The large deviation approach to statistical mechanics

Hugo Touchette

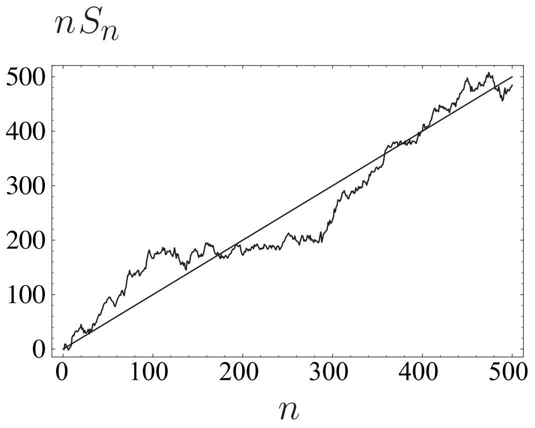
School of Mathematical Sciences, Queen Mary, University of London

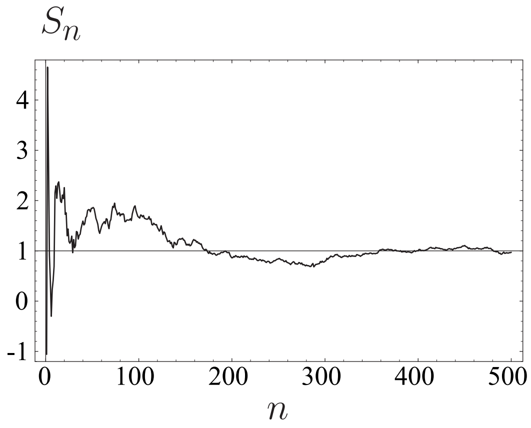
Statistical Mechanics Study Group
Queen Mary, University of London
January 24, 2008

Outline

Example 1: Sum of Gaussian random variables

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad X \sim N(\mu, \sigma^2)$$





Basic observations

- $S_n \to \mu$ in probability
- Fluctuations $\sim 1/\sqrt{n} \to 0$

Sum of Gaussian random variables (cont'd)

- Probability density of S_n:

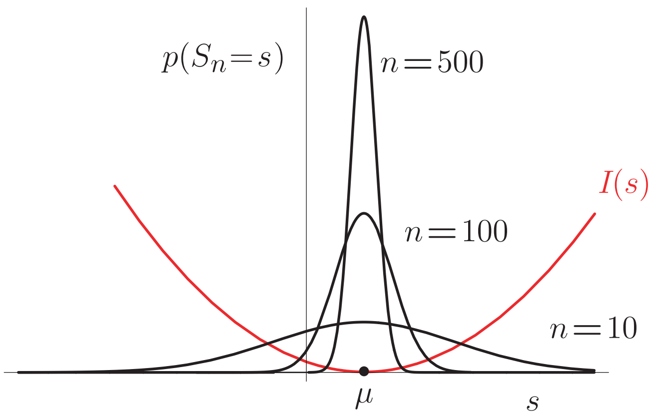
$$P(S_n = s) = \sqrt{\frac{n}{2\pi\sigma^2}}\, e^{-n(s-\mu)^2/(2\sigma^2)}$$

- Dominant part:

$$P(S_n = s) \approx e^{-nI(s)}$$

- Rate function:

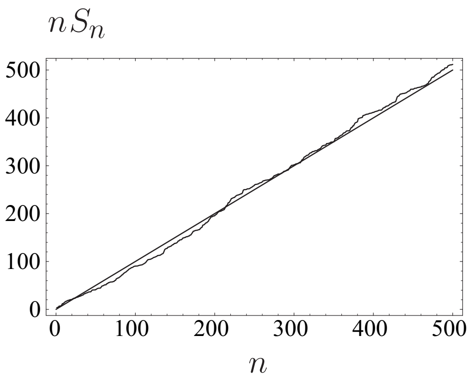
$$I(s) = \frac{(s-\mu)^2}{2\sigma^2}$$

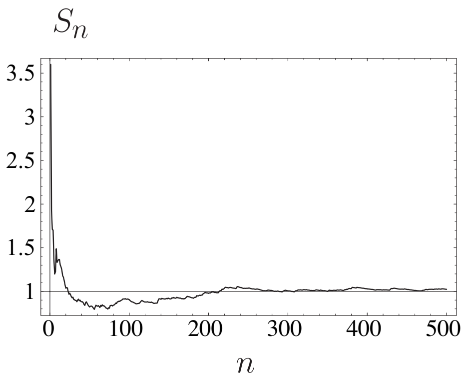


Example 2: Sum of Poisson random variables

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$p(x) = \frac{1}{\mu} e^{-x/\mu}$$



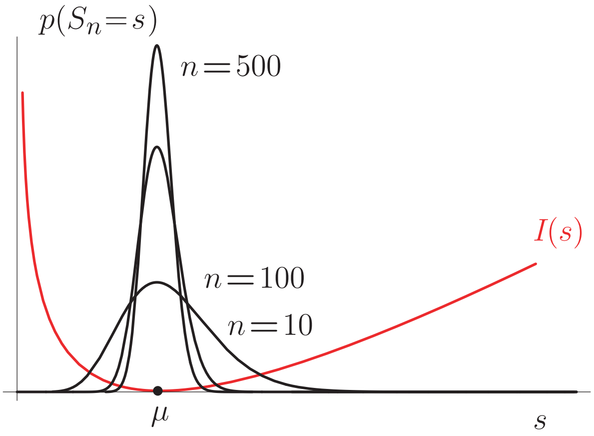


- Large deviation probability:

$$P(S_n = s) \approx e^{-nI(s)}$$

- Rate function:

$$I(s) = \frac{s}{\mu} - 1 - \ln\frac{s}{\mu}$$



Example 3: Equilibrium fluctuations

- N-particle system
- Macrostate: m

Microcanonical

Einstein (1925)

$$P_U(m) = e^{S(m,U)}$$

- Extensivity: $S \sim N$
- Large deviation probability:

$$P_U(m) \approx e^{-NI_U(m)}$$

- Equilibrium: $I_U(m^*) = 0$

Canonical

Landau (1935)

$$P_\beta(m) = e^{-F(m,\beta)}$$

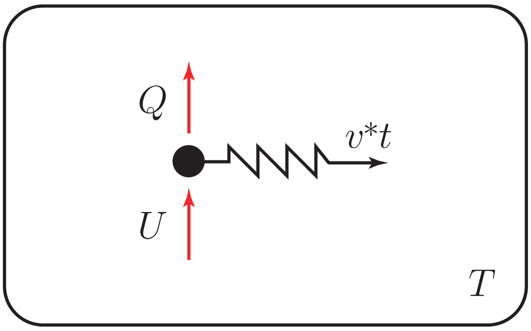
- Extensivity: $F \sim N$
- Large deviation probability:

$$P_\beta(m) \approx e^{-NI_\beta(m)}$$

- Equilibrium: $I_\beta(m^*) = 0$

Example 4: Nonequilibrium fluctuations

van Zon and Cohen (2003)



- Pulling force: $F = -k[x(t) - v^* t]$
- Langevin dynamics:

$$\dot{x}(t) = -\tau_r^{-1}[x(t) - v^* t] + \alpha^{-1}\xi(t)$$

- Work: $W_\tau = \frac{1}{\tau}\int_0^\tau F v^* dt = \Delta U + Q_\tau$

Work fluctuations

- Large deviation probability:

$$P(W_\tau = w) \approx e^{-\tau I(w)}$$

- Fluctuation Theorem:

$$\frac{P(w)}{P(-w)} = e^{a\tau w}, \qquad I(w) = I(-w) - a\tau w$$

Large deviation theory

- Random variable: A_n
- Probability density: $P(A_n = a)$

Large deviation principle

$$P(A_n = a) \approx e^{-nI(a)}$$

$$\begin{aligned} \ln P(a) &= -nI(a) + o(n) \\ \lim_{n\to\infty} -\frac{1}{n}\ln P(a) &= I(a) \end{aligned}$$

- Rate function: $I(a) \geq 0$

Goals of large deviation theory

1. Prove that a large deviation principle exists
2. Calculate the rate function

Large deviation theory (cont'd)

Large deviation principle

- Upper bound for closed sets C:

$$\limsup_{n\to\infty} \frac{1}{n} \ln P(A_n \in C) \leq -\inf_{a\in C} I(a)$$

- Lower bound for open sets O:

$$\liminf_{n\to\infty} \frac{1}{n} \ln P(A_n \in O) \geq -\inf_{a\in O} I(a)$$

- $I(a)$ is lower semicontinuous
- Bounds related to weak convergence

The Gärtner-Ellis Theorem

Log-generating function or free energy function

$$\lambda(k) = \lim_{n\to\infty} \frac{1}{n} \ln\langle e^{nkA_n}\rangle, \qquad k \in \mathbb{R}$$

Theorem: Gärtner (1977), Ellis (1984)

If $\lambda(k)$ is differentiable, then

1. Existence of large deviation principle:

$$P(A_n = a) \approx e^{-nI(a)}$$

2. Expression for $I(a)$:

$$I(a) = \max_k \{ka - \lambda(k)\}$$

- $I(a)$ is the Legendre transform of $\lambda(k)$
- $I(a)$ is convex

Contraction principle

- Large deviation principle for A_n:

$$P(A_n = a) \approx e^{-nI_A(a)}$$

- Contraction: $B_n = f(A_n)$

Contraction principle

- Large deviation principle for B_n:

$$P(B_n = b) \approx e^{-nI_B(b)}$$

- Rate function:

$$I_B(b) = \inf_{a:f(a)=b} I_A(a) = \inf_{f^{-1}(b)} I_A(a)$$

Sum of IID random variables

Cramér (1938)

- Random variable:

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad X_i \sim p(x), \quad IID$$

- Free energy function:

$$\lambda(k) = \lim_{n\to\infty} \frac{1}{n} \ln\langle e^{nkS_n}\rangle = \lim_{n\to\infty} \frac{1}{n} \ln\left\langle \prod_{i=1}^{n} e^{kX_i}\right\rangle = \ln\langle e^{kX}\rangle$$

Gaussian

$$\lambda(k) = \mu k + \frac{\sigma^2}{2}k^2, \quad k \in \mathbb{R}$$

$$I(s) = \frac{(s-\mu)^2}{2\sigma^2}, \quad s \in \mathbb{R}$$

Poisson

$$\lambda(k) = -\ln(1-\mu k), \quad k < \frac{1}{\mu}$$

$$I(s) = \frac{s}{\mu} - 1 - \ln\frac{s}{\mu}, \quad s > 0$$

Sanov's Theorem

Sanov (1961), Boltzmann (1877)

- Sequence of IID random variables: $X_1, X_2, \dots, X_n$
- Empirical frequencies:

$$L_n(x) = \frac{1}{n}\sum_{i=1}^{n}\delta(X_i - x) = \frac{\#\ occurrences\ of\ X_i = x}{n}$$

Gärtner-Ellis Theorem

- Free energy:

$$\lambda[k] = \lim_{n\to\infty}\frac{1}{n}\ln\langle e^{nk\cdot L_n}\rangle = \ln\langle e^{k(X)}\rangle$$

- Large deviation principle:

$$P(L_n = \ell) \approx e^{-nD(\ell||p)}, \qquad D(\ell||p) = \int dx\ \ell(x)\ln\frac{\ell(x)}{p(x)}$$

Markov processes

Donsker and Varadhan (1975)

- Markov chain: $X_1 \to X_2 \to \cdots \to X_n$
- Probability:

$$P(x) = \mu(x_1)\pi(x_2|x_1)\cdots\pi(x_n|x_{n-1})$$

- Additive process:

$$S_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

Gätner-Ellis Theorem

- Free energy function:

$$\lambda(k) = \ln\zeta(\Pi_k), \quad \pi_k(x_n|x_{n-1}) = \pi(x_n|x_{n-1})e^{kf(x_n)}$$

- Large deviation principle:

$$P(S_n = s) \approx e^{-nI(s)}, \quad I(s) = \sup_k\{ks - \lambda(k)\}$$

Noise-pertubed dynamics

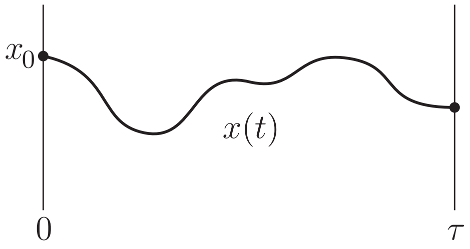
Freidlin and Wentzell (1984)

- Perturbed dynamics:

$$\dot{X}_\epsilon(t) = b(X_\epsilon) + \sqrt{\epsilon}\eta(t), \quad \epsilon \to 0$$

- Gaussian white noise:

$$\langle \eta(t) \rangle = 0, \quad \langle \eta(t)\eta(t') \rangle = \delta(t-t')$$



LDP for the paths

$$P[x] \approx e^{-J[x]/\epsilon}, \quad J[x] = \frac{1}{2}\int_0^\tau [\dot{x}(t) - b(x(t))]^2 \, dt$$

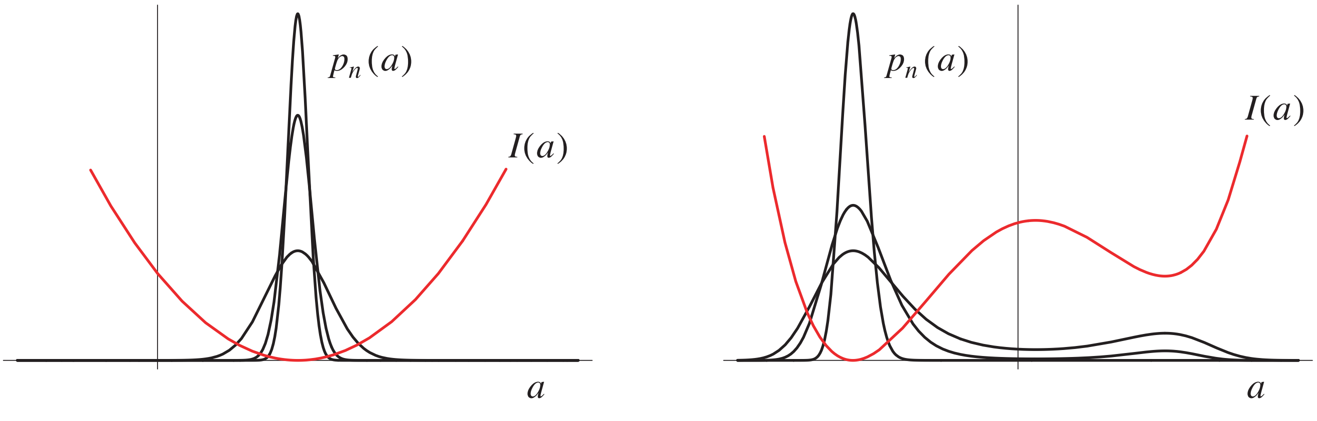
WKB approximations

$$P(x,\tau|x_0) = \int_{x(0)=x_0}^{x(\tau)=x} \mathcal{D}[x] \; P[x] \approx e^{-V(x,\tau|x_0)/\epsilon}$$

$$V(x,\tau|x_0) = \inf_{x(t):x(0)=x_0, x(\tau)=x} J[x]$$

Properties of I

- $I \geq 0$
- I is often convex
- Most probable value ⇔ min and zero of I
- Zero of I ⇔ Law of Large Numbers
- Parabolic minimum ⇔ Central Limit Theorem



Equilibrium statistical mechanics

Lanford (1973), Ellis (1985)

- N-particle system
- Macrostate: m

Microcanonical

- Large deviation probability:

$$P_U(m) \approx e^{-NI_U(m)}$$

- Equilibrium:

$$\inf_m I_U(m) = 0$$

- Max entropy principle

Canonical

- Large deviation probability:

$$P_\beta(m) \approx e^{-NI_\beta(m)}$$

- Equilibrium:

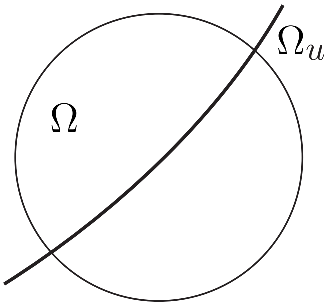
$$\inf_m I_\beta(m) = 0$$

- Min free energy principle

- Equilibrium states determined by variational principles

Microcanonical entropy

- Microstate: $\omega \in \Omega$
- Energy: $U(\omega)$
- Density of states: $\rho(u) = \int_\Omega d\omega\ \delta(U(\omega)/N - u)$
- Large deviation form: $\rho(u) \approx e^{Ns(u)}$



Gärtner-Ellis Theorem

- Free energy:

$$\varphi(\beta) = \lim_{N\to\infty} -\frac{1}{N} \ln Z(\beta), \qquad Z(\beta) = \int_\Omega d\omega\ e^{-\beta U(\omega)}$$

- If $\varphi(\beta)$ is differentiable, then

$$\rho(u) \approx e^{Ns(u)} \qquad s(u) = \inf_\beta \{\beta u - \varphi(\beta)\}$$

- Basis for Legendre transforms in statistical mechanics

Fluctuation Theorems

Farago (2002)

- Langevin dynamics:

$$\dot{v}(t)+\gamma v(t)=\underbrace{\psi(t)}_{noise}$$

- Energetics:

$$\frac{d}{dt}\left(\frac{1}{2}v^2\right)=\underbrace{-\gamma v^2}_{dissipation}+\underbrace{\psi v}_{injection}$$

- Injected energy:

$$W_\tau=\frac{1}{\tau}\int_0^\tau \psi v\ dt \qquad (Markov\ additive\ process)$$

- Large deviation probability:

$$P(W_\tau=w)\approx e^{-\tau I(w)}$$

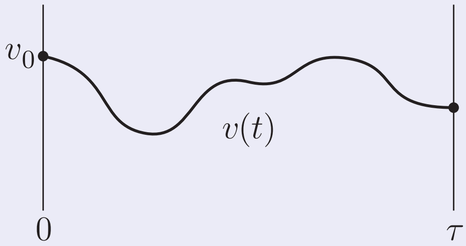
  - Large deviations of Markov processes
  - Donsker and Varadhan (1975)

Fluctuations Theorems (cont'd)

Gärtner-Ellis Theorem

- Free energy function:

$$\lambda(k)=\lim_{\tau\to\infty}\frac{1}{\tau}\ln\langle e^{\tau k W_\tau}\rangle$$



- Path integral:

$$\langle e^{\tau k W_\tau}\rangle=\int \mathcal{D}[v]\ P[v]\ e^{k\int_0^\tau v(\dot{v}+\gamma v)dt}$$

- Large deviation result:

$$P(W_\tau=w)\approx e^{-\tau I(w)}, \qquad I(w)=\max_k\{kw-\lambda(k)\}$$

- Fluctuation Theorem:

$$\frac{P(w)}{P(-w)}=e^{-\tau[I(w)-I(-w)]}\neq e^{a\tau w}$$

Multifractals

- Measure: $\mu(x)$
- Coarse-grained measure:

$$p_{\varepsilon,i} = \int_{i\text{th box}} \mu(x)$$

- Structure function:

$$Z_\varepsilon(q) = \sum_i p_{\varepsilon,i}^q \approx \varepsilon^{\tau(q)}$$

- Multifractal spectrum:

$$n_\varepsilon(\alpha) \approx \varepsilon^{-f(\alpha)}$$

Gärtner-Ellis Theorem

If $\tau(q)$ is differentiable, then

$$f(\alpha) = \inf_q \{q\alpha - \tau(q)\}$$

Thermodynamic formalism

- Map: $x_{n+1} = f(x_n)$
- Invariant density: $\rho(x)$
- Time average:

$$S_n = \frac{1}{n}\sum_{i=1}^n g(x_i) \longrightarrow \langle g(x)\rangle = \int g(x)\,\rho(x)\,dx$$

- Empirical density:

$$L_n(x) = \frac{1}{n}\sum_{i=1}^n \delta(x_i - x) \longrightarrow \rho(x)$$

- Large deviation principles:

$$P(S_n = s) \approx e^{-nI(s)}, \quad P(L_n = \ell) \approx e^{-nJ(\ell)}$$

The big picture

The mathematics of statistical mechanics
is large deviation theory

- Equilibrium statistical mechanics
  - Entropy = rate function
  - Free energy = log-generating function
  - Why variational principles for equilibrium states?
  - Why the Legendre transform?
- Nonequilibrium statistical mechanics
  - Fluctuation Theorems are large deviation results
  - Variational principles for steady states
  - Entropy = rate function
- Limits of statistical mechanics = limits of large deviation theory
  - $\lambda(k)$ non-differentiable
  - Nonconvex rate function, nonconcave entropies
  - First-order phase transitions
  - Non-exponential deviations: $\lambda(k) = 0$ or ∞

Further reading

R. S. Ellis
Entropy, Large Deviations, and Statistical Mechanics
Springer, New York, 1985.

A. Dembo and O. Zeitouni
Large Deviations Techniques and Applications
Springer, New York, 1998.

F. den Hollander
Large Deviations
American Mathematical Society, Providence, 2000.

Y. Oono
Large deviation and statistical physics
Prog. Theoret. Phys. Suppl. **99**, 165-205, 1989.

Further reading (cont'd)

R. S. Ellis
An overview of the theory of large deviations and applications to statistical mechanics
Scand. Actuar. J. **1**, 97-142, 1995.

R. S. Ellis
The theory of large deviations: From Boltzmann's 1877 calculation to equilibrium macrostates in 2D turbulence
Physica D **133**, 106136, 1999.

H. Touchette
The large deviation approach to statistical mechanics
Physics Reports, *to appear in* 2008.

http://www.maths.qmul.ac.uk/~ht
ht@maths.qmul.ac.uk