

# シャノンの定理を物理のモデルを通して眺める

樺島 祥介

## 1. はじめに

対象をモデル化することの効用の一つは一見つながりのない物事に関連づけができることである。電気の LCR 回路と機械のバネ・マス・ダンパ系はまったく異なるモノであるし、用途も違うのだが同一の微分方程式でモデル化できる。この同一性への気づきによって、1つの微分方程式の性質を調べるだけで同時に2つの事柄に対して有用な知見が得られる。つまり「1粒で2度オイシイ」のである。

こういうことをもっと積極的にやろう、というのがいわゆるシステム科学というものなのであろう。ならば、出来るだけ関係なさそうなものが実は関係している、というオチであるほど面白い。そういった意味で、以下で紹介する内容はかなりイケている部類に入ると思う。

具体的には、アモルファス金属やタンパク質の物性とシャノンの情報理論が同一の数理モデルで関係づけられるという話である。サイバネティクスの提唱以来、システム科学的な視点が科学や工学に本格的に導入されて既に60年ほどが経過している。そのため同一性に関する基本事例はほぼ調べ尽くされていると思われるがちなのだが、ここで述べる話が注目されはじめたのは高々この10年足らずの間である。学問が成熟し様々な分野で閉塞状況を嘆く声を耳にするが、我々が気づいていないことがまだまだ沢山ある、そういったことを示す証拠だと考えたい。

以下、物理のモデルの紹介、それを通して見たシャノンの通信路符号化定理、の順で述べていく。

## 2. ランダムエネルギーモデル

### 2.1 不規則系の統計力学

世の中のすべてのモノは数種類の基本粒子からできている。気体や金属の性質は基本粒子が多数集まることで生まれてくると考えられる。統計的な枠組みによりミクロな基本粒子の性質と気体や金属などマクロな系の振る舞いとをつなげる物理学の理論は統計力学と呼ばれる。

さて、気体や金属は多数の粒子から構成されているものの、粒子どうしの相互作用は比較的単純であり同一の構造が規則的に現れることが多い。統計力学はまずこういった規則的な系を中心に発展してきたが、それらに関する対処法は1970年代くらいにほぼ出揃った。そこで、1980年代からは粒子間相互作用の規則性が乱された不規則系に関する理論にも大きな関心が寄せられている。

具体的には、結晶化していないため相互作用の規則性が失われるアモルファス金属、基本粒子が異なる個性を持っているため相互作用に単純な規則性のないヘテロポリマー（タンパク質など）などを思い浮かべていただきたい。こういった系の特徴を簡潔に表現するためにしばしば行われるのはシステムを凍結されたランダム変数  $\mathbf{R}$  に依存するエネルギー関数  $H(\mathbf{S}|\mathbf{R})$  で特徴づけるモデル化である。ただし、 $\mathbf{S}$  はシステムを記述するミクロな状態変数であり、アモルファス金属なら各原子に付随する電子の状態や磁気モーメントの向き、タンパク質なら各原子の位置を表した変数のようなものである。

### 2.2 ランダムエネルギーモデル

規則性という便利な性質を使えない不規則系の解析は一般に難しい。そこで、個別問題の定量的な解析というよりは、手に負えるくらい単純化したモデルに基づいて生じ得る現象を網羅的に調べる構成的研究がしばしば行われる。ランダムエネルギーモデルはそういった研究の代表的題材である[1]。いくつかのバリエ

かばしま よしゆき  
東京工業大学 大学院総合理工学研究科  
〒226-8502 横浜市緑区長津田町 4259

ーションが提案されているが、ここで紹介するのは次節の内容と直接的につながる離散版のモデルである。

このモデルでは長さ  $K$  の 2 元ベクトル  $\mathbf{S} = \{+1, -1\}^K$  によって表された各状態に対し独立に、そのエネルギーの値  $E(\mathbf{S})$  を同一の 2 項分布

$$P(E) = \binom{N}{\frac{N-E}{2}} \times 2^{-N} \quad (1)$$

に従って、 $-N, -N+2, -N+4, \dots, N-2, N$  のいずれかの値を取るように定める。このように定めた各状態に関するエネルギーの値の組  $\mathbf{E}(\mathbf{S})$  がこの場合の凍結されたランダム変数  $\mathbf{R}$  に対応する。つまり、

$$H(\mathbf{S}|\mathbf{E}) = E(\mathbf{S}) \quad (2)$$

のように、エネルギー関数の値自体をランダム変数によって定めたモデルである。

エネルギー関数が  $H(\mathbf{S}|\mathbf{E})$  によって与えられると(平衡)統計力学では絶対温度  $T = \beta^{-1}$  の熱平衡状態においてミクロな状態  $\mathbf{S}$  が出現する確率はボルツマン分布

$$P(\mathbf{S}|\mathbf{E}) = Z^{-1}(\beta; \mathbf{E}) e^{-\beta H(\mathbf{S}|\mathbf{E})} \quad (3)$$

によって与えられると仮定する。ただし、 $Z(\beta; \mathbf{E}) = \sum_{\mathbf{S}} e^{-\beta H(\mathbf{S}|\mathbf{E})}$  は分配関数と呼ばれる規格化定数である。

統計力学では、ボルツマン分布(3)に関して 1 自由度あたりのエネルギー関数  $H(\mathbf{S}|\mathbf{E})$  の平均値 (内部エネルギー)  $u = K^{-1} \sum_{\mathbf{S}} H(\mathbf{S}|\mathbf{E}) P(\mathbf{S}|\mathbf{E})$  やエントロピー  $s = -K^{-1} \sum_{\mathbf{S}} P(\mathbf{S}|\mathbf{E}) \ln P(\mathbf{S}|\mathbf{E})$  を評価することで温度  $T$  のマクロな性質を議論する。この評価に重要な役割を果たすのが分配関数  $Z(\beta; \mathbf{E})$  である。具体的には、一般に

$$u = -\frac{1}{K} \frac{\partial}{\partial \beta} \ln Z(\beta; \mathbf{E}) \quad (4)$$

$$s = -\frac{1}{K} \beta^2 \frac{\partial}{\partial \beta} (\beta^{-1} \ln Z(\beta; \mathbf{E})) \quad (5)$$

となることが簡単な計算から確かめられる。つまり、 $Z(\beta; \mathbf{E})$  さえ評価できればあとは知りたいことが芋づる式にわかるのである。

### 2.3 統計力学的解析

分配関数の評価は一般に難しく、すべての場面で上の方針に沿った解析が実際的であるわけではない。しかしながら、ランダムエネルギーモデルでは、 $K, N$  がそれらの比  $R = K/N$  を  $O(1)$  に保ちつつ十分大きいとみなせる場合、下記のような評価が可能になる。

場合分けが必要になると説明が面倒なので、以下、 $R < 1$  の場合のみを想定する。説明が増えるだけで  $R$

$> 1$  のときも同様である。まず、分配関数の評価の中で  $H(\mathbf{S}|\mathbf{E}) = Ne$  となる部分の和を先に実行し

$$Z(\beta; \mathbf{E}) = \sum_e e^{-N\beta e} \times W(e|\mathbf{E}) \quad (6)$$

という形に書き換える。ただし、 $W(e|\mathbf{E})$  は凍結されたランダム変数  $\mathbf{E}$  で定まるエネルギー関数  $H(\mathbf{S}|\mathbf{E})$  に対し  $H(\mathbf{S}|\mathbf{E}) = Ne$  となる微視的状态  $\mathbf{S}$  の個数である。これは  $\mathbf{E}$  に依存するランダム変数であるが、 $N, K \rightarrow \infty$  の場合には典型値から任意の許容幅以上に外れる確率が多数の法則によりゼロに収束していく。

$|e| < e_c(R)$  に対する典型値は

$$\begin{aligned} \Sigma(e|\mathbf{E}) &= \frac{1}{N} \ln W(e|\mathbf{E}) \\ &= \ln 2 \times \left( H_2\left(\frac{1+e}{2}\right) + R - 1 \right), \end{aligned} \quad (7)$$

により特徴付けられる。ただし、

$$H_2(x) = -x \log_2 x - (1-x) \log_2 (1-x) \quad (8)$$

は 2 元エントロピー関数であり、 $e_c(R)$  は方程式  $H_2\left(\frac{1+e}{2}\right) + R - 1 = 0$  の解のうち正のものである。一方、 $|e| > e_c(R)$  の領域に対する典型値は  $W(e|\mathbf{E}) = 0$  で与えられる。これらの結果を利用すると、式(6)を式(7)のグラフを用いて評価することが可能となる。

$N$  が十分大きな場合、式(6)の和はそのうちの最大項のみを用いて近似できる。最大項を与えるのは  $-\beta e + \Sigma(e|\mathbf{E})$  を最大化する  $e$  の値、 $e^*$  である。図 1 に示すように、これは  $|e| \leq e_c(R)$  の領域で曲線  $y = \Sigma(e|\mathbf{E})$  に対し傾き  $\beta$  の接線を引くことで求められる。 $\beta$  に応じて接点の位置は

$$e^* = \begin{cases} -\tanh(\beta), & \beta < \beta_c(R) \\ -e_c(R), & \beta > \beta_c(R) \end{cases} \quad (9)$$

のように 2 つの関数形で表現される。ただし、 $\beta_c(R) = \frac{1}{2} \ln \frac{1+e_c(R)}{1-e_c(R)}$  である。これを用いると規格化され

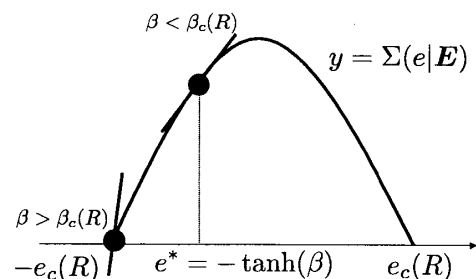


図 1 和(6)での最大項を与える  $e^*$  の決め方。  $|e| \leq e_c(R)$  の範囲で傾き  $\beta$  の直線が関数  $y = \Sigma(e|\mathbf{E})$  に接する点が  $e^*$  となる。このことから、 $\beta < \beta_c(R)$  のとき  $e^* = -\tanh(\beta)$ 、 $\beta > \beta_c(R)$  のとき  $e^* = -e_c(R)$  であることがわかる。

た分配関数の対数  $\frac{1}{K} \ln Z(\beta; \mathbf{E})$  が求まるが、実は式(9)から1自由度あたりの内部エネルギー  $u$ 、エントロピー  $s$  を直接評価することができる。結果は

$$u = R^{-1} e^* \quad (10)$$

$$s = R^{-1} \sum (e^* | \mathbf{E}) \quad (11)$$

である。

#### 2.4 相転移

図2は式(10), (11)により与えられる内部エネルギー、エントロピーを絶対温度  $T = \beta^{-1}$  の関数として表現したグラフを示している。このグラフからわかることは、 $T > T_c(R) = \beta_c^{-1}(R)$  では内部エネルギー、エントロピーともに温度  $T$  の増加関数となっているのに対し、 $T < T_c(R)$  では温度に依存せず一定値にとどまり続けるという振る舞いである。つまり  $T < T_c(R)$  でシステムは“凍りついて”いる。詳細に議論すると異なるカテゴリに分けられるのだが、この振る舞いは広い意味で、水が氷になる、といった制御変数（いまの場合は温度）の変化に対してマクロな性質が不連続に変化する相転移現象の一種だと考えられる。

分布関数(3)は正の温度  $T = \beta^{-1}$  に対して解析的であり、有限の臨界温度  $T = T_c(R)$  で不連続性が生じることはあり得ない。それでもなお、上述のような相転移現象が現れることのタネは熱力学的極限  $N, K \rightarrow \infty$  にある。 $N, K \rightarrow \infty$  での状態数  $W(e | \mathbf{E})$  の典型値の振る舞いが  $|e| < e_c(R)$  と  $|e| > e_c(R)$  でまったく異なる。これがここでの相転移メカニズムの肝である。

エントロピーは与えられた温度  $T$  のマクロな性質にどれだけ多くのミクロな状態が関与しているかの目安を与える。ランダムエネルギーモデルが示す相転移

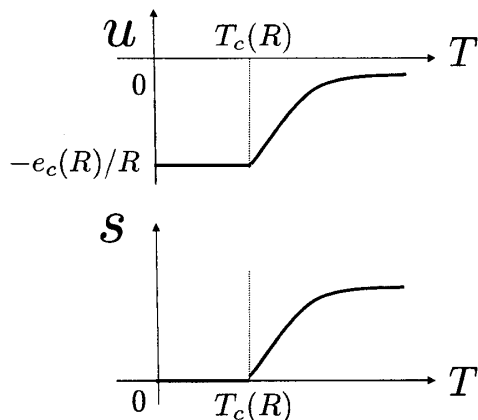


図2 ランダムエネルギーモデルでの1自由度あたりの内部エネルギー  $u$  およびエントロピー  $s$  の温度依存性 (概念図)。  $T < T_c$  でシステムは凍りついてしまう。

では  $T < T_c(R)$  でエントロピーがゼロとなる。このことは、高温相  $T > T_c(R)$  ではシステムは熱運動により沢山の微視的状态を遷移する一方で、低温相  $T < T_c(R)$  では最低エネルギーの微視的状态にトラップされてしまうことを意味している。これは例えば臨界温度以上では比較的自由な形を取るタンパク鎖が臨界温度以下である決まった形に折りたたまれる生体内での現象を彷彿とさせる。ランダムエネルギーモデルは、こういった特徴的な振る舞いがランダムな相互作用の結果、典型的に現れる現象であることを示唆している。

### 3. 通信路符号化定理と相転移

#### 3.1 符号化・劣化・復号

ノイズのある通信路を介して  $K$  ビットのメッセージ  $\mathbf{S} = \{+1, -1\}^K$  を送受信することを考える<sup>1</sup>。  $\mathbf{S}$  をそのまま送るとノイズによって情報が劣化してしまうので  $K$  よりも長い  $N$  ビットの2元ベクトル  $\mathbf{x}(\mathbf{S}) = \{+1, -1\}^N$  に符号化して送信する。  $K$  ビット送れば済むものをそれより長い  $N$  ビットにして送るので無駄(冗長性)がある。しかしながら、この冗長性を活かすことである程度少ない誤りなら元に戻せるようになる。これが通信路符号化、より実際的には誤り訂正符号、の基礎となるアイデアである。

簡単のため、通信路のモデルとしてはアルファベット  $+1, -1$  が各成分独立に確率  $p = (1 - \tanh(\beta_n))/2$  で反転する2元対称通信路(BSC)を仮定する。受信語を  $\mathbf{y} = \{+1, -1\}^N$  とすればBSCは条件付確率

$$P(\mathbf{y} | \mathbf{x}(\mathbf{S})) = \frac{e^{\beta_n \mathbf{y} \cdot \mathbf{x}(\mathbf{S})}}{(2 \cosh(\beta_n))^N} \quad (12)$$

によって表現される。さらに、メッセージ  $\mathbf{S}$  は一様な事前分布  $P(\mathbf{S}) = 2^{-K}$  に従うとしよう。ベイズ決定の一般理論によれば、劣化した受信語  $\mathbf{y}$  を得た際に真のメッセージをもっとも精度良く推定(復号)する方法は事後分布

$$\begin{aligned} P(\mathbf{S} | \mathbf{y}) &= \frac{P(\mathbf{y} | \mathbf{x}(\mathbf{S})) P(\mathbf{S})}{\sum_{\mathbf{S}} P(\mathbf{y} | \mathbf{x}(\mathbf{S})) P(\mathbf{S})} \\ &= Z^{-1}(\beta_n; \mathbf{y}) e^{\beta_n \mathbf{y} \cdot \mathbf{x}(\mathbf{S})} \end{aligned} \quad (13)$$

を用いて構成される。ただし、 $Z(\beta_n; \mathbf{y}) = \sum_{\mathbf{S}} e^{\beta_n \mathbf{y} \cdot \mathbf{x}(\mathbf{S})}$  は分配関数に相当する規格化定数である。

<sup>1</sup> 情報理論では2元アルファベットとして0,1を用いることが多いが、ここでは前節とのつながりを明確にするため  $\pm 1$  を用いる。

### 3.2 ランダム符号とランダムエネルギーモデル

任意の符号化  $x(S)$  に対して事後分布(13)に基づく復号法は必ず最良である。ただし、得られる誤り訂正能力は  $x(S)$  の構成法による。シャノンの通信路符号化定理は  $N, K \rightarrow \infty, R = K/N \sim O(1)$  の状況では各  $S$  に対し  $x(S)$  の各成分を独立かつ確率  $1/2$  で  $\pm 1$  となるようにランダムに決めた符号 (ランダム符号) がほとんどの場合にほぼ最良の誤り訂正能力を達成することを示している[2]。

ランダム符号は前節のランダムエネルギーモデルと深いかわりを持っている[3]。これを示すために受信語  $y$  は真のメッセージ  $S_0$  に対する符号語  $x(S_0)$  が劣化したものと仮定しよう。式(13)はエネルギー関数  $H(S|y) = -y \cdot x(S)$  に関する温度  $T_n = \beta_n^{-1}$  のボルツマン分布であるとみなすことができる。ここで、符号語  $x(S)$  は独立かつランダムに生成されているので  $S \neq S_0$  についてはそれぞれのエネルギーの値  $E(S) = -y \cdot x(S)$  は前節のランダムエネルギーモデルと同様、2項分布(1)に従うと考えてよい。

ただし、真のメッセージ  $S_0$  の場合のみは事情が異なる。なぜなら、 $y$  はもともと真の符号語が劣化したものなので  $x(S_0)$  と高い相関を有するからである。より詳しく言えば、 $S_0$  のエネルギー値  $E(S_0)$  に対してほぼ確率 1 で

$$e_0(\beta_n) = E(S_0)/N = -\tanh(\beta_n) \quad (14)$$

となることが簡単な計算から示される。

つまり、ランダム符号に対する事後分布(13)とは  $2^K - 1$  個の状態で構成されるランダムエネルギーモデルに式(14)で特徴づけられるエネルギー値を持つ1つの状態を付け加えたもの、とみなすことができる。

### 3.3 相転移現象としての通信路符号化定理

上述のランダム符号とランダムエネルギーモデルとのかかわりに気づくことにより、前節と同様、符号の性能をグラフを用いて評価することが可能となる。ノイズの大きさ  $p = (1 - \tanh(\beta_n))/2$  の値に応じて図3(a), (b)の2つの場合があり得る。図3(a)はノイズの値が比較的小さい場合である。このときには事後分布  $P(S|y)$  の中で真のメッセージ  $S_0$  の取る確率が圧倒的に1に近くなる。そのため、ほとんど誤ることなく真のメッセージ  $S_0$  を復号することができる。つまり、ほとんどの場合に対して誤りのない情報伝達が可能になる。一方、図3(b)に対応するノイズが大きい場合には真のメッセージに対するエネルギー値(14)がランダムエネルギーモデルのエネルギー分布の中に埋もれてし

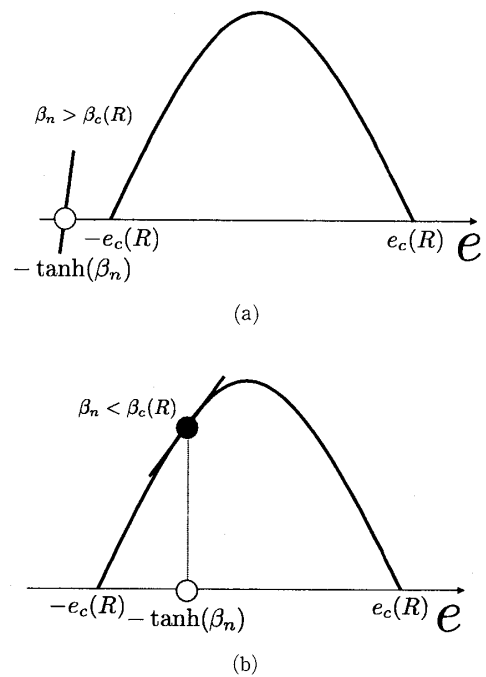


図3 グラフを用いたランダム符号の解析。白丸は正解  $S_0$  のエネルギー値に対応する式(14)の値を示している。(a):ノイズの小さな場合は正解  $S_0$  が事後分布の確率値のほとんどを占める。(b):ノイズの大きな場合は正解  $S_0$  のエネルギー値は連続的なエネルギー分布の中に埋もれてしまう。

まう。これは事後分布(13)の中で  $S_0$  と異なる状態の方が支配的になることを示しており、復号を誤る確率がほぼ1となる。

$p = (1 - \tanh(\beta_n))/2$  の変化に伴う図3(a), (b)間の質的な変化は一種の相転移現象とみなすことができる。相転移が生じる条件を整理すると

$$R = 1 - H_2(p) \quad (15)$$

となる。これは通信路符号化定理が与える BSC のシャノン限界 (符号化による完全な誤り訂正の限界) に他ならない。

## 4. おわりに

アモルファスやタンパク質のモデルとシャノンの定理が関係しているという話を述べた。物理の用語は使わないものの3.3節と本質的に等価な内容はシャノンの定理の荒っぽい説明として昔から情報理論でも知られている。なので、本稿のレベルでは、関係がつかないからといって殊更新しい結果が得られたわけでもない。

にも関わらずこういった関連づけが有用と思われるのは文化の異なる分野との結びつきを通じて、閉じられた分野にありがちな研究スタイルに関する暗黙のシバリや思い込みに風穴を空ける可能性があるからであ

る。実験や観測という手段で対象の真の姿を垣間見ることができない情報科学では数理が研究を支える大きな柱になる。そのため「理論＝証明」という傾向が伝統的に強い。しかしながら、ここに示したような関連があれば、少なくとも、コトの問題をモノの研究スタイルにならって荒っぽく捉えることに躊躇しなくなる。証明スタイルに沿って精密に・厳密に考えると手に負えなかった複雑な問題でも荒く眺めることにより輪郭のレベルでゴロツとした新しい知見が得られる。そういったことがあるかもしれない。実際、高性能な符号のスタンダードとなりつつある低密度パリティ検査符号や第3世代の携帯電話に採用されているCDMA通信方式の基礎研究において統計力学的な解析は既に無視できない存在となっている[4][5]。

物理と情報という大きな分類の根元での同一性の導入はここに示した例にとどまらず、さまざまな問題に大きな影響を与える可能性がある。その可能性を追究

する研究が平成18年度から文科省科研費特定領域研究「情報統計力学の深化と展開」によって推進されている。公開シンポジウム、研究会・講習会や領域ウェブ[6]を通じて研究活動を積極的に公開しているので多くの方に関心を持っていただけると幸いである。

#### 参考文献

- [1] B. Derrida: Phys. Rev. B, 24 (1981), 2613-2626.
- [2] C. E. Shannon: Bell. Syst. Tech. J., 27 (1948), 379-423; *ibid*, 623-656.
- [3] N. Sourlas: Nature, 399 (1989), 693-695.
- [4] Y. Kabashima, T. Murayama and D. Saad: Phys. Rev. Lett., 84 (2000), 1355-1358.
- [5] T. Tanaka: IEEE Trans. Inform. Theory, 11 (2002), 2888-2910.
- [6] 文科省科研費特定領域研究「情報統計力学の深化と展開」(URL) <http://dex-smi.sp.dis.titech.ac.jp/DEX-SMI/>