MaxEnt and the Boltzmann principle

Wednesday, September 7, 2016
SBU: CHE/PHY/BME 558, Physical & Quantitative Biology
Rutgers University: Chemical Thermodynamics
Lecturer: Gábor Balázsi
Recap: Path integrals. Exact differentials
Path integrals: Traveling on landscapes

Walking on a landscape, moving either in the x or y direction. How tired are you?

At each point, your effort can be described by: $X(x, y)dx + Y(x, y)dy$

Total effort, to reach point B from point A: $\int \int_{path \ A \rightarrow \ B} [X(x, y)dx + Y(x, y)dy]$
Path independence and exact differentials

Exact differentials have the following property:

\[ X(x, y)\,dx + Y(x, y)\,dy = \frac{\partial f}{\partial x} \,dx + \frac{\partial f}{\partial y} \,dy = df \]

\[ \int_{\text{path } A \to B} [X(x, y)\,dx + Y(x, y)\,dy] = \int_{\text{path } A \to B} [\frac{\partial f}{\partial x} \,dx + \frac{\partial f}{\partial y} \,dy] = \int_{A \to B} df = f_B - f_A \]

Path integral depends only on the start and end points!
The integral is path-independent.

For \( X(x, y)\,dx + Y(x, y)\,dy \) to be an exact differential, it is required that:

(Euler Test) \[ \frac{\partial X(x, y)}{\partial y} = \frac{\partial Y(x, y)}{\partial x} \]
Example: Clausius Theorem

Consider a cyclical process driven by heat exchange.

\[
\oint_{\text{path } A \rightarrow A'} \left[ \frac{\delta Q_{V,T}(p,V,T)}{T} + \frac{\delta Q_{p,T}(p,V,T)}{T} + \frac{\delta Q_{p,V}(p,V,T)}{T} \right] = \oint_{A \rightarrow A'} \frac{\delta Q}{T} \leq 0
\]

If \( \oint_{A \rightarrow A'} \frac{\delta Q}{T} dpdVdT = 0 \) then \( dS = \frac{\delta Q}{T} \) is an exact differential.

\( S \) is the entropy. What is the total entropy change for melting ice?
The chain rule

If the variable(s) of a function \( f \) are themselves function(s) of other variables:

\[
df = \frac{\partial X}{\partial x} \, dx + \frac{\partial Y}{\partial y} \, dy \quad \text{and} \quad X = X(u); Y = Y(u)
\]

We can use the chain rule:

\[
\frac{dX}{du} = \frac{\partial X}{\partial x} \frac{dx}{du} + \frac{\partial X}{\partial y} \frac{dy}{du} \quad \text{and} \quad \frac{dY}{du} = \frac{\partial Y}{\partial x} \frac{dx}{du} + \frac{\partial Y}{\partial y} \frac{dy}{du}
\]

\[
df = \frac{\partial X}{\partial x} \frac{dx}{du} du + \frac{\partial X}{\partial y} \frac{dy}{du} du + \frac{\partial Y}{\partial x} \frac{dx}{du} du + \frac{\partial Y}{\partial y} \frac{dw}{du} du = \frac{dX}{du} du + \frac{dY}{du} du
\]

Example: Assume that we are changing slowly \( T \) and \( V \) for an ideal gas: \( pV = nRT \)

Knowing \( \frac{dT}{dt} = \alpha \) and \( \frac{dV}{dt} = \beta \), what is \( \frac{dp}{dt} \)?

Answer:

\[
\frac{dp}{dt} = \frac{\partial p}{\partial V} \frac{dV}{dt} + \frac{\partial p}{\partial T} \frac{dT}{dt} = \alpha \frac{nR}{V} - \beta \frac{nRT}{V^2}
\]
The Boltzmann Law

Boltzmann’s grave. Zentralfriedhof, Vienna
A bit of history: Ludwig Boltzmann

Europe during Boltzmann’s life

Ludwig Boltzmann
1844 – 1906

Founder of statistical physics
Definition of thermodynamic entropy, $S$

\[ S = k_B \ln(W) \]

$W$ = “Wahrscheinlichkeit”
Old German meaning: multiplicity
New German meaning: probability
We mean multiplicity by $W$.

\[ k_B \approx 1.38 \times 10^{-23} \text{JK}^{-1} = \text{Boltzmann’s constant} \]
Another bit of history: Claude Shannon

Claude E. Shannon (1916–2001)

Founder of information theory

Imagine sending a binary sequence of messages. Which message sequence is more informative?

Message #1: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Message #2: 0 1 1 0 0 0 1 0 1 0 0 1 1 1 0 1 0 1 1 0 0

Shannon Entropy $H$ (in bits):

$$H = - \sum_{i} p_i \log_2(p_i)$$

$p_i$ = probability of message $i$ taken from all possible messages

The higher $H$, the more informative each message. Many symbols, uniformly distributed are great!
What relates S to H? Both refer to distributions.

Imagine casting an $M$-faced die $N$ times. If outcome $i$ occurs $n_i$ times, the multiplicity $W$ is:

$$W = \frac{N!}{n_1! n_2! n_3! \ldots n_M!}$$

Use Stirling’s approximation: $$n! \approx \left(\frac{n}{e}\right)^n$$

$$W = \frac{N^n}{n_1^{n_1} n_2^{n_2} n_3^{n_3} \ldots n_M^{n_M}}$$

$$W = \frac{1}{p_1^{n_1} p_2^{n_2} p_3^{n_3} \ldots p_M^{n_M}}$$

$$S = \frac{\ln(W)}{k} = -\sum_i p_i \ln(p_i)$$
Why the logarithm in $k_B \ln(W)$?

Take two subsystems, A and B.

The entropy of the joint system should be the sum of individual entropies.

$$W_{A+B} = W_A W_B$$

$$S(W_{A+B}) = S(W_A W_B) = k \ln(W_A W_B) = S(W_A) + S(W_B)$$
Max W equivalent to MaxEnt

Spin an arrow around a disk with 4 colors $N$ times.

$$W = \frac{N!}{n_r! n_y! n_g! n_b!}$$

$$S = -\sum_{\{r,y,g,b\}} p_i \ln(p_i)$$

(a) Ordered

$N = 8$

$W = 1$

$S = 0$

(b) Biased

$N = 8$

$W = 2520$

$S = \ln(4)$

(c) Biased

$N = 8$

$W = 2520$

$S = \ln(4)$

(d) Random
Max-multiplicity and MaxEnt for coin flips

Flip a coin $N$ times. Find the MaxEnt distribution!
If Heads=1 occur $n_1$ times, with probability $n_1/N$:

$$W = \frac{N!}{n_1!n_0!}$$

Use Stirling’s approximation: $n! \approx \left(\frac{n}{e}\right)^n$

$$\frac{1}{W} = \frac{n_1^{n_1}n_0^{n_0}}{N^N} = p^{n_1}(1-p)^{n_0} = p^{pN}p^{(1-p)N}$$

$$\frac{d(\ln W^{-1})}{dp} = \frac{d[pN \ln p + (1-p)N \ln(1-p)]}{dp}$$

$$\ln \frac{p}{1-p} = 0 \Rightarrow p = 1 - p = \frac{1}{2}$$

MaxEnt and MaxW imply a flat distribution.
Maximum Entropy + no constraints
= Flat distributions

\[ S(p_1, p_2, \ldots p_M) = -\sum_i p_i \ln(p_i) \]

Goal: to maximize \( S \).

Constraint:
\[ \sum_i p_i = 1 \rightarrow \sum_i dp_i = 0 \]

\[ \frac{\partial S}{\partial p_i} = - \frac{\partial [p_i \ln(p_i)]}{\partial p_i} = -\ln(p_i) - 1 \]

Applying Lagrange multipliers.

\[-\ln(p_i) - 1 - \alpha = 0 \rightarrow p_i = e^{-1-\alpha} = \text{const} \]

\[ \sum_i p_i = \sum_i e^{-1-\alpha} = Me^{-1-\alpha} = 1 \rightarrow p_i = e^{-1-\alpha} \frac{e^{-1-\alpha}}{\sum_i e^{-1-\alpha}} = \frac{1}{M} \]

Conclusion: a flat (uniform) distribution maximizes \( S \).
Maximum Entropy + Average given = Exponential distributions

\[ S(p_1, p_2, \ldots p_M) = -\sum_i p_i \ln(p_i) \]

Goal: to maximize \( S \).

Constraints:
\[ \sum_i p_i = 1 \quad \sum_i dp_i = 0 \quad \sum_i \epsilon_i p_i = \mu \quad \sum_i \epsilon_i dp_i = 0 \]

Applying Lagrange multipliers:
\[-\ln(p_i) - 1 - \alpha - \beta \epsilon_i = 0 \quad \Rightarrow \quad p_i = e^{-(1-\alpha-\beta \epsilon_i)} \]

Application: How would particles distribute on energy levels if their average energy is given?

\[ \sum_i p_i = e^{-1-\alpha} \sum_i e^{-\beta \epsilon_i} = 1 \]

\[ p_i = \frac{e^{-\beta \epsilon_i}}{\sum_i e^{-\beta \epsilon_i}} = \frac{e^{-\beta \epsilon_i}}{Z} \]

Conclusion: the exponential distribution maximizes \( S \).
Example: estimating the bias of loaded dice

For a die, $i=1..6$

$\varepsilon_i = i$

$x = e^{-\beta}$

$p_i = \frac{x^i}{x + x^2 + \ldots + x^6}$

$\langle \varepsilon \rangle = \langle i \rangle = \frac{x + 2x^2 + \ldots + 6x^6}{x + x^2 + \ldots + x^6}$
Example: estimating the bias of coins

For a coin, $i=1,2$ (Heads=1, Tails=2)

$$\varepsilon_i = i$$

$$x = e^{-\beta}$$

$$p_1 = p(Heads) = \frac{x}{x + x^2} \quad p_2 = p(Tails) = \frac{x'}{x + x^2}$$

$$\langle \varepsilon \rangle = \langle i \rangle = \frac{x + 2x^2}{x + x^2} = \frac{1 + 2x}{1 + x}$$

$$x = \frac{\langle \varepsilon \rangle - 1}{2 - \langle \varepsilon \rangle} \quad \text{if} \quad \langle \varepsilon \rangle = 1.8 \quad x = \frac{0.8}{0.2} = 4$$

$$p_1 = \frac{1}{5} \quad p_2 = \frac{4}{5}$$
The Postulates of Thermodynamics by Herbert B. Callen

Through any two points there is exactly one line.

*Postulate (or axiom) = A statement considered true without proof.*
Basic Problem of Thermodynamics

The single, all-encompassing problem of thermodynamics is the determination of the equilibrium state that eventually results after the removal of internal constraints in a closed, composite system.
The single, all-encompassing problem of thermodynamics is the determination of the equilibrium state that eventually results after the removal of internal constraints in a closed, composite system.
Postulate #1

There exist particular states (called equilibrium states) of simple systems that, macroscopically, are characterized completely by the internal energy $U$, the volume $V$, and the amounts of the $M$ components $n_1, n_2, \ldots, n_M$.

\[
S^{(1)} = S^{(1)} (U^{(1)}, V^{(1)}, N_1^{(1)}, N_2^{(1)}, \ldots, N_M^{(1)})
\]
Postulate #2

For all equilibrium states there exists a function (the entropy \([S]\)) of the extensive parameters such that:
In the absence of constraints, the extensive parameters \(U, V, n_1, n_2, \ldots, n_M\) take on values that maximize the entropy \(S\) over all microstates.

\[
S_{\text{max}} = S(U^{(F)}, V^{(F)}, N_1^{(F)}, N_2^{(F)}, \ldots, N_M^{(F)})
\]
Postulate #3

The entropy $S$ of a composite system is additive over the constituent subsystems.

The entropy $S$ is continuous and differentiable and is a monotonically increasing function of the energy.

\[ S^{(1+2)} = S^{(1)} + S^{(2)} \]
Postulate #4

The entropy of any system vanishes in the state for which:

\[ \left( \frac{\partial U}{\partial S} \right)_{V,N_1,N_2,...,N_M} = 0 \] (that is, at the zero of temperature).