Summary of the 6th class

- **Stationary solution from Liouville's theorem: Mechanical invariants and the energy function**

The energy function $E(q, p)$ is the *only* mechanical invariant relevant to the equilibrium distribution.

- **Statistical independence from quasi-closed subsystems: $\log f$ is additive**

$f_{XY}(x, y) = f_X(x) f_Y(y)$ is obtained from statistical independence, and hence $\log f$ is additive.

- **$\log f$ is an additive mechanical invariant: The Gibbs distribution**

The Gibbs distribution is given by $\log f(q, p) = -\alpha - \beta E(q, p)$, in which α is for the normalization of f and β is related to temperature. Note that β is the same for all the quasi-closed subsystems in thermal equilibrium.

- **Entropy, additivity of entropy due to statistical independence, and the entropy for the Gibbs distribution**

Entropy is interpreted as the "width" of a macroscopic state that traverses many microscopic states.

- **The Metropolis algorithm to generate the Gibbs distribution**

A Markov chain is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. That means a Markov chain "memoryless".

Markov chain Monte Carlo methods create samples from a continuous random variable, with probability density proportional to a known function. These samples can be used to evaluate an integral over that variable.

Practically, an ensemble of chains is generally developed, starting from a set of points arbitrarily chosen and sufficiently distant from each other. These chains are **stochastic processes** of "walkers" which move around randomly according to an algorithm that *looks for places with a reasonably high contribution to the integral*, assigning them higher probabilities.

The Metropolis–Hastings algorithm works by generating a sequence of sample values in such a way that, as more and more sample values are produced, the distribution of values more closely approximates the desired distribution *P*(*x*). These sample values are produced iteratively, with the distribution of the next sample being dependent only on the current sample value (thus making the sequence of samples into a Markov chain). Specifically, at each iteration, the algorithm picks a candidate for the next sample value based on the current sample value. Then, with some probability, the candidate is either accepted (in which case the candidate value is used in the next iteration) or rejected (in which case the candidate value is discarded, and current value is reused in the next iteration).

In the next page, this algorithm is shown to generate the desired distribution for a two-energy-level system.

$$E_H - E_L = \Delta > 0, \quad w = e^{-\Delta/k_B T}, \quad \rho_H : \rho_L = e^{-E_H/k_B T} : e^{-E_L/k_B T} = w,$$

with $1/k_B T \equiv \beta$. (k_B is the Boltzmann constant and T is the temperature.)

Markov chain and canonical distribution

$E_H - E_L = \Delta > 0$

$W = e^{-\Delta/k_BT} < 1$

H is always followed by L, while L is followed by H with a probability W.

HLH, W

HLLH, $(1-W)W$

HLLLH, $(1-W)^2W$

HLLLLH, $(1-W)^3W$

$W + (1-W)W + (1-W)^2W \cdots = W[1+(1-W)+(1-W)^2\cdots]$

$= W \times \frac{1}{1-(1-W)} = 1$, normalized.

On average, H is followed by n L's.

$n = W + 2(1-W)W + 3(1-W)^2W + \cdots$

$= W[1 + 2(1-W) + 3(1-W)^2 \cdots] = W\frac{1}{(1-(1-W))^2}$

$= \frac{1}{W}$

$\rho_H : \rho_L = 1 : n = 1 : \frac{1}{W} = W = e^{-\Delta/k_BT}$

as desired.

$S = 1 + x + x^2 + \cdots$

$xS = \quad x + x^2 + \cdots$ $\qquad (1-x)S = 1,\ S = \frac{1}{1-x}$

$T = 1 + 2x + 3x^2 + 4x^3 + \cdots$

$xT = \quad x + 2x^2 + 3x^3 \cdots$

$(1-x)T = 1 + x + x^2 + \cdots = S = \frac{1}{1-x}$

$T = \frac{1}{(1-x)^2}$