}{k_B T} = \frac{3N}{2E} \Rightarrow E = \frac{3}{2} N k_B T \quad (60)$$

となり、熱力学で習った単原子の運動エネルギーと温度の関係式が得られる。

Stirling の公式の証明:

$$\begin{aligned}
 \log N! &= \sum_{j=1}^N \log j \\
 &= N \log N + N \sum_{j=1}^N \frac{1}{N} \log \frac{j}{N} \\
 &\simeq N \left(\log N + \int_0^1 dx \log x \right) = N (\log N - 1)
 \end{aligned}$$

$x = e^{-t}$ と変数変換すれば $\int_0^1 dx \log x = -\int_0^\infty dt e^{-t} t = -1$ となるのを利用した

3.6 希薄極限における分布

ミクロカノニカル分布は、全系のエネルギーを固定しそれを制約条件とした上で、エントロピーを最大化する事で求める事が出来た。ここでは、希薄極限を取った際にミクロなエネルギー状態からこの分布を表現するとどのように表されるかを見ていきたい。

ある箱の中の N 粒子の気体原子を考える。1つの粒子が取りうるエネルギー準位が L 個あり、準位 l のエネルギーを E_l とし、準位 l の状態が M_l 個あり、準位 l をいる粒子が N_l 個とする。系全体のエネルギーの和を E とすると、以下の条件を満たす。

$$\sum_{l=1}^L N_l = N, \quad \sum_{l=1}^L N_l E_l = E \quad (61)$$

全粒子の状態は $\{N_l\}_{l=1}^L$ で特定できるので、その状態をとる確率を $p(\{N_l\}_{l=1}^L) = p_n$ とすると、エントロピー $S = -k_B \sum_n p_n \log p_n$ を条件式 (61) の下で最大化するように分布が決定されるとして分布を求めてみよう。これは制約条件下での最適化問題であり、解法として Lagrange の未定乗数法があるので紹介したい。

3.6.1 Lagrange の未定乗数法

制約条件: $\vec{g}(\vec{X}) = 0$ の下で、下に凸 (/上に凸) な関数 $f(\vec{X})$ を最小化 (/最大化) する \vec{X} を同定する問題を考える。

Lagrange の未定乗数法 (等式条件の場合):

$$\tilde{f}(\vec{X}) = f(\vec{X}) - \vec{c}^T \cdot \vec{g}(\vec{X})$$

で定義される Lagrange 関数 \tilde{f} について

$$\frac{\partial \tilde{f}}{\partial X_i} = 0, \quad \frac{\partial \tilde{f}}{\partial c_i} = 0 \quad (62)$$

を解いて得られた \vec{X}^* は、制約条件: $\vec{g}(\vec{X}) = 0$ の下で、下に凸 (/上に凸) な関数 $f(\vec{X})$ を最小化 (/最大化) する \vec{X} に等しい。

(例 1) 使わないで解く場合：下に凸な関数 $f(x, y)$ を 1 つの制約条件 $g(x, y) = 0$ の下で最小化する問題を考える。制約条件があるので、 x, y は必ずしも独立ではない ($dy/dx \neq 0$) とすると、

$$\frac{dg}{dx} = \frac{\partial g}{\partial x} + \frac{dy}{dx} \frac{\partial g}{\partial y} = 0 \quad (63)$$

$$\left. \frac{df}{dx} \right|_{(x^*, y^*)} = \left. \frac{\partial f}{\partial x} \right|_{(x^*, y^*)} + \left. \frac{dy}{dx} \frac{\partial f}{\partial y} \right|_{(x^*, y^*)} = 0 \quad (64)$$

を満たすような x, y を探せば良い。(1 つ目の式は制約条件から)

2 つ目の式に $\frac{\partial g}{\partial y}$ をかけると

$$0 = \left. \frac{\partial g}{\partial y} \frac{\partial f}{\partial x} \right|_{(x^*, y^*)} + \left. \frac{dy}{dx} \frac{\partial g}{\partial y} \frac{\partial f}{\partial y} \right|_{(x^*, y^*)} \quad (65)$$

$$= \left. \frac{\partial g}{\partial y} \frac{\partial f}{\partial x} \right|_{(x^*, y^*)} - \left. \frac{\partial g}{\partial x} \frac{\partial f}{\partial y} \right|_{(x^*, y^*)} \quad (66)$$

$$\Rightarrow \left. \frac{\partial f}{\partial x} / \frac{\partial g}{\partial x} \right|_{(x^*, y^*)} = \left. \frac{\partial f}{\partial y} / \frac{\partial g}{\partial y} \right|_{(x^*, y^*)} = c(\text{const.}) \quad (67)$$

よって求めるべき (x^*, y^*) については以下の式が成立する。

$$\frac{\partial}{\partial x}(f - cg) = \frac{\partial f}{\partial x} - c \frac{\partial g}{\partial x} = 0 \quad (68)$$

$$\frac{\partial}{\partial y}(f - cg) = \frac{\partial f}{\partial y} - c \frac{\partial g}{\partial y} = 0 \quad (69)$$

これはラグランジュ関数の偏微分から得られる式に他ならない。

3.6.2 希薄極限における分布

step1: 希薄条件 (多体極限)

希薄極限 $1 \ll N_l \ll M_l$ が各準位 l に対して成立するとする。この時、準位 l において、

同じ状態に2つ以上粒子がいる状態の数は、同じ状態に高々1つの粒子がいる状態の数に対して $N_l/M_l \ll 1$ のオーダーになるので、これを無視することにとすると、

$$\log W_l \simeq \log \frac{M_l^{N_l}}{N_l!} \simeq N_l \log M_l - N_l(\log N_l - 1) = N_l(\log \frac{M_l}{N_l} + 1) \quad (70)$$

$$S \simeq k_B \sum_l \left(N_l \left(\log \frac{M_l}{N_l} + 1 \right) \right) \quad (71)$$

と書ける。このエントロピーを制約条件の式 (61) の下で最大化する事を考える

step2: 拘束条件の考慮

制約条件の式 (61) に対応する未定乗数をそれぞれ λ, ν とすると、Lagrange 関数は

$$\tilde{S}(\{N_l\}, \lambda, \nu) = S(\{N_l\}) + \lambda(N - \sum_l N_l) + \nu(E - \sum_l E_l N_l) \quad (72)$$

であり、最大化する $\{N_l\}$ は

$$0 = \frac{\partial \tilde{S}}{\partial N_l} = k_B \log \frac{M_l}{N_l} - \lambda - \nu E_l \quad \Leftrightarrow \quad \frac{M_l}{N_l} = \exp[\lambda + \nu E_l] \quad (73)$$

$$0 = \frac{\partial \tilde{S}}{\partial \lambda} = N - \sum_l N_l \quad \Leftrightarrow \quad \sum_l N_l = \sum_l M_l e^{-\lambda - \nu E_l} = N \quad (74)$$

$$0 = \frac{\partial \tilde{S}}{\partial \nu} = E - \sum_l E_l N_l \quad \Leftrightarrow \quad \sum_l N_l E_l = \sum_l M_l E_l e^{-\lambda - \nu E_l} = E \quad (75)$$

を満たす。式 (73) より

$$S(\{N_l\})/k_B = \sum_l \left(M_l e^{-\lambda - \nu E_l} \right) (\lambda + \nu E_l + 1) = (\lambda + 1)N + \nu E \quad (76)$$

$$\Rightarrow \frac{1}{T} = \frac{\partial S}{\partial E} = k_B \nu \quad \Leftrightarrow \quad \nu = \frac{1}{k_B T} = \beta \quad (77)$$

β を逆温度と呼び、統計力学ではよく用いる。また式 (74) より、

$$e^\lambda = \frac{1}{N} \sum_l M_l e^{-\beta E_l} \quad (78)$$

$$\Rightarrow \frac{N_l}{N} = \frac{M_l \exp[-\beta E_l]}{\sum_l M_l \exp[-\beta E_l]} \quad (79)$$

式 (79) で与えられる分布はカノニカル分布と呼ばれる分布である。

4 カノニカル分布

4.1 問題設定

前節までで、孤立 (多体) 系において等重率の原理を仮定し、そこから得られる帰結と熱力学との整合性を確認した。しかし一方で、現実的な話として

1. この世に観測できる孤立系は存在しない (観測する瞬間には外部と相互作用する)
2. マクロな物理量を測定する場合、系全体での測定は行えず、系の一部から系全体の状態を推察

であるので、特に (2) について、もう少し現実的かつ理論的に扱いやすい問題設定を考えるべきである。よってこれ以降は以下の問題設定を考える。注目する系 (System) に熱浴 (Bath) が接している系を考え、以下を仮定する。

1. system+Bath の全系は孤立系
2. system と Bath の間には微小な相互作用があり、エネルギーのやり取りが可能
3. 自由度 (相空間の次元) について $d_B \gg d_s$ ($E_B \gg E_s$)

4.2 導出

全系・注目系・熱浴を表す添字をそれぞれ T, S, B とする。全系の状態は

$$\vec{X}_T = (\vec{X}_S, \vec{X}_B) \quad (80)$$

$$(81)$$

のように、システムの状態と熱浴の状態で指定できる。また全系のエネルギーは

$$E_T = E_S + E_B + E_{int} \simeq E_S + E_B \quad (82)$$

と書ける。ここで2つ目の問題より第3項は無視した。全系は孤立系であるから、ミクロカノニカル分布でかけるとして、

$$p(\vec{X}_T|E_T) = p(\vec{X}_S, \vec{X}_T|E_T) = 1/W_T(E_T) \quad (83)$$

ここで $W_T(E_T)$ はエネルギー E_T における全系の状態数である。注目系と熱浴の相関がない場合、これらは以下を満たす。

$$W_T(E_T) = \sum_{E_S} W_S(E_S) * W_B(E_T - E_S) = \sum_{\vec{X}_S} W_B(E_T - E_S(\vec{X}_S)), \quad (84)$$

$$p(\vec{X}_T|E_T) = p(\vec{X}_S, \vec{X}_T|E_T) = 1 / \left(\sum_{\vec{X}_S} W_B(E_T - E_S(\vec{X}_S)) \right), \quad (85)$$

$$\Rightarrow p(\vec{X}_S|E_T) = \sum_{\vec{X}_B} p(\vec{X}_S, \vec{X}_T|E_T) \quad (86)$$

$$= W_B(E_T - E_S(\vec{X}_S)) / \left(\sum_{\vec{X}_S} W_B(E_T - E_S(\vec{X}_S)) \right) \quad (87)$$

ここで熱浴がミクロカノニカルである時に正しいエントロピー $\tilde{S}_B(E) = k_B \log W_B(E)$ を考えると

$$\tilde{S}_B(E_B) = \tilde{S}_B(E_T - E_S) \quad (88)$$

$$\simeq \tilde{S}_B(E_T) - \left. \frac{\partial \tilde{S}_B}{\partial E} \right|_{E_T} E_S = \tilde{S}_B(E_T) - \frac{E_S}{T_B} \quad (89)$$

二行目において、 $E_B = E_T$ のとき $E_S = 0$ 注目系に何もないような状況であるので、熱浴そのものが全体系であり孤立系になっているので、この極限では上記の \tilde{S} は真のエントロピーであり、このエネルギー微分は温度の逆数となる事を用いた。

よって

$$p(\vec{X}_S|E_T) \propto W_B(E_T - E_S(\vec{X}_S)) = \exp \left[\frac{\tilde{S}_B(E_T)}{k_B} - \frac{E_S(\vec{X}_S)}{k_B T_B} \right] \quad (90)$$

$$\propto \exp \left[-\frac{E_S(\vec{X}_S)}{k_B T_B} \right] \quad (91)$$

$$p(\vec{X}_S) = \frac{1}{Z} \exp \left[-\frac{E_S(\vec{X}_S)}{k_B T_B} \right], \quad (92)$$

$$Z = \sum_{\vec{X}_S} \exp \left[-\frac{E_S(\vec{X}_S)}{k_B T_B} \right] = \sum_{E_S} W_S(E_S) \exp \left[-\frac{E_S}{k_B T_B} \right] \quad (93)$$

と書ける。最後の行で書き表される確率分布をカノニカル分布と呼び、 Z を分配関数と呼ぶ。(この時点では Z はただの規格化定数のように見えるが、後で分配関数と呼ばれる所いがわかるはず。) また温度 T_B は熱浴の温度であるが、並行状態においては孤立系 (全系) 内部の温度は均等であるはずなので $T_S = T_B = T$ であり、以降は T と書く。ここ

まででは、状態を指定した場合の分布関数を書いたが、注目系のエネルギーを指定した分布を考えてもよく、その場合

$$p(E_S) = \frac{1}{Z} W_S(E_S) \exp\left[-\frac{E_S}{k_B T}\right] \quad (94)$$

$$= \frac{1}{Z} \exp\left[-\frac{(E_S - T\tilde{S}_S(E_S))}{k_B T}\right] \quad (95)$$

\tilde{S} はミクロカノニカルを仮定した際のエントロピーの値である。 $f(E_S) = E_S - T\tilde{S}_S(E_S)$ とすると、「自由エネルギーのような関数」でミクロカノニカル分布を表現する事が出来る。

一方で真に熱力学と対応させるべきはマクロな値であり、統計力学における期待値である。カノニカル分布からエントロピーを計算すると

$$S(p) = -k_B \sum_{\vec{X}_S} p(\vec{X}_S) \log[p(\vec{X}_S)] \quad (96)$$

$$= -k_B \sum_{\vec{X}_S} p(\vec{X}_S) \left(-\frac{E_S(\vec{X}_S)}{k_B T} - \log Z \right) \quad (97)$$

$$= \frac{\langle E \rangle}{T} + k_B \log Z \quad (98)$$

$$\Leftrightarrow F = \langle E \rangle - ST = -k_B T \log Z \quad (99)$$

であり、最後の行が熱力学における自由エネルギーの統計力学の対応物である。

4.3 キュムラント生成関数と分配関数による物理量の計算

前節で定義した分配関数 Z による物理量の期待値の計算方法を考えてみよう。形を見て思い出した人もいるかもしれないが、分配関数の \log をとったものはある種のキュムラント生成関数として見る事が出来る。まずキュムラント生成関数について思い出してみよう。ある分布関数 $p(X)$ に対して、物理量 $f(X)$ のキュムラント生成関数は

$$C_f(\theta) = \log\left[\sum_X p(X) \exp[-\theta f(X)]\right] \quad (100)$$

$$C_f^{(n)} = \left(\frac{\partial}{\partial \theta}\right)^n C_f(\theta) \Big|_{\theta=0} \quad (101)$$

のように定義し、 n 次のキュムラントを計算する事が出来る。(物理で使う方のキュムラント生成関数 (\exp の肩が $-\theta$) を使っている。)

ここで1次のキュムラントは分布 $p(X)$ における $f(X)$ の期待値になっている。ところで、二行目の式の $|\theta=0$ を取らずに、キュムラントを定義する事も出来る。すなわち

$$C_f^{(n)}(\theta) = \left(\frac{\partial}{-\partial\theta} \right)^n C_f(\theta) \quad (102)$$

としてもよく、この場合のキュムラントは新たに定義される分布関数

$$p(X; \theta) = p(X) \exp[-\theta f(X)] / \sum_X \left(p(X) \exp[-\theta f(X)] \right) \quad (103)$$

によって期待値を取ることになる。(これは確かに分布関数として満たすべき性質 $\sum_X p(X; \theta) = 1, p(X; \theta) \geq 0$ を満たす。) これはモーメント生成関数には出来ない事であり、「log をとっているために分布関数をスケール倍しても、微分をした際に分母にそのスケールが (分子と) 同数現れるのでキャンセルし、新たな分布関数で期待値を表す事が出来る」ことに由来する。(練習問題によって確かめてみよ)

さて話を分配関数に戻そう。今カノニカル分布を考えると、分配関数は

$$Z = \sum_X \exp[-\beta E(X)] \quad (104)$$

のように書いており、 $\log Z$ をキュムラント生成関数として見るならば $p(X)$ は状態に依らず一定の確率を返す確率分布で、 $p(X; \theta)$ に相当するものがカノニカル分布 (β が θ に相当) である。よって例えば1次のキュムラントを計算するならば、

$$C_E^{(1)}(\beta) = \langle E \rangle_{p(X; \beta)} = -\frac{\partial}{\partial\beta} \log Z \quad (105)$$

となり、これによってカノニカル分布における (注目系の) エネルギーの期待値が計算できる。さらに話を進めてみよう。例えばエネルギーを与える関数の中身が $E(X) = E_0(X) + \theta f(X)$ のような形で書けているとしよう。今カノニカル分布における物理量 $f(X)$ の期待値が知りたい時、 $p(X)$ は $E_0(X)$ によってエネルギーが決定される場合のカノニカル分布として、 $p(X; \theta)$ に相当するものが真の (エネルギーが $E(X)$ で表現される) カノニカル分布 ($\beta\theta$ が θ に相当) であるとしてみなせば良い。よって

$$C_f^{(1)}(\beta\theta) = \langle f \rangle_{p_C} = -\frac{\partial}{\partial\theta} \log Z = \frac{\partial}{\partial\theta} F \quad (106)$$

のように、これも分配関数 (自由エネルギー) によって計算する事が出来る。このようにして、気になる物理量を全部何かしらのパラメータ (外場の強さである事が多い) と結合させてエネルギーの関数の中に入れておいて分配関数を一度計算しておけば、log をとっ

てパラメータで微分するだけでその期待値を計算する事が出来る。ので物理の人間は「分配関数が計算できる」＝「その系の平衡系の物理を (ほぼ) 全て理解できる」と思っている (はず)。

5 具体的な物理量の計算

5.1 ゴム弾性

高分子化合物のようなものを考えて、原子が 1 次元的に連なっていることを考える。また、その取れる状態は連なって直線に伸びている/折りたたまれている状態の 2 つとする。さらに、エネルギーは状態に依らないとする。(簡単のため $E(X) = 0$ としておく) ここでは右に伸びている状態を \rightarrow , 左に伸びている状態を \leftarrow として以降記述する。以下の問題設定を考える

- 高分子が全部で N 個

$$\rightarrow \text{の数} : N_+, \quad \leftarrow \text{の数} : N_-, \quad N_- + N_+ = N$$

- 高分子の長さ $\ell \Rightarrow$ 全体の長さ $L = \ell(N_+ - N_-)$

$$\begin{aligned} N_+ - N_- &= \frac{L}{\ell}, \quad N_+ + N_- = N \\ \Rightarrow N_+ &= \frac{1}{2}\left(N + \frac{L}{\ell}\right) = \frac{1}{2}(N + x), \quad N_- = \frac{1}{2}\left(N - \frac{L}{\ell}\right) = \frac{1}{2}(N - x) \end{aligned}$$

ただし $x = L/\ell$ とした。系の長さ (始点と終点の距離。右向きを正とする。今回の定義では負の値を取ることもある。) を L とした時に取りうる状態の数 W_L を考えたい。N 個の高分子を 2 つの状態のうちどちらかを選ぶ場合の数は、N 個のボールを 2 つの箱に分ける時と同じなので、

$$W_L = \frac{N!}{N_+!N_-!}$$

より、Stirling の公式を用いて、

$$\begin{aligned} \log W_L &\simeq N(\log N - 1) - N_+(\log N_+ - 1) - N_-(\log N_- - 1) \\ &= N \log N - \frac{1}{2}(N + x) \log \frac{N + x}{2} - \frac{1}{2}(N - x) \log \frac{N - x}{2} \end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial L} \log W_L &= \frac{1}{\ell} \frac{\partial}{\partial x} \log W_L \\
&= \frac{1}{\ell} \left(-\frac{1}{2} \log \frac{N+x}{2} - \frac{1}{2} + \frac{1}{2} \log \frac{N-x}{2} + \frac{1}{2} \right) \\
&= \frac{1}{2\ell} \log \frac{N-x}{N+x} = -\frac{1}{2\ell} \frac{2x}{N} = -\frac{L}{N\ell^2}
\end{aligned}$$

さらに、

$$\begin{aligned}
F(L) &= -k_B T \log W_L \\
p &= \frac{\partial F}{\partial L} = -k_B T \frac{\partial}{\partial L} \log W_L = \frac{1}{\beta} \frac{L}{N\ell^2} \\
F(L, T) - F(0, T) &= \frac{1}{2} \left(\frac{1}{\beta N \ell^2} \right) L^2
\end{aligned}$$

よってバネ定数 $k = 1/(\beta N \ell^2)$ が得られた。

5.2 磁性

5.2.1 Ising 模型の導入

統計力学の適用先として、よく挙げられるのは固体物性である。(おそらく理由は 1. 実用上の重要性, 2. 物質相の多彩さ, 3. ミクロな描像から始めて比較的簡単な模型を使える, あたりだと思う。) 固体中には大まかにはアボガドロ数個のオーダー ($O(10^{23})$) の原子が集まって、結晶格子を組んでいる。これは容易に大数の法則を満たし、統計力学を適用可能である。

さて、一般に固体の電磁気的な性質は、ほとんど全て物質中の電子が決定している。電子は陽の電荷を持つ原子核による静電ポテンシャルによって、原子核の周りにいる方が低いエネルギー状態であり、カノニカル分布を考えれば、そのような状態が実現しやすい。ここで思い切った近似として、電子はほぼ原子核の周辺にしかいないとして、 \vec{X} での電子の位置の指定から、原子核にインデックス i を割り振り、近くの原子核のインデックスによる指定に切り替えよう。

さらに原子核の”付近”とは、十分小さな領域であり、そこに電子が2つ存在すれば、お互いに働くクーロン反発によって、エネルギーが非常に大きな状態となり、カノニカル分布を考えれば、ほとんどそういった状態が起こらないという仮定をしよう。そのようにすれば、各サイト i には高々電子は1つしか存在せず、さらに注目している電子とサイトの数が同数であれば、全てのサイトに電子が1つずつ存在する状態が安定となる。

この時、あるサイトの電子が隣のサイトに移動してしまうと、隣のサイトでの電子数が2となり、エネルギーの高い状態となってしまう。このため、この状況において電子はサイトからほとんど動けず、サイトに局在しており、ほとんど運動量の情報を失いスピンとしての情報のみを残す。

このような状況を表す有効模型として用いられるのが以下のハミルトニアンで記述される Ising 模型である。

$$\mathcal{H}(\{S_i\}) = -J \sum_{\langle i,j \rangle} S_i S_j - g\mu_B h \sum_i S_i \quad (S_i = \pm 1) \quad (107)$$

を考える。 i, j は結晶の原子一位置のサイトのインデックスであり、 $\langle i, j \rangle$ は隣りあっている場合のみ計算することを表す。第一項は隣り合うスピン間の相互作用であり、第二項は外部磁場とスピンの coupling の項である。これは結晶中の原子の周りの電子の計算などに用いる。

- 実際に物質の有効模型として使って磁性などを解析
- 数理的な模型としての面白さを探究

の2つの文脈で研究されている。物理の面白さとして、人間が物性を解析するために(簡単な模型に落とすために)注目する系の詳細な情報を落としていって抽象化(粗視化)していった先で得られた有効モデルの性質が、元々は全く異なる現象のための有効モデルの性質と全く同じになる事が起こりうる。このように粗視化した先での、同一視されるようなモデル群とその性質をユニバーサリティクラスと呼ぶ。(物理学の「普遍性」の代表例である。)(気になる方は「イジング ユニバーサリティクラス」で検索してみよう。)

以降、統計力学(カノニカル分布)を用いて、このイジング模型を簡単に解析してみよう。

5.2.2 転送行列による解析

5.2.3 平均場近似による解析

5.3 比熱

物性実験のゼミをとる人向け。カノニカル分布からの計算の練習にもちょうど良いかも。試験範囲にはしない。

6 非平衡統計力学

6.1 Markov 過程

6.2 Kramers-Moyal 展開

6.3 Pawula Theory

6.4 Fokker-Planck 方程式と Onsager-Machlup 形式

6.5 Langevin 方程式

6.6 Boltzmann 方程式による輸送応答の計算