Dynamical large deviations of Markov processes

Hugo Touchette

National Institute for Theoretical Physics (NITheP)
Stellenbosch, South Africa

2017 Summer School on
Fundamental Problems in Statistical Physics
Bruneck (Brunico), Italy

Lecture notes: arxiv:1705.06492

Detailed plan

Large deviations

Quantum mechanics

Statistical physics

Markov

Italy

Steady states

Stochastic processes

Grappa

State of the world

Mountains

Nonequilibrium systems

Observables

South African vs Italian Wines

Fluctuations

Bar questions

Question 1

Is there,
can there be,
will there be
a complete theory of nonequilibrium systems?

Question 2

What is nonequilibrium?

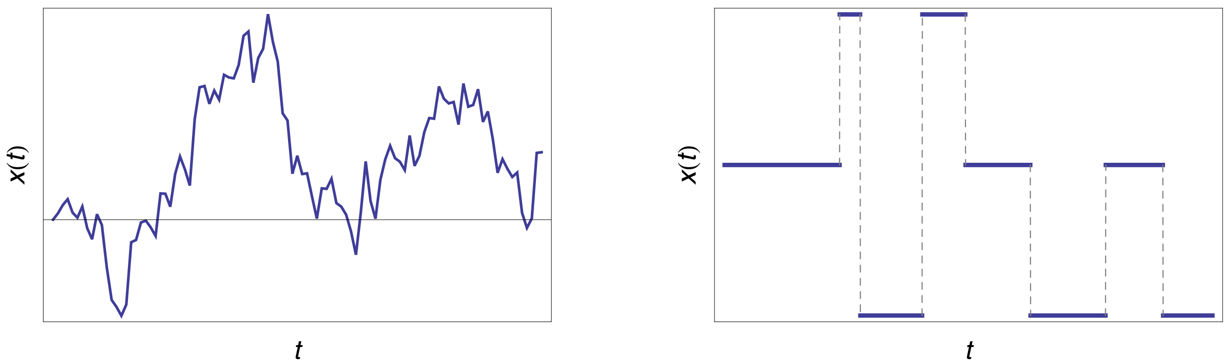
Question 3

Why is nonequilibrium more difficult than equilibrium?

Question 4

Where's the bar?

Langevin equations



- SDE:

$$dX_t = F(X_t)dt + \sigma dW_t, \qquad X_t \in \mathbb{R}^n, W_t \in \mathbb{R}^m$$

  - Drift: $F(x) \in \mathbb{R}^d$
  - Noise matrix: σ $(n \times m)$

- Noisy ODE:

$$\dot{X}_t = F(X_t) + \sigma \xi_t$$

- White noise:

$$\langle \xi_t \rangle = 0, \qquad \langle \xi_t^i \xi_{t'}^j \rangle = \delta_{ij}\delta(t - t'),$$

Fokker-Planck equation

- Fokker-Planck equation:

$$\partial_t p(x,t) = L^\dagger p(x,t), \qquad p(x,t) = P(X_t = x)$$

- Fokker-Planck generator:

$$L^\dagger = -\nabla \cdot F + \frac{1}{2}\nabla \cdot D\nabla$$

- Conservation equation:

$$\partial_t p(x,t) + \nabla \cdot J_t(x) = 0$$

- Fokker-Planck current:

$$J_t(x) = F(x)p(x,t) - \frac{D}{2}\nabla p(x,t)$$

- Stationary distribution:

$$L^\dagger p_s = 0$$

Examples

Kramers or underdamped Langevin equation

$$\begin{aligned} dq_t &= \frac{p_t}{m}dt \\ dp_t &= \left(-\nabla V(q_t) + \phi_t - \Gamma\frac{p_t}{m}\right)dt + \sqrt{2\Gamma/\beta}\, dW_t \end{aligned}$$

Overdamped Langevin equation

$$dq_t = \Gamma^{-1}(-\nabla V + \phi_t)dt + \sqrt{2\Gamma^{-1}/\beta}\, dW_t.$$

Gradient SDEs

$$dX_t = -\nabla U(X_t)dt + \sigma dW_t$$

Stationary distribution:

$$p_s(x) = c\, e^{-2U(x)/\varepsilon^2}$$

Examples (cont'd)

Linear diffusions

$$dX_t = -MX_t dt + \sigma dW_t$$

- Stationary distribution:

$$p_s(x) = \sqrt{\frac{\det C}{2\pi}} \exp\left(-\frac{1}{2}x \cdot C^{-1}x\right)$$

- Lyapunov equation:

$$D = MC + CM^T, \qquad D = \sigma\sigma^T$$

- Ornstein-Uhlenbeck process:

$$dX_t = -\gamma X_t dt + \sigma dW_t$$

Equilibrium vs nonequilibrium

Equilibrium

- Example: Gradient SDE
- Detailed balance
- Reversible process
- $J_s(x) = 0$
- Spectrum of L real
- Has no "rotation"

Nonequilibrium

- Example: Non-gradient SDE
- No detailed balance
- Nonreversible process
- $J_s(x) \neq 0$ but $\nabla \cdot J_s = 0$
- Spectrum of L complex
- Has "rotation"

Question 5

Why study Markov processes?

Examples

- Linear process:

$$dX_t = -MX_t dt + \sigma dW_t$$

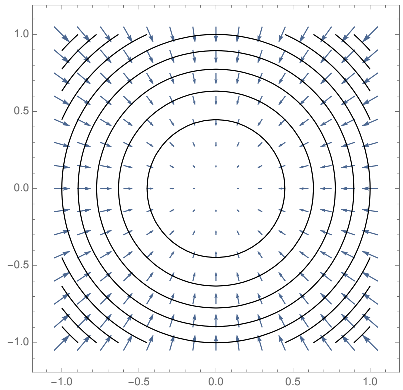
- Noise matrix: $\sigma = \epsilon \mathbb{1}$
- Gradient system: $M = \mathbb{1}$

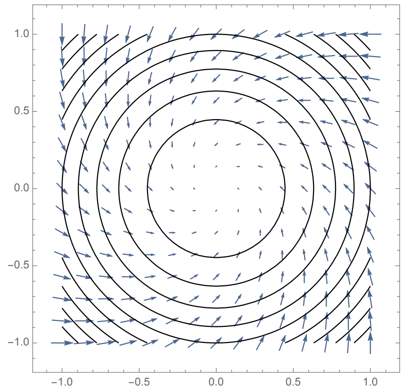
$$F = -\nabla U, \qquad U(x) = \frac{x^2}{2}$$

- Non-gradient system:

$$M = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

- Depends in general on M and $D = \sigma\sigma^T$





Evolution of expectations

- Average or expectation:

$$\langle f(X_t)\rangle = E[f(X_t)] = \int f(x)\, p(x,t)\, dx$$

- Evolution:

$$\partial_t \langle f(X_t)\rangle = \langle (Lf)(X_t)\rangle$$

- Generator:

$$L = (L^\dagger)^\dagger = F \cdot \nabla + \frac{1}{2}\nabla \cdot D\nabla$$

- $L \neq L^\dagger$ if $F \neq 0$

Scrödinger picture

$$\partial_t p(x,t) = L^\dagger p(x,t)$$

Heisenberg picture

$$\partial_t \langle f(X_t)\rangle = \langle (Lf)(X_t)\rangle$$

Non-Hermitian operators

- Inner or scalar product:

$$\langle f(X)\rangle = \int p(x)f(x)\,dx = \langle p, f\rangle$$

- Duality or adjoint:

$$\langle p, Lf\rangle = \langle L^\dagger p, f\rangle$$

- Fokker-Planck equation:

$$\partial_t\langle f(X_t)\rangle = \langle p_t, Lf\rangle = \langle L^\dagger p_t, f\rangle$$

- Differential operator:

$$\langle p, \nabla f\rangle = \int p\nabla f\,dx = \int pdf = \underbrace{pf|_{\text{boundary}}}_{=0} - \int fdp$$

  - $\nabla^\dagger = -\nabla$ (skew symmetric)
  - $\Delta^\dagger = \Delta$

Non-Hermitian operators (cont'd)

- Direct problem:

$$Lv(x) = \lambda v(x)$$

- Dual problem:

$$L^\dagger u(x) = \beta u(x), \qquad \beta = \lambda^*$$

- Orthonormal basis:

$$\langle u_i, v_j\rangle = \int u_i^*(x)v_j(x)\,dx = \delta_{ij}$$

- Completeness:

$$\delta(x - x') = \sum_i u_i^*(x)v_i(x')$$

Comparison with quantum mechanics

	Markov	**Quantum**
State	X_t	$\lvert\psi(t)\rangle$ or $\psi(x,t)$
Distribution	$p(x,t)$	$\lvert\psi(x,t)\rvert^2$
Evolution	Fokker-Planck	Schrödinger
Generator	L	H
Propagator	$U(t)=e^{L^\dagger t}$	$U(t)=e^{-iHt/\hbar}$
Inner product	$\langle p,f\rangle$	$\langle\psi,\psi\rangle=\langle\psi\lvert\psi\rangle$
Duality	$\langle p,Af\rangle=\langle A^\dagger p,f\rangle$	$\langle\psi,A\psi\rangle=\langle A\psi,\psi\rangle=\langle\psi\lvert A\rvert\psi\rangle$
Self-adjoint?	Not necessarily	Always (closed systems)

Dynamical observables

- General observable

$$A_T=\frac{1}{T}\int_0^T f(X_t)dt+\frac{1}{T}\int_0^T g(X_t)\circ dX_t$$

- Increment: $dX_t=X_{t+dt}-X_t$
- Stratonovich product:

$$g(X_t)\circ dX_t=g(X_t+dX_t/2)dX_t=\underbrace{g(X_t)dX_t}_{\text{Ito}}+\frac{g'(X_t)}{2}dt$$

- Interpretation:

$$A_T=\underbrace{\frac{1}{T}\int_0^T f(X_t)dt}_{\text{state}}+\underbrace{\frac{1}{T}\int_0^T g(X_t)\dot{X}_t dt}_{\text{velocity}}$$

Examples

- Potential energy

$$\Delta U_T = U(X_T) - U(X_0) = \int_0^T \nabla U(X_t) \circ dX_t$$

- Work:

$$W_T = \int_0^T F(X_t) \circ dX_t$$

- Entropy production:

$$\Sigma_T = 2 \int_0^T (D^{-1} F(X_t)) \circ dX_t$$

- Empirical density:

$$\rho_T(x) = \frac{1}{T} \int_0^T \delta(X_t - x)\, dt \to p_s(x)$$

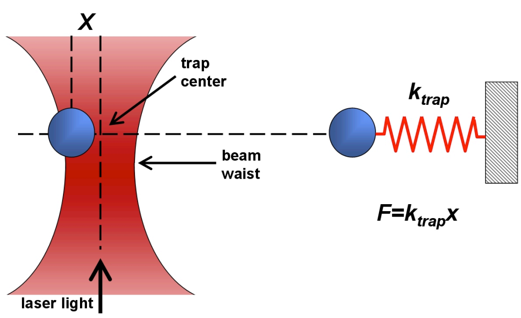
- Empirical current:

$$J_T(x) = \frac{1}{T} \int_0^T \delta(X_t - x) \circ dX_t \to J_s(x)$$

Example: Pulled Brownian particle

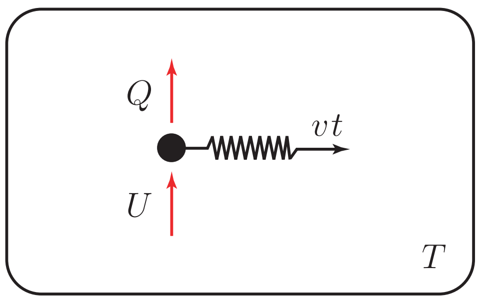
- Glass bead in water
- Laser tweezers
- Langevin dynamics:

$$m\ddot{x}(t) = \underbrace{-\alpha\dot{x}}_{\text{drag}} \underbrace{-k[x(t) - vt]}_{\text{spring force}} + \underbrace{\xi(t)}_{\text{noise}}$$



- Fluctuating work:

$$\underbrace{W_T}_{\text{work}} = \underbrace{\Delta U}_{\text{potential}} + \underbrace{Q_T}_{\text{heat}}$$



- Work observable:

$$W_T = -kv \int_0^T [x(t) - vt] dt$$

- Work distribution: $P(W_T = w)$

Large deviation theory

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

Large deviation principle (LDP)

$$P(A_T = a) \approx e^{-TI(a)}$$

- Meaning of $\approx$:

$$\ln P(a) = -TI(a) + o(T)$$

$$\lim_{T\to\infty} -\frac{1}{T}\ln P(a) = I(a)$$

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

Goals of large deviation theory

1. Prove large deviation principle exists
2. Calculate rate function

Varadhan's Theorem

- LDP:

$$P(A_T = a) \approx e^{-TI(a)}$$

- Exponential expectation:

$$\langle e^{Tf(A_T)}\rangle = \int e^{Tf(a)}\, P(A_T = a)\, da$$

- Limit functional:

$$\lambda(f) = \lim_{T\to\infty} \frac{1}{T}\ln\langle e^{Tf(A_T)}\rangle$$

S. R. Srinivasa Varadhan

Abel Prize 2007

Theorem: Varadhan (1966)

$$\lambda(f) = \max_a\{f(a) - I(a)\}$$

Special case: $f(a) = ka$

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

Gärtner-Ellis Theorem

Scaled cumulant generating function (SCGF)

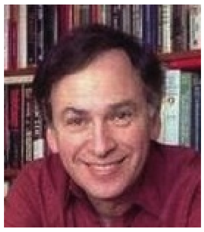
$$\lambda(k) = \lim_{T\to\infty} \frac{1}{T} \ln\langle e^{TkA_T}\rangle, \qquad k \in \mathbb{R}$$

Theorem: Gärtner (1977), Ellis (1984)

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

1. LDP:
$$P(A_T = a) \approx e^{-TI(a)}$$
2. Rate function:
$$I(a) = \max_k \{ka - \lambda(k)\}$$

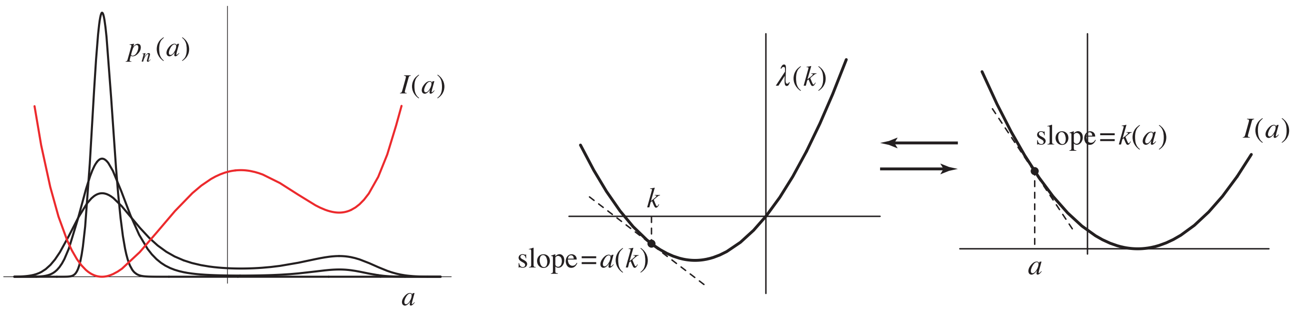
- $I(a) =$ Legendre transform of $\lambda(k)$



Richard S. Ellis

J. Gärtner

Important properties



- Duality:
$$\lambda'(k) = a \quad \Leftrightarrow \quad I'(a) = k$$
- Mean:
$$\lambda'(0) = \mu \quad \Leftrightarrow \quad I(\mu) = 0$$
- Variance:
$$\text{var}(A_T) \sim \frac{\lambda''(0)}{T}$$
- CLT:
$$I(a) = \frac{(a-\mu)^2}{2\sigma^2} + O(a^3)$$

Examples

- Sample mean:

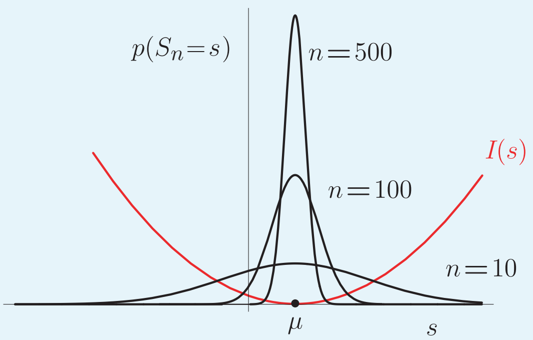
$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad X_i \sim p(x), \quad \text{IID}$$

- SCGF:

$$\lambda(k) = \lim_{n\to\infty} \frac{1}{n} \ln E[e^{nkS_n}] = \ln E[e^{kX}]$$

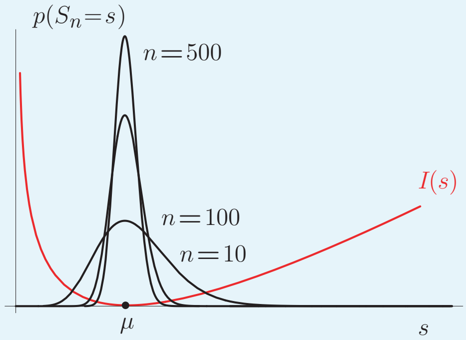
Gaussian

$$\lambda(k) = \mu k + \tfrac{\sigma^2}{2}k^2, \quad k \in \mathbb{R}$$
$$I(s) = \tfrac{1}{2\sigma^2}(s-\mu)^2, \quad s \in \mathbb{R}$$



Exponential

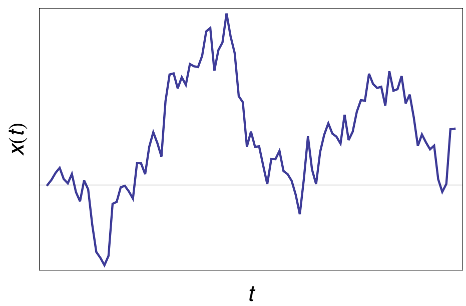
$$\lambda(k) = -\ln(1-\mu k), \quad k < \tfrac{1}{\mu}$$
$$I(s) = \tfrac{s}{\mu} - 1 - \ln \tfrac{s}{\mu}, \quad s > 0$$



Large deviations of Markov processes

- Process:

$$dX_t = F(X_t)dt + \sigma dW_t$$

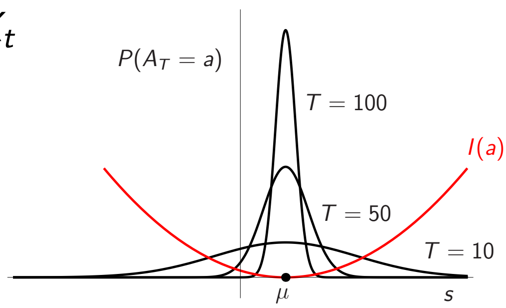


- Observable:

$$A_T = \frac{1}{T}\int_0^T f(X_t)\, dt + \frac{1}{T}\int_0^T g(X_t) \circ dX_t$$

- LDP:

$$P(A_T = a) \approx e^{-T I(a)}, \quad T \to \infty$$



Question 6

Why study fluctuations?

Question 7

Do all observables satisfy the LDP?

Dual problem

Scaled cumulant function

$$\lambda(k) = \lim_{T\to\infty} \frac{1}{T} \ln E[e^{TkA_T}]$$

Gärtner-Ellis Theorem

$\lambda(k)$ differentiable, then

1. LDP for A_T
2. $I(a) = \sup_k \{ka - \lambda(k)\}$

Feynman-Kac-Perron-Frobenius

For Markov processes,

$$\lambda(k) = \zeta_{\max}(\mathcal{L}_k)$$

- SCGF = dominant eigenvalue
- Tilted (twisted) operator:

$$\mathcal{L}_k = F\cdot(\nabla + kg) + \frac{1}{2}(\nabla + kg)\cdot D(\nabla + kg) + kf$$

- $\mathcal{L}_{k=0} = L$

Spectral problem

$\mathcal{L}_k$ non-Hermitian in general

- Spectral problem:

$$\mathcal{L}_k r_k(x) = \lambda(k) r_k(x)$$

- Dual problem:

$$\mathcal{L}_k^\dagger l_k(x) = \lambda(k) l_k(x)$$

- Boundary condition:

$$r_k(x) l_k(x) \xrightarrow{|x|\to\infty} 0$$

- Normalization:

$$\int r_k(x)\, l_k(x)\, dx = 1, \qquad \int l_k(x)\, dx = 1$$

- Extract dominant eigenvalue

Question 8

Why must wavefunctions decay at infinity?

Feynman-Kac formula

- Generating function:

$$G(x,t) = \left\langle e^{\int_0^t c(X_s)ds} \right\rangle_x = \langle e^{tkA_t} \rangle_x$$

- FK equation:

$$\partial_t G(x,t) = \mathcal{L}_c G(x,t), \qquad \mathcal{L}_c = L + c$$

- Initial condition: $G(x,0) = 1$
- Solution:

$$G(x,t) = (e^{t\mathcal{L}_k} 1)(x)$$

- Spectral decomposition:

$$G(x,t) = \sum_i e^{\zeta_i t} r_k^{(i)}(x) \sim \text{const} \cdot e^{\zeta_{\max} t}$$

Dominant eigenvalue is real (Perron-Frobenius)

Spectral problem

Equilibrium

- X_t reversible, g gradient, f arbitrary
- $\mathcal{L}_k$ non-Hermitian but conjugated to Hermitian
- Real spectrum (quantum problem)
- Easy

Nonequilibrium

- X_t nonreversible OR g non-gradient, f arbitrary
- $\mathcal{L}_k$ non-Hermitian, not conjugated to Hermitian
- Complex spectrum
- Full spectral problem
- Not so easy

Symmetrization

- Symmetrization:

$$\mathcal{H}_k = p_s^{1/2}\mathcal{L}_k p_s^{-1/2} = \sqrt{p_s}\mathcal{L}_k \frac{1}{\sqrt{p_s}}$$

- Hamiltonian:

$$\mathcal{H}_k = \frac{\varepsilon^2}{2}\Delta - V_k$$

- Effective potential:

$$V_k(x) = \frac{|\nabla U(x)|^2}{2\varepsilon^2} - \frac{\Delta U(x)}{2} - kf(x)$$

- Spectral problem:

$$\mathcal{H}_k\psi_k = \lambda(k)\psi_k,$$

- Eigenfunctions:

$$\psi_k(x) = \sqrt{p_s(x)}r_k(x), \qquad \psi_k(x) = l_k(x)/\sqrt{p_s(x)}$$

- Normalization:

$$\psi(x)^2 \stackrel{|x|\to\infty}{\longrightarrow} 0, \qquad \int \psi(x)^2\, dx = 1$$

Example

- Process:

$$dX_t = -\gamma X_t dt + \sigma dW_t$$

- Observable:

$$A_T = \frac{1}{T}\int_0^T X_t\, dt$$

- Tilted generator:

$$\mathcal{L}_k = L + kf = -\gamma x\frac{d}{dx} + \frac{\sigma^2}{2}\frac{d^2}{dx^2} + kx, \qquad k \in \mathbb{R}$$

- Stationary distribution:

$$p_s(x) = \sqrt{\frac{\gamma}{\pi\varepsilon^2}}\, e^{-\gamma x^2/\varepsilon^2} \propto e^{-2U(x)/\varepsilon^2}$$

- Effective hamiltonian:

$$\mathcal{H}_k = \frac{\varepsilon^2}{2}\frac{d^2}{dx^2} - \frac{\gamma^2 x^2}{2\varepsilon^2} + \frac{\gamma}{2} + kx.$$

Example (cont'd)

- Dominant eigenvalue:

$$\lambda(k) = \frac{\varepsilon^2 k^2}{2\gamma^2}$$

- Rate function:

$$I(a) = \frac{\gamma^2 a^2}{2\varepsilon^2}$$

- Eigenfunctions:

$$\psi_k(x) = \left(\frac{\gamma}{\pi\varepsilon^2}\right)^{1/4} \exp\left(-\frac{\gamma\left(x-\varepsilon^2 k/\gamma^2\right)^2}{2\varepsilon^2}\right)$$

$$r_k(x) = \exp\left(\frac{kx}{\gamma} - \frac{3\varepsilon^2 k^2}{4\gamma^3}\right)$$

$$l_k(x) = \sqrt{\frac{\gamma}{\pi\varepsilon^2}} \exp\left(-\frac{\gamma\left(2x-\varepsilon^2 k/\gamma^2\right)^2}{4\varepsilon^2}\right)$$

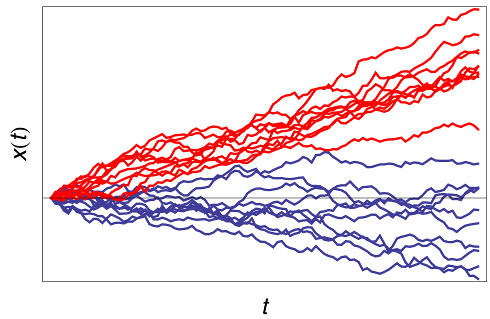
How are fluctuations created?

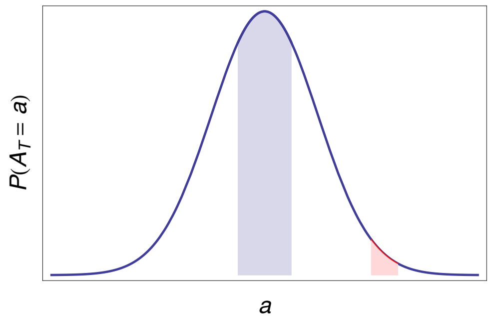
Physical problem

- Stochastic process: X_t
- Observable: A_T
- Look at trajectories leading to $A_T = a$
- Find effective process describing these trajectories

Mathematical problem

- Markov process: $\{X_t\}_{t=0}^T$
- Conditioned process: $X_t | A_T = a$
- Is it a Markov process?
- Construct its generator





Fluctuation process

Driven diffusion

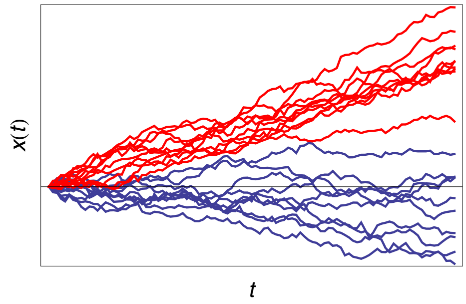
$$d\hat{X}_t = F_k(\hat{X}_t)dt + \sigma dW_t$$

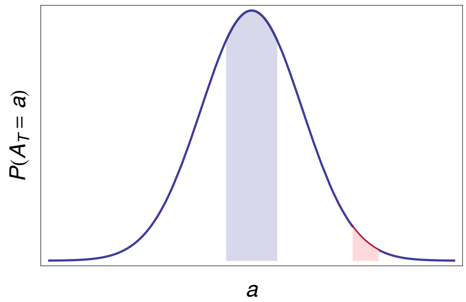
- Modified drift:

$$F_k(x) = F(x) + D(kg + \nabla \ln r_k), \quad I'(a) = k$$

Conditioning

$$\underbrace{X_t \,|\, A_T = a}_{\text{conditioned microcanonical}} \overset{T\to\infty}{\cong} \underbrace{\hat{X}_t}_{\text{driven canonical}}$$





Effective process creating the fluctuation

[Chetrite HT, PRL 2013, AHP 2015, JSTAT 2015]

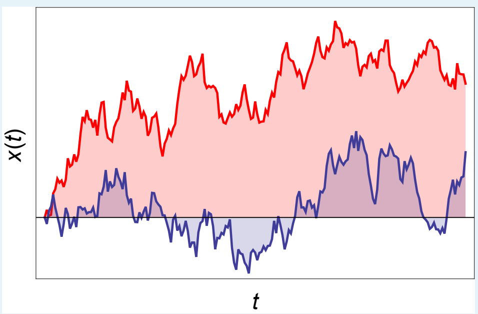
Examples

$$dX_t = -\gamma X_t dt + \sigma dW_t$$

Area

$$A_T = \frac{1}{T}\int_0^T X_t dt = a$$

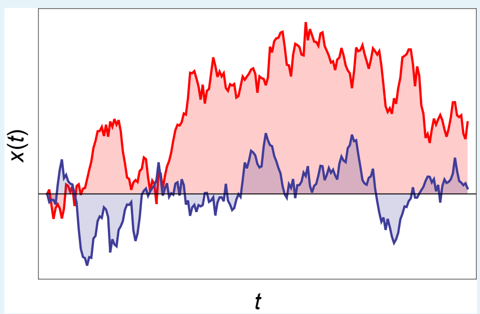
- $F_a(x) = -\gamma x + \gamma a$
- Modified drift



Empirical variance

$$A_T = \frac{1}{T}\int_0^T X_t^2 dt = a$$

- $F_a = -\sigma^2 x/(2a)$
- Modified friction



Research topics

- Jump processes
- Many particles dynamics
- Interacting particle systems
- Interacting diffusions
- Macroscopic fluctuation theory (stochastic PDEs)
- Fluctuation relations
- Level 2.5 large deviations
- Stochastic thermodynamics
- Entropy production
- Disordered systems (spin glasses, random graphs)
- Quantum systems (thermalization, nonequilibrium)
- Numerical large deviations
- Large deviation simulations

Final question

Does statistical mechanics have any relevance/future?

References

H. Touchette
The large deviation approach to statistical mechanics
Physics Reports 478, 1-69, 2009

H. Touchette
Introduction to large deviations: Theory, applications, simulations
2011 Oldenburg School Lecture Notes, arxiv:1106.4146

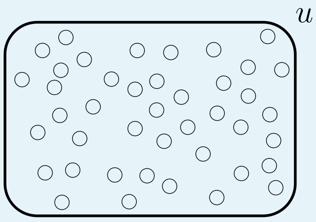
H. Touchette
Introduction to dynamical large deviations of Markov processes
2017 FPSP Lecture Notes, arxiv:1705.06492

More at
www.physics.sun.ac.za/~htouchette/ldt
www.physics.sun.ac.za/~htouchette/ldtcourse/

Extra: Analogy with equilibrium statistical mechanics

Microcanonical

[Einstein (1910)]



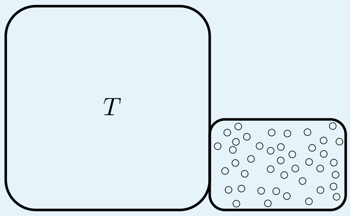
$$P_u(M_N = m) = e^{S(u,m)/k_B}$$

- Extensivity: $S \sim N$
- LDP:

$$P_u(M_N = m) \approx e^{-NI_u(m)}$$

Canonical

[Landau (1937)]



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

- Extensivity: $F \sim N$
- LDP:

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

- Exponential concentration of probability
- Equilibrium states = minima and zeros of I