Active matter – Week 4: Large deviation theory

DRSTP Advanced Topics in Theoretical Physics (Spring 2025)

Yann-Edwin Keta
keta@lorentz.leidenuniv.nl

Large deviation theory and its application to (nonequilibrium) statistical physics are very well introduced in reviews by Hugo Touchette [1, 2].

I. STATISTICS OF AVERAGES

I.A. Large deviation principle

Consider a sequence $X_1, \ldots, X_N$ of independent and identically distributed random variables. We denote

$$\langle X_i \rangle = \mu, \tag{1a}$$

$$\left\langle (X_i - \mu)^2 \right\rangle = \sigma^2, \tag{1b}$$

their mean and variance. We introduce the sample average

$$R_N = \frac{1}{N} \sum_{i=1}^{N} X_i, \tag{2}$$

which is itself a random variable. We can quantify the statistics of R_N at different levels in the $N \to \infty$ limit. First, the *law of large numbers* indicates that X_N converges almost certainly to the expected value μ,

$$\mathrm{Prob}\left(\lim_{N \to \infty} R_n = \mu \right) = 1. \tag{3}$$

Second, the *central limit theorem* tells us that X_N converges in distribution to a normal distribution of mean μ and variance σ^2/N,

$$\mathrm{Prob}(R_N) \underset{N \to \infty}{\sim} \mathcal{N}\left(\mu, \frac{\sigma^2}{N} \right). \tag{4}$$

Both of these results give us information about the statistics of R_N close to the mean μ. Third and lastly, the *large deviation principle* characterises the statistics of R_N away from its mean[1]. R_N is said to satisfy a large deviation principle if the following limit exists,

$$\lim_{N \to \infty} \frac{1}{N} \log \mathrm{Prob}(R_N = r) = -I(r), \tag{5}$$

where I is called the *rate function* which characterises the exponential decay of probabilities of R_N away from its mean μ. We denote the asymptotic convergence of probability of R_N as

$$\mathrm{Prob}(R_N = r) \asymp \exp(-N\, I(r)). \tag{6}$$

We introduce the *scaled cumulant generating function*

$$\psi(s) = \lim_{N \to \infty} \frac{1}{N} \log \left\langle \exp(sNR_N) \right\rangle. \tag{7}$$

The Gärtner-Ellis theorem states that if ψ exists and is differentiable, then R_N satisfies a large deviation principle and the rate function is given by

$$I(r) = \sup_s \{sr - \psi(s)\}, \tag{8a}$$

[1] Realisations of R_N away from μ in the limit $N \to \infty$ are characterised as *rare*.

and conversely

$$\psi(s) = \sup_r \{sr - I(r)\}, \tag{8b}$$

such that the rate function I and the scaled cumulant generating function ψ are related by Legendre-Fenchel transforms.

I.B. Properties of the rate function and the scaled cumulant generating function

It follows from (7) that

$$\psi(0) = 0, \tag{9a}$$

and the derivatives of ψ at $s=0$ are linked to the asymptotic mean

$$\psi'(0) = \lim_{N\to\infty} \frac{1}{N} \left. \frac{\langle NR_N \exp(sNR_N)\rangle}{\langle \exp(sNR_N)\rangle} \right|_{s=0} = \lim_{N\to\infty} \langle R_N \rangle = \mu \tag{9b}$$

and variance

$$\begin{aligned}\psi''(0) &= \lim_{N\to\infty} \frac{1}{N} \frac{\big\langle (NR_N)^2 \exp(sNR_N)\big\rangle\big\langle \exp(SNR_N)\big\rangle - \big\langle NR_N \exp(sNR_N)\big\rangle\big\langle NR_N \exp(sNR_N)\big\rangle}{\big\langle \exp(sNR_N)\big\rangle^2} \\ &= \lim_{N\to\infty} N \left(\big\langle R_N^2\big\rangle - \big\langle R_N\big\rangle^2\right) \\ &= \sigma^2\end{aligned} \tag{9c}$$

of the sample average R_N. Moreover it follows from (9a)

$$\psi(0) = \sup_r \{-I(r)\} = \inf_r I(r) = 0, \tag{10}$$

thus the rate function I is a positive function,

$$I(r) \geq 0. \tag{11}$$

We denote $s(r)$ the value which maximises (8a), thus

$$I(r) = s(r)r - \psi(s(r)). \tag{12}$$

Since this is a maximum this implies that

$$\left.\frac{\partial}{\partial s}(sr - \psi(s))\right|_{s=s(r)} = 0 \Rightarrow \psi'(s(r)) = r. \tag{13}$$

It is possible to show that rate functions I and scaled cumulant generating function ψ obtained from the Gärtner-Ellis theorem are strictly convex by property of the Legendre-Fenchel transform [1], and that their curvatures are related by

$$I''(r) = \frac{1}{\psi''(s(r))}. \tag{14}$$

Their convexity implies that there is an unique $s(r)$ which satisfies (13) and in particular, considering (9b) and (12), this indicates that

$$s(\mu) = 0, \tag{15}$$

$$I(\mu) = 0, \tag{16}$$

with the latter being the global minimum of I thus

$$I'(\mu) = 0. \tag{17}$$

We note that (16) implies from (6) that

$$\mathrm{Prob}(R_N = \mu) \asymp 1, \tag{18}$$

which is equivalent to the law of large numbers (3), and moreover, using (14), we obtain

$$I''(\mu) = \frac{1}{\psi''(s(\mu))} = \frac{1}{\psi''(0)} = \frac{1}{\sigma^2}, \tag{19}$$

thus with a Taylor expansion of I close to μ,

$$I(r) = I(\mu) + I'(\mu)(r-\mu) + \frac{1}{2}I''(\mu)(r-\mu)^2, \tag{20}$$

and again from (6),

$$\mathrm{Prob}(R_n = r) \asymp \exp\left(-\frac{(r-\mu)^2}{2(\sigma^2/N)}\right), \tag{21}$$

which is equivalent to the central limit theorem (4).

II. LARGE DEVIATION THEORY AND STATISTICAL MECHANICS

II.A. Analogy with equilibrium statistical mechanics

Consider a system of N particles and E_N the mean energy per particle. At thermal equilibrium, the distribution of microstates ω only depends on their energy $NE_N(\omega)$ and is given by the Boltzmann distribution

$$\mathrm{Prob}_\beta(\omega) = \frac{e^{-\beta N E_N(\omega)}}{Z(\beta)}, \tag{22}$$

where $\beta = 1/(k_B T)$, with β the Boltzmann constant and T the temperature, and where

$$Z(\beta) = \sum_\omega e^{-\beta N E_N(\omega)} \tag{23}$$

is the partition function. It is noteworthy that the Boltzmann weights for the mean energy per particle (22) feature a similar exponential form as sample averages following a large deviation principle (6), which is the basis of our analogy.

We can define the scaled cumulant generating function of the mean energy E_N

$$\begin{aligned}\psi_\beta(\Delta\beta) &= \lim_{N\to\infty} \frac{1}{N} \log \sum_\omega e^{-\Delta\beta\, N\, E_N(\omega)}\, \mathrm{Prob}_\beta(\omega) \\ &= \lim_{N\to\infty} \frac{1}{N} \log \frac{Z(\beta+\Delta\beta)}{Z(\beta)} \\ &= \beta F(\beta) - (\beta+\Delta\beta)F(\beta+\Delta\beta),\end{aligned} \tag{24}$$

which is then related to the free energy density

$$\beta F(\beta) = -\lim_{N\to\infty} \frac{1}{N} \log Z_N(\beta). \tag{25}$$

According to the Gärtner-Ellis theorem, if the scaled cumulant generating function ψ_β is differentiable, *i.e.* if the free energy density βF is differentiable, then E_N satisfies a large deviation principle,

$$P_\beta(E_N) \asymp \exp(-N\, I_\beta(E_N)). \tag{26}$$

Given the Boltzmann distribution (22) and the number $C(E_N)$ of microstates ω with mean energy E_N per particle, we know from equilibrium statistical mechanics that E_N has the following distribution

$$P_\beta(E_N) = \mathcal{N}\, C(E_N)\, \frac{e^{-\beta N E_N}}{Z(\beta)}, \tag{27}$$

where $\mathcal{N}$ is some normalisation constant, and thus in the $N \to \infty$ limit

$$P_\beta(E_N) \asymp \exp(N(S(E_N) - \beta E_N + \beta F(\beta))) \tag{28}$$

where we have introduced the entropy density

$$S(E_N) = \lim_{N\to\infty} \frac{1}{N} \log C(E_N). \tag{29}$$

We obtain the rate function by identification with (6)

$$I_\beta(E_N) = -S(E_N) + \beta E_N - \beta F(\beta), \tag{30}$$

from which it follows that the most probable mean energy per particle E_N^*, obtained from the property $I_\beta(E_N^*) = 0$, then satisfies

$$F(\beta) = E_N^* - \frac{1}{\beta} S(E_N^*), \tag{31}$$

which defines the Helmoltz free energy at inverse temperature β.

II.B. Dynamical phase transitions in nonequilibrium statistical physics

Consider a stochastic process $\boldsymbol{X}(t)$ and an observable

$$O_\tau = \frac{1}{\tau} \int_0^\tau \mathrm{d}t\, f(\boldsymbol{X}(t)), \tag{32}$$

where f is some arbitrary deterministic function, so that we can define a scaled cumulant generating function of O_τ,

$$\psi(s) = \lim_{\tau\to\infty} \frac{1}{\tau} \log \left\langle e^{s\tau O_\tau} \right\rangle. \tag{33}$$

We can draw an analogy between the equilibrium canonical ensemble, discussed in the previous section II.A, and biased ensembles of realisations of $\boldsymbol{X}(t)$. A realisation $\boldsymbol{x}(t)$ of the stochastic process $\boldsymbol{X}(t)$, for example the trajectory of the degrees of freedom of some stochastic system, may be considered as a microstate characterised by a mean energy O_τ. Given a positive (respectively negative) value of s, the average in (33) will be dominated by trajectories with larger (respectively smaller) values of O_τ with respect to the expected mean value of O_τ, and these correspond to trajectories sampled by a *biased* path probability distribution

$$P_s[\boldsymbol{x}, t] = P_0[\boldsymbol{x}, t]\, e^{s\tau O_\tau}, \tag{34}$$

where $P_0[\boldsymbol{x}, t]$ is the unbiased path probability of realisations of our stochastic process and where s is then interpreted as a biasing field. In analogy with (24), the biasing field s corresponds to a difference in inverse temperature $\Delta\beta$, such that we are probing trajectories typical of equilibrium systems at higher or lower temperatures (which are by definition rare), and the scaled cumulant generating function ψ is analogous to a difference in free energy density. This analogy suggests that singularities in ψ indicate *dynamical phase transitions* where symmetries of the trajectories are either created or destroyed.

Analytically, we can compute the scaled cumulant generating function (33) by solving a spectral problem. We introduce the Fokker-Planck linear operator $\mathcal{L}$ of stochastic process $\boldsymbol{X}(t)$ such that its Fokker-Planck equation is written

$$\frac{\partial}{\partial t} P(\boldsymbol{x}, t) = \mathcal{L} P(\boldsymbol{x}, t), \tag{35}$$

then $\psi(s)$ is given by the largest eigenvalue of the following *tilted* operator[2]

$$\mathcal{W}_s = \mathcal{L} + sf. \tag{36}$$

Numerically, ensembles of biased trajectories can be generated using population-dynamics methods, and the scaled cumulant generating function can be computed from these ensembles [3]. An analogy with optimal control theory can also be drawn in this context [4]: solving the large deviation problem (35) can be thought as finding the *least unlikely* control mechanism[3] of our trajectories which samples the biased distribution (34).

[2] You can find a derivation of this result using Feynman-Kac formalism in Ref. [2].
[3] In the case of an active particles system, it may for example be an additional interaction force between particles.

III. PROBLEMS

A number of examples can be found in the aforementioned reviews [1, 2] as well as in Ref. [5]. In the context of active particles, Ref. [6] provides some analytical and numerical derivations.

REFERENCES

[1] H. Touchette, The large deviation approach to statistical mechanics, Physics Reports **478**, 1 (2009).
[2] H. Touchette, Introduction to dynamical large deviations of Markov processes, Physica A: Statistical Mechanics and its Applications **504**, 5 (2018).
[3] T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, Population-dynamics method with a multicanonical feedback control, Physical Review E **93**, 062123 (2016).
[4] R. L. Jack, Ergodicity and large deviations in physical systems with stochastic dynamics, The European Physical Journal B **93**, 74 (2020).
[5] R. Chetrite and H. Touchette, Nonequilibrium Markov Processes Conditioned on Large Deviations, Annales Henri Poincaré **16**, 2005 (2015).
[6] Y.-E. Keta, É. Fodor, F. van Wijland, M. E. Cates, and R. L. Jack, Collective motion in large deviations of active particles, Physical Review E **103**, 022603 (2021).