Large deviations: their use in equilibrium statistical mechanics and in a broader perspective

STEFANO RUFFO

SISSA, Trieste, Italy

`IFPU Focus Week (Feb 6-10 2023): Dynamical complexity in astrophysical context`

February 6, 2023

Plan

- Law of large numbers and central limit theorem
- Coin tossing
- Large deviations
- Cramèr's theorem
- Three-states Potts model
- Modified XY model
- Free electron laser
- Mean-field ϕ^4 model

Citation from the book of Kerson Huang

There seems to be little hope that we can straightforwardly carry out the recipe of the microcanonical ensemble for any system but the ideal gas.

Law of large numbers

Consider a sample of N independent, identically distributed (i.i.d.) random variables

$$x_1, x_2, \ldots, x_N$$

with PDF $f(x)$ and expectation μ: $< x >= \int f(x) x dx = \mu$
Then, the sample mean

$$X_N = \frac{1}{N} \sum_{i=1}^{N} x_i$$

converges to μ almost surely

$$\text{Prob}\left\{ \lim_{N \to \infty} X_N = \mu \right\} = 1$$

Central limit theorem

Consider a function $g(x)$ of the random variable x and the sample mean

$$G_N = \frac{1}{N}\sum_{i=1}^{N} g(x_i)$$

Define

$$t_N = \frac{G_N - < g(x) >}{\sqrt{\mathrm{var}\left\{G_N\right\}}} = \frac{\sqrt{N}(G_N - < g(x) >)}{\sqrt{\mathrm{var}\{g(x)\}}}$$

Then $(\sigma^2 = \mathrm{var}\{g\})$

$$\lim_{N\to\infty} \mathrm{Prob}\{a < t_N < b\} = \int_a^b \frac{\exp[-t^2/2]}{\sqrt{2\pi}} dt$$

$$f(G_N) = \frac{1}{\sqrt{2\pi(\sigma^2/N)}} \exp\left[\frac{N(G_N - < g >)^2}{2\sigma^2}\right]$$

Coin tossing and large deviations

$$X_k = \pm 1 \quad , \quad S_N = \frac{1}{N}\sum_{k=1}^{N} X_k$$

$$P(S_N = x) = \frac{N!}{N_+!N_-!2^N} = \frac{N!}{\left(\frac{(1+x)N}{2}\right)! \ \left(\frac{(1-x)N}{2}\right)! \ 2^N}$$

Using the Stirling's formula in the large N limit

$$\ln P(x) \sim -N\left(\frac{(1+x)}{2}\ln(1+x) + \frac{(1-x)}{2}\ln(1-x)\right) \quad \sim -NI(x)$$

The *rate function* $I(x)$ has a single minimum in $x = 0$, the most probable value and is in this case symmetric around the minimum. S_N fulfills a *large deviation principle*, characterized by the rate function $I(x)$.

The coin toss experiment can be thought as a microscopic realization of a chain of N non-interacting Ising spins. $I(x)$ corresponds to the opposite of the Boltzmann entropy of a macrostate characterized by a fraction x of up-spins.

Cramèr's theorem

Let $\mathbf{X} \in R^d$ be a random variable with given PDF and $\mathbf{X}_i, i = 1, \ldots, N$, a sample of $\mathbf{X}$.
Let $\mathbf{M}_N = \frac{1}{N} \sum_i \mathbf{X}_i$ be sample mean
Which is the PDF of the sample mean? (`Cramèr`)
Compute the generating function

$$\Psi(\lambda) =< \exp(\lambda \cdot \mathbf{X}) >,$$

with $\lambda \in R^d$ and the average $< \cdot >$ performed on the PDF of $\mathbf{X}$ If $\Psi(\lambda) < \infty$ and differentiable, then

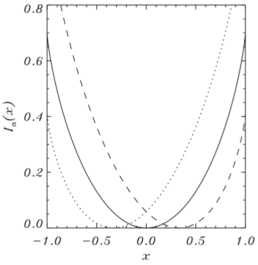
$$P(\mathbf{M}_N = \mathbf{x}) \sim \exp(-NI(\mathbf{x}))$$

where the rate function $I(\mathbf{x})$ is given by the Legendre-Fenchel transform of $\ln(\Psi(\lambda))$

$$I(\mathbf{x}) = \sup_{\lambda \in R^d} (\lambda \cdot \mathbf{x} - \ln(\Psi(\lambda)))$$

Unbiased/Biased coin tossing using Cramèr

- **Unbiased:** $d\mu = [\delta(X-1) + \delta(X+1]dX/2$;
  $\Psi(\lambda) = \langle \exp(\lambda X) \rangle = \cosh \lambda$; $I(x) = \sup_\lambda (\lambda \cdot x - \ln \cosh \lambda)$, whose critical point is $\lambda = \operatorname{arth} x$.
- **Biased:** $d\mu = [(1-\alpha)\delta(X-1) + \alpha\delta(X+1]dX$, with $\alpha \in [0,1]$ and $\alpha = 1/2$ corresponding to the unbiased case; $\Psi_\alpha(\lambda) = \exp(\lambda) - 2\alpha \sinh \lambda$. $I_\alpha(\lambda)$ is plotted in the figure for $\alpha = 1/3, 1/2, 2/3$. This model corresponds to an ensemble of non-interacting Ising spins whose probability to take the upper value is different from the one for the down value.



Entropy and free energy

Step 1 Express the Hamiltonian in terms of global variables γ

$$H_N(\omega_N) = \widetilde{H}_N\left(\gamma(\omega_N)\right) + R_N(\omega_N)$$

(ω_N a phase-space configuration) leading to
$h(\gamma) = lim_{N\to\infty} \widetilde{H}_N\left(\gamma(\omega_N)\right)/N$.
Step 2 Compute the entropy functional in terms of the global variables using, e.g., Cramèr's theorem

$$s(\gamma) = \lim_{N\to\infty} \frac{1}{N} \ln \Omega_N(\gamma)$$

with $\Omega_N(\gamma)$ the number of microscopic configurations with fixed γ.
Step 3 Solve the microcanonical and canonical variational problems

$$s(\varepsilon) = \sup_{\gamma}\left(s(\gamma) \mid h(\gamma) = \varepsilon\right) \ ,$$

$$\beta f(\beta) = \inf_{\gamma}\left(\beta h(\gamma) - s(\gamma)\right)$$

Potts model-I

$$H_N^{Potts} = -\frac{J}{2N} \sum_{i,j=1}^{N} \delta_{S_i, S_j} \quad .$$

$S_i = a, b, c$

Step 1

$$\widetilde{H}_N^{Potts} = -\frac{JN}{2}(n_a^2 + n_b^2 + n_c^2)$$

Step 2

$$\gamma = \left(\frac{1}{N} \sum_i \delta_{S_i,a}, \frac{1}{N} \sum_i \delta_{S_i,b}, \frac{1}{N} \sum_i \delta_{S_i,c} \right) \quad .$$

Local random variables

$$\mathbf{X}_k = (\delta_{S_k,a}, \delta_{S_k,b}, \delta_{S_k,c})$$

Potts model-II

Generating function

$$\begin{aligned}\Psi(\lambda_a, \lambda_b, \lambda_c) &= \frac{1}{3} \sum_{S=a,b,c} \left(e^{\lambda_a \delta_{S,a} + \lambda_b \delta_{S,b} + \lambda_c \delta_{S,c}} \right) \\ &= \frac{1}{3} \left(e^{\lambda_a} + e^{\lambda_b} + e^{\lambda_c} \right)\end{aligned}$$

Rate function

$$I(\gamma) = \sup_{\lambda_a, \lambda_b, \lambda_c} \left(\lambda_a n_a + \lambda_b n_b + \lambda_c n_c - \ln \Psi(\lambda_a, \lambda_b, \lambda_c) \right) \quad .$$

Exact solution $\lambda_\ell = \ln n_\ell$, with $\ell = a, b, c$

$$I(\gamma) \;=\; n_a \ln n_a + n_b \ln n_b + (1 - n_a - n_b) \ln(1 - n_a - n_b) + \ln 3$$

Entropy

$$s(\gamma) = -I(\gamma) + \ln \mathcal{N}$$

where the normalization factor is $\mathcal{N} = 3$

Potts model-III

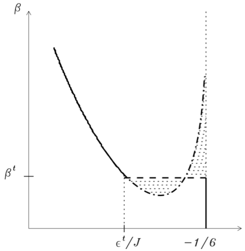
Step 3

Microcanonical entropy

$$s(\varepsilon) = \sup_{n_a,n_b}\Big(-n_a \ln n_a - n_b \ln n_b - (1-n_a-n_b)\ln(1-n_a-n_b)$$
$$\Big| -\frac{J}{2}\left(n_a^2+n_b^2+(1-n_a-n_b)^2\right) = \epsilon\Big)$$

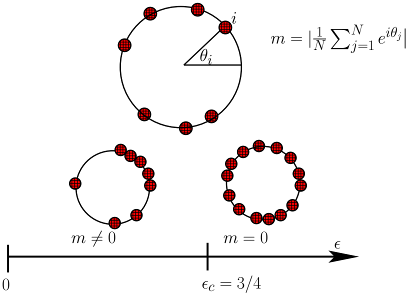
Canonical free energy

$$\beta f(\beta) = \inf_{n_a,n_b,n_c}\left(n_a \ln n_a + n_b \ln n_b + n_c \ln n_c - \frac{\beta J}{2}\left(n_a^2+n_b^2+n_c^2\right)\right)$$



Generalized XY model

$$H_{XY} = \sum_{i=1}^{N} \frac{p_i^2}{2} - \frac{J}{2N}(\sum_{i=1}^{N} \vec{s}_i)^2 - \frac{K}{4N^3}\left[(\sum_{i=1}^{N} \vec{s}_i)^2\right]^2, \quad \vec{s}_i = (\cos\theta_i, \sin\theta_i)$$



Entropy of XY model

Step 1 Global variables

$$\gamma = (m_x, m_y, \mathcal{E}_K) \text{ with } \mathcal{E}_K = \lim_{N\to\infty} \sum_i p_i^2/N$$

$$h(\gamma) = \frac{1}{2}\left(\mathcal{E}_K - Jm^2 - Km^4/2\right)$$

Step 2

$$\mathbf{X} = \left(\cos\theta, \sin\theta, p^2\right) \text{ Local random variable}$$

$$\Psi(\lambda) \simeq I_0(\sqrt{\lambda_x^2 + \lambda_y^2})/\sqrt{-\lambda_K} \text{ where } \lambda = (\lambda_x, \lambda_y, \lambda_K)$$

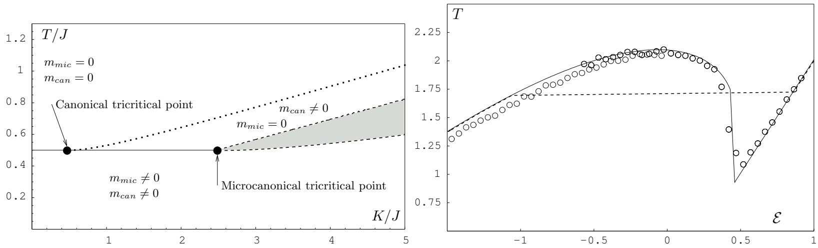
$$I(\gamma) = -s(\gamma) = \sup_\lambda(\lambda_K \mathcal{E}_K + \lambda_x m_x + \lambda_y m_y +$$

$$+ \ln(-\lambda_K)/2 - \ln(I_0(\sqrt{\lambda_x^2 + \lambda_y^2})))$$

Step 3 Entropy

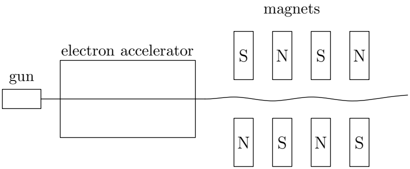
$$s(\varepsilon) = \sup_\gamma\{s(\gamma) \mid \mathcal{E}_K = 2\varepsilon + Jm^2 + Km^4/2)\}$$

Phase diagram and caloric curves



- At $K/J = 0$ (HMF model), second order phase transition at $T/J = 0.5$. Ensembles are equivalent.
- For $K/J < 1/2$ ensembles are inequivalent. Negative specific heat for $1/2 < K \le 5/2$; Temperature jumps for $K > 5/2$.
- Right figure shows the caloric curve for $K/J = 10$. The points are results of a molecular dynamics simulation with $N = 100$

Free Electron Laser



Colson-Bonifacio model

$$
\begin{aligned}
\frac{d\theta_j}{dz} &= p_j \\
\frac{dp_j}{dz} &= -\mathbf{A}e^{i\theta_j} - \mathbf{A}^* e^{-i\theta_j} \\
\frac{d\mathbf{A}}{dz} &= i\delta\mathbf{A} + \frac{1}{N}\sum_j e^{-i\theta_j}
\end{aligned}
$$

Microcanonical solution

Hamiltonian

$$H_N = \sum_{j=1}^{N} \frac{p_j^2}{2} - N\delta A^2 + 2A \sum_{j=1}^{N} \sin(\theta_j - \varphi)$$

where $A = \sqrt{\mathbf{A}\mathbf{A}^*}$.

Entropy

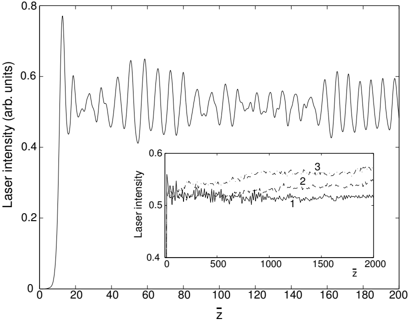
$$s(\varepsilon, \sigma, \delta) = \sup_{A,m} \left[\frac{1}{2} \ln \left[2 \left(\varepsilon - \frac{\sigma^2}{2} \right) + 4Am + 2(\delta - \sigma) A^2 - A^4 \right] + s_{conf}(m) \right.$$

where $m = \sqrt{m_x^2 + m_y^2}$, $m_x = \sum_i \cos\theta_i / N$, $m_y = \sum_i \sin\theta_i / N$, σ is the total average momentum $\sum_i p_i / N + A^2$ and

$$s_{conf}(m) = -\sup_{\lambda} \left[\lambda m - \ln I_0(\lambda) \right]$$

Ensembles are equivalent for this model. There is a second order phase transition at $\varepsilon = -1/(2\delta)$, $\delta < 0$.

Time relaxation of the laser intensity



$N = 5000$ (curve 1), $N = 400$ (curve 2), $N = 100$ (curve 3)

On a first stage the system converges to a quasi-stationary state. Later it relaxes to equilibrium on a time $O(N)$. The quasi-stationary state is a Vlasov equilibrium, sufficiently well described by Lynden-Bell's distributions.

Mean-field ϕ^4 model

$$H = \sum_{i=1}^{N}\left(\frac{p_i^2}{2} - \frac{1}{4}q_i^2 + \frac{1}{4}q_i^4\right) - \frac{1}{4N}\sum_{i,j=1}^{N} q_i q_j .$$

Global variables

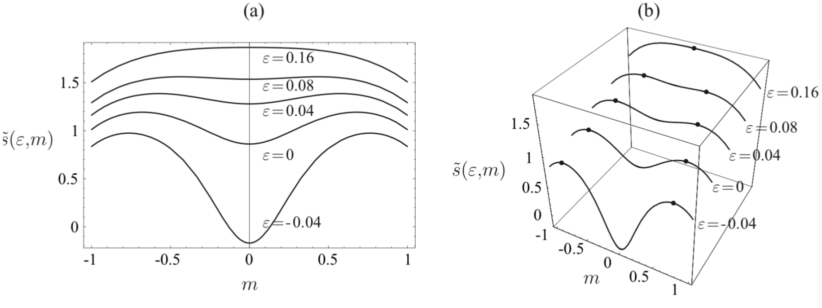
$$u = \frac{1}{N}\sum_{i=1}^{N} p_i^2 \quad , \quad z = \frac{1}{4N}\sum_{i=1}^{N}(q_i^4 - q_i^2) \quad , \quad m = \frac{1}{N}\sum_{i=1}^{N} q_i$$

$$\ln\Psi(\lambda_u, \lambda_z, \lambda_m) = -\frac{\ln\lambda_u}{2} + \ln\int dq\,\exp(-\lambda_m q - \lambda_z(q^4 - q^2)) + \mathrm{const}$$

$$s(u, z, m) = \inf_{\lambda_u,\lambda_z,\lambda_m}(\lambda_u u + \lambda_z z + \lambda_m m - \ln\Psi)$$

$$s(\varepsilon, m) = \sup_{u,z}(s(u, z, m)|\varepsilon = \frac{u}{2} + z - \frac{m^2}{4})$$

Entropy of the mean-field ϕ^4 model



Negative susceptibility

Thermodynamics first law for magnetic systems $TdS = dE - hdM$.
In the microcanonical ensemble

$$h(\varepsilon, m) = -\frac{\partial s}{\partial m} / \frac{\partial s}{\partial \varepsilon} = -\frac{1}{\beta(\varepsilon, m)} \frac{\partial s}{\partial m}.$$

In the canonical ensemble

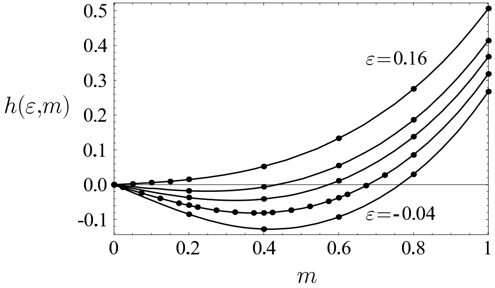
$$f(\beta, h) = \inf_{\varepsilon, m} \left[\varepsilon - hm - \frac{1}{\beta} s(\varepsilon, m) \right].$$

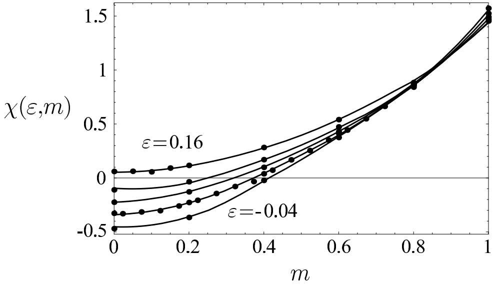
which gives $\partial s / \partial m = -hm$, $\partial s / \partial \varepsilon = \beta$, in agreement with the microcanonical expressions for h and β.

$$\chi = \frac{\partial m}{\partial h} = \beta \frac{s_{\varepsilon\varepsilon}}{s_{\varepsilon m}^2 - s_{\varepsilon\varepsilon} s_{mm}}$$

In the canonical ensemble $s_{\varepsilon\varepsilon} > 0$ and the denominator is positive as a consequence of stationarity, hence $\chi > 0$. In the microcanonical ensemble $s_{mm} < 0$ and, at free energy saddles, $s_{\varepsilon\varepsilon} < 0$, hence susceptibility can be negative.

Comparison with numerics





Conclusions

- Large deviations are a powerful tool to derive microcanonical entropies.
- Examples: Potts model, generalized XY model, Colson-Bonifacio model of the free electron laser, ϕ^4 theory