Théorie des grandes déviations: Des mathématiques à la physique

Hugo Touchette

Department of Mathematical Sciences
Stellenbosch University, Afrique du Sud

Département de physique, Université de Sherbrooke
7 novembre 2018

Plan

Thèmes

- États typiques
- Fluctuations
- Plusieurs particules/composantes

Résumé

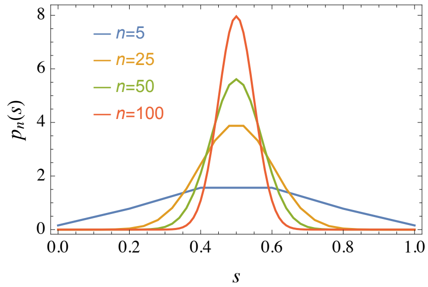
- Un peu d'histoire
- Bases des grandes déviations
- Systèmes à l'équilibre
- Systèmes hors équilibre

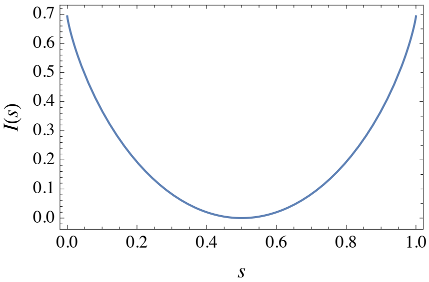
⋮	⋮
Lewis (80s) Graham (80s)	Ellis (1984)
	Gärtner (1977)
	Freidlin-Wentzell (70s)
Lanford (1973)	Donsker-Varadhan (70s)
	Sanov (1957)
Onsager (1953)	Cramér (1938)
Einstein (1910)	
Boltzmann (1877)	

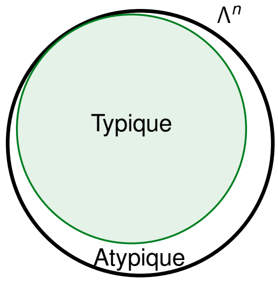
Pile ou face

P	F	F	P	F	$\cdots$
1	0	0	1	0	$\cdots$
X_1	X_2	X_3	X_4	X_5	$\cdots$

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







Distribution

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

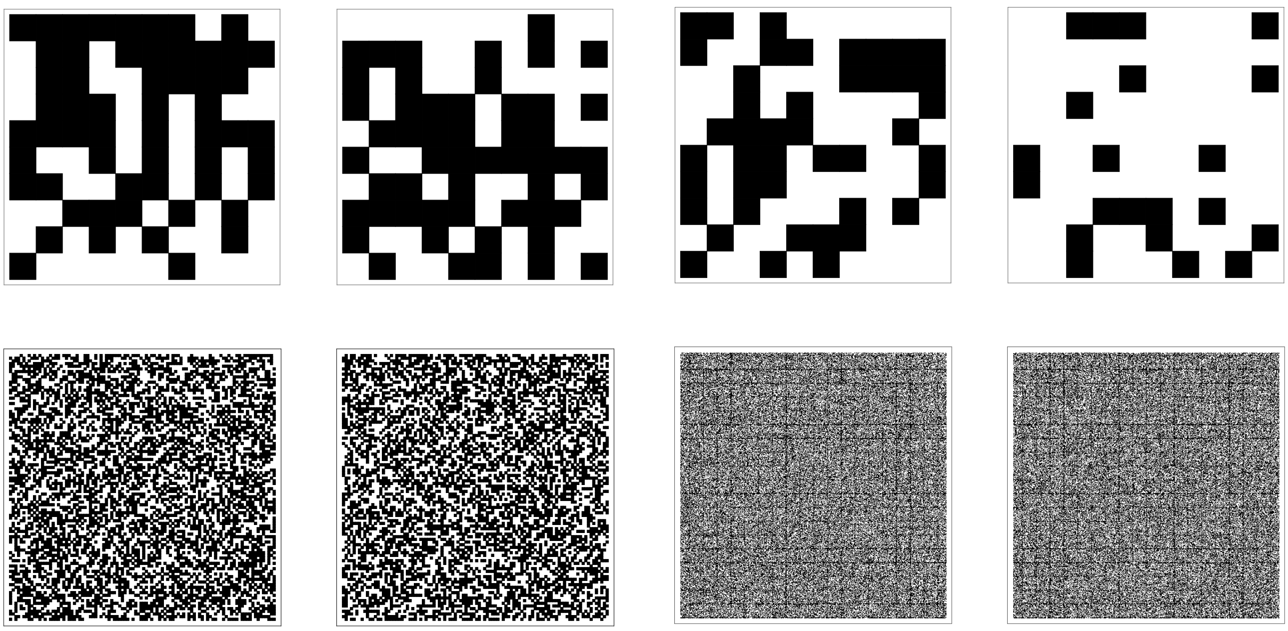
$$I(s) = s\log s + (1-s)\log(1-s) + \log 2$$

Séquences typiques

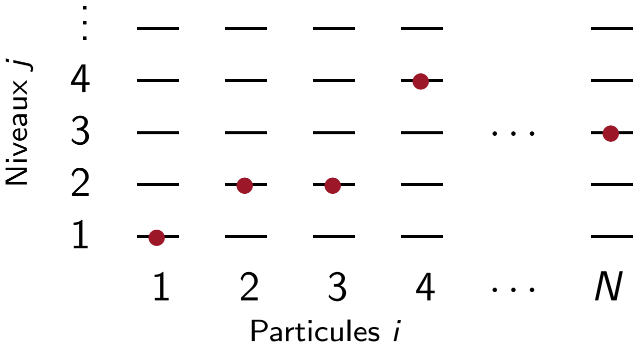
$$\#\{\vec{X} : S_n = 0.5\} \approx 2^n$$

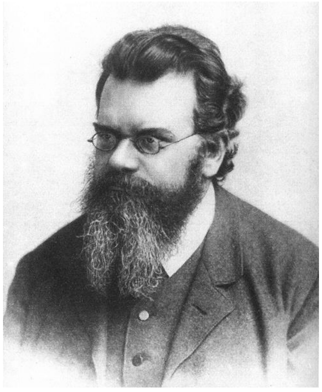
Images aléatoires: la télé de Bernoulli

- Pixel aléatoire ($p = 0.5$)
- $n \times n$ pixels
- $2^{n\times n}$ configurations



Boltzmann (1877)





- Distribution d'énergie:

$$w_j = \#\text{ particules au niveau } j$$

- Distribution multinomiale:

$$\ln \frac{N!}{\prod_j w_j!} \approx -N \sum_j w_j \ln w_j = Ns(\mathbf{w})$$

- $P(\mathbf{w}) \approx e^{Ns(\mathbf{w})}$

188 42. Bezieh. zw. zweitem Hauptsatze u. Wahrscheinlichkeitsrechnung.

hinzutritt, welche ein Minimum werden soll; führen wir ferner statt der Bedingung, daß der Nenner ein Minimum werden muß, die gleichbedeutende ein, daß dessen Logarithmus ein Minimum werden muß: dann erhalten wir für das Wärmegleichgewicht die Bedingung, daß die Größe

$$M = w_0 l w_0 + w_1 l w_1 + w_2 l w_2 + \ldots - n$$

ein Minimum sei, während gleichzeitig wieder die beiden Bedingungen erfüllt sein müssen:

(20) $n = w_0 + w_1 + w_2 + \ldots,$

(21) $L = \varepsilon w_1 + 2\varepsilon w_2 + 3\varepsilon w_3 + \ldots,$

welche mit den Gleichungen (1) und (2) des ersten Abschnittes identisch sind. Führen wir hier zunächst statt der Größen w

Einstein (1910)

- Généralise Boltzmann
- Macro-état: M_N
- Densité d'états (complexion):

$$W(m) = \#\text{ micro-états avec } M_N = m$$

Postulat d'Einstein

$$W(m) = e^{Ns(m)}$$

- Probabilité:

$$P(m) = e^{N[s(m)-s(m^*)]}$$

- Équilibre: $s(m^*)$ max

Aus Gleichung (1) folgt

$$W = \text{konst.}\, e^{\frac{N}{R} S}.$$

Diese Gleichung gilt der Größenordnung nach, wenn man jedem Zustand Z ein kleines Gebiet, von der Größenordnung wahrnehmbarer Gebiete, zuordnet. Die Konstante bestimmt sich der Größenordnung nach durch die Erwägung, daß W für den Zustand des Entropiemaximums (Entropie S_0) von der Größenordnung Eins ist, so daß man der Größenordnung nach hat

$$W = e^{\frac{N}{R}(S - S_0)}.$$

Cramér (1938)

- Somme aléatoire:

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

- Fonction cumulant:

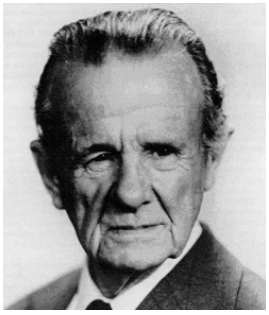
$$\lambda(k) = \ln E[e^{kX}] = \int_{\mathbb{R}} p(x)\, e^{kx}\, dx$$

- Densité de probabilité:

$$P(S_n = s) = e^{-nI(s)} \frac{1}{\sqrt{n}} \left(b_0 + \frac{b_1}{n} + \cdots \right)$$

- Fonction de décroissance:

$$I(s) = \max_{k\in\mathbb{R}} \{ks - \lambda(k)\}$$



Harald Cramér (1893-1985)

On a d'ailleurs $b_0 = \frac{1}{h\bar{\sigma}\sqrt{2\pi}}$. En introduisant dans (21), on obtient donc

(28) $$1 - F_n\left(\frac{\bar{m}\sqrt{n}}{\sigma}\right) = \frac{1}{\sqrt{n}} e^{-(h\bar{m} - \log R)n}\left[b_0 + \frac{b_1}{n} + \cdots + \frac{b_{k-1}}{n^{k-1}} + O\left(\frac{1}{n^k}\right)\right].$$

Soit maintenant c un nombre donné tel que $0 < c < C_1$, et prenons h égal à la racine (unique) positive de l'équation (27). En introduisant cette valeur dans (28) et en posant

(29) $$\alpha = h\bar{m} - \log R$$

(où l'on voit facilement que α est toujours positif), on a le théorème suivant.

Sanov (1957)

- Séquence de variables aléatoires:

$$X_1, X_2, \ldots, X_n \qquad X_i \sim p(x)$$

- Distribution empirique:

$$L_n(x) = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i,x}$$

$$P(L_n = \rho) \approx e^{-nD(\rho||p)}$$

- Entropie relative:

$$D(\rho||p) = \int dx\, \rho(x) \ln \frac{\rho(x)}{p(x)}$$

- Loi des grands nombres: $L_n \to \rho$

Ivan Nikolaevich Sanov (1919-1968)

Теорема 10. *Пусть $F(x)$ — функция распределения случайной величины ξ, $F_N(x)$ — эмпирическая функция распределения после N независимых наблюдений случайной величины ξ. Пусть $\Phi(x)$ — другая функция распределения, такая, что $\int_{-\infty}^{+\infty} \ln \frac{dF}{d\Phi} d\Phi$ существует. Пусть V_n — последовательность ε-окрестностей, содержащих $\Phi(x)$ и F-сходящихся к ней.*

Тогда

$$P(F_N \in V_n) = e^{N\left[\int_{-\infty}^{+\infty} \ln \frac{dF}{d\Phi} d\Phi + \delta_n + O\left(\frac{n \ln N}{N}\right)\right]}, \qquad (45)$$

Théorie des grandes déviations

- Variable aléatoire: A_n
- Distribution: $P(A_n = a)$

Principe de grandes déviations (PGD)

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

- Signification de $\approx$:

$$\ln P(a) = -nI(a) + o(n)$$

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

- Fonction taux: $I(a) \geq 0$

Buts de la théorie

1. Prouver qu'un PGD existe
2. Calculer la fonction taux

Théorème de Gärtner-Ellis

Fonction cumulant limite

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

Théorème: Gärtner (1977), Ellis (1984)

Si $\lambda(k)$ is dérivable, alors

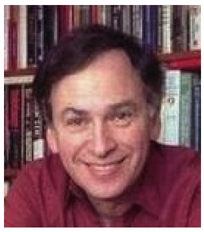
1. PGD:

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

2. Fonction taux:

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

- $I(a) =$ transformée de Legendre de $\lambda(k)$



Richard S. Ellis

J. Gärtner

Cramer's Theorem

- Somme:

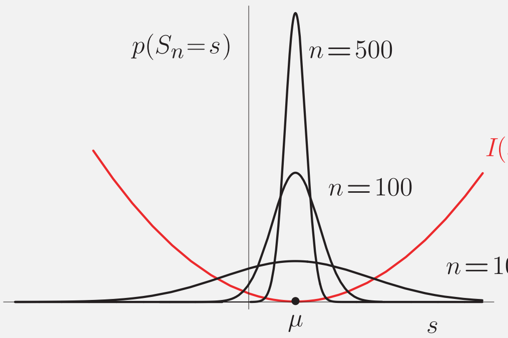
$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad X_i \sim p(x), \quad \text{IID}$$

- Fonction cumulant:

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

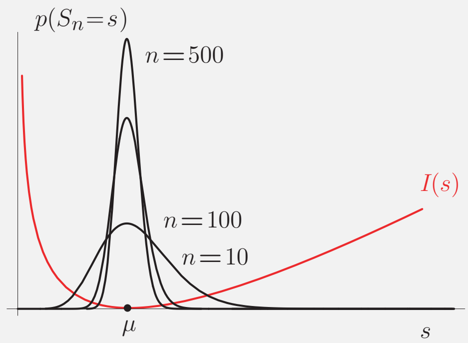
Cas Gaussien

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



Cas exponentiel

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



Théorème de Sanov

- n variables aléatoires:

$$\omega = \omega_1, \omega_2, \ldots, \omega_n, \qquad P(\omega_i = j) = p_j$$

- Fréquences empiriques:

$$L_{n,j} = \frac{1}{n}\sum_{i=1}^{n} \delta_{\omega_i,j} = \frac{\#(\omega_i = j)}{n}, \qquad \mathbf{L}_n = (L_{n,1}, L_{n,2}, \ldots)$$

Gärtner-Ellis

- Fonction cumulant:

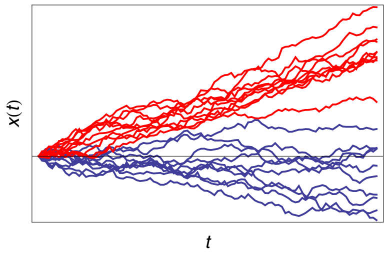
$$\lambda(\mathbf{k}) = \lim_{n\to\infty}\frac{1}{n}\ln E[e^{n\mathbf{k}\cdot\mathbf{L}_n}] = \ln\sum_j p_j\, e^{k_j}$$

- Fonction taux:

$$D(\boldsymbol{\mu}) = \inf_{\mathbf{k}}\{\mathbf{k}\cdot\boldsymbol{\mu} - \lambda(\mathbf{k})\} = \sum_j \mu_j \ln\frac{\mu_j}{p_j}$$

Au delà de l'indépendance: Processus de Markov

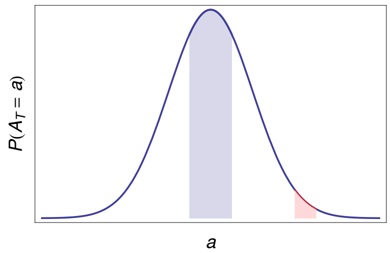
- Processus: $\{X_t\}_{t=0}^{T}$
- Observable: $A_T = \frac{1}{T}\int_0^T f(X_t)\,dt$



PGD

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

- Limite ergodique
- Donsker & Varadhan ('70)

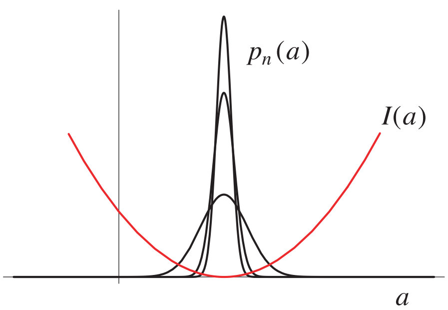


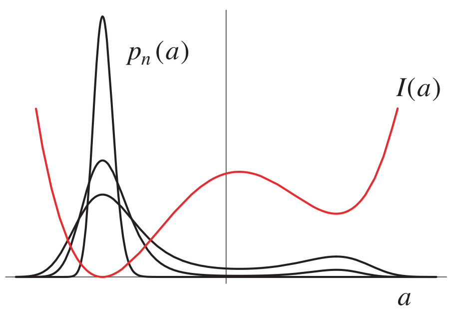
Applications

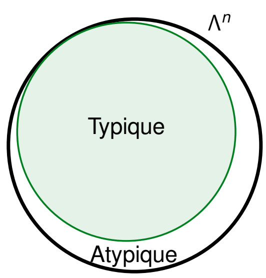
- Marches aléatoires
- Systèmes dynamiques bruités
- EDP stochastiques
- Particules en interaction
- Graphes aléatoires
- Files d'attente
- Statistiques, estimation
- Théorie de l'information

Résumé

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







- Valeur typique: Loi des grands nombres
- Forme quadratique locale: Théorème limite centrale
- Forme non-quadratique en général: Grandes déviations

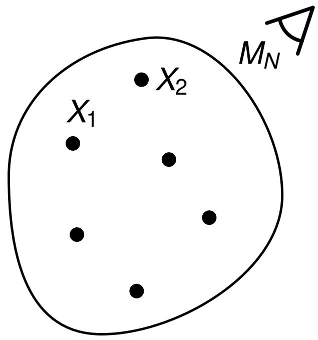
Théorie générale des états typiques et fluctuations

Systèmes à l'équilibre

- N particules
- Micro-état: $\vec{X} = X_1, X_2, \ldots, X_N$
- Ensemble: $P(\vec{X})$
- Macro-état: $M_N(\vec{X})$
- Distribution:

$$P(M_N = m) = \sum_{\vec{X}: M_N(\vec{X}) = m} P(\vec{X})$$

M_N	Macro
$\vec{X}$	Micro



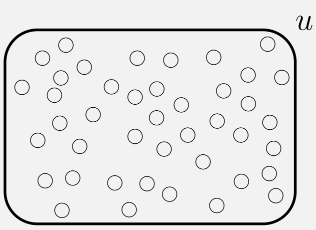
Problèmes

- $P(M_N = m)$
- Valeurs typiques de M_N (= états d'équilibre)
- Fluctuations autour de l'équilibre
- Limite thermodynamique $N \to \infty$

Grandes déviations à l'équilibre

Microcanonique

Einstein (1910)



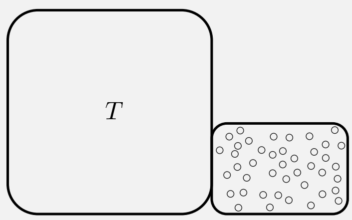
$$P_u(M_N = m) = e^{S(u,m)/k_B}$$

- Extensivité: $S \sim N$
- PGD:

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

Canonique

Landau (1937)



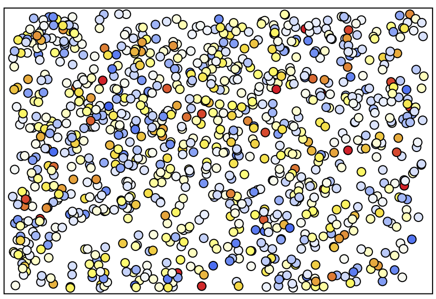
$$P_\beta(M_N = m) = e^{-F(\beta,m)}$$

- Extensivité: $F \sim N$
- PGD:

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

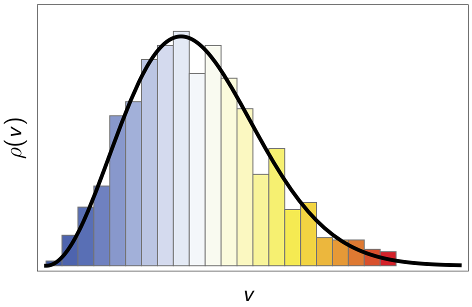
- Concentration exponentielle
- États à l'équilibre = minima et zéros de I

Distribution de Maxwell



- Énergie:

$$U_N = \sum_{i=1}^{N} \frac{v_i^2}{2m}, \qquad N \sim 10^{23}$$



- Distribution des vitesses:

$$L_N(v) = \frac{\#\text{ particles } v_i \in [v, v+\Delta v]}{N\Delta v}$$

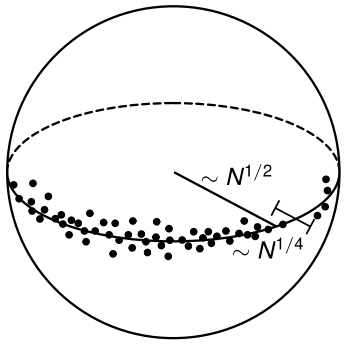
LDP

$$P(L_N = \rho) \approx e^{-NI(\rho)}$$

- Distribution d'équilibre:

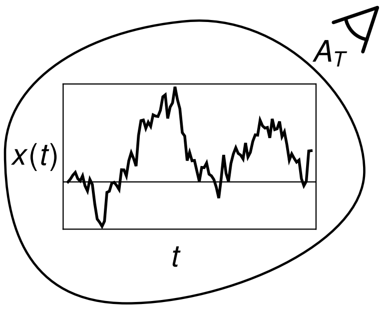
$$\rho^*(v) = c\, v^2 e^{-\frac{mv^2}{2k_B T}}$$

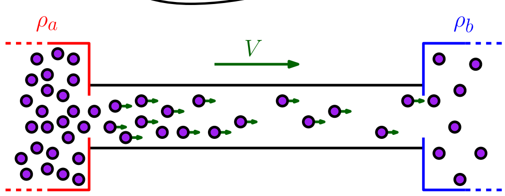
- Distribution de Maxwell



Systèmes hors équilibre

- Processus: X_t
  - Une ou plusieurs particules
  - Modèle: Processus de Markov
  - Forces externes
  - Réservoirs
- Trajectoire: $\{x_t\}_{t=0}^{T}$
- Distribution: $P[x]$
- Observable: $A_{N,T}[x]$

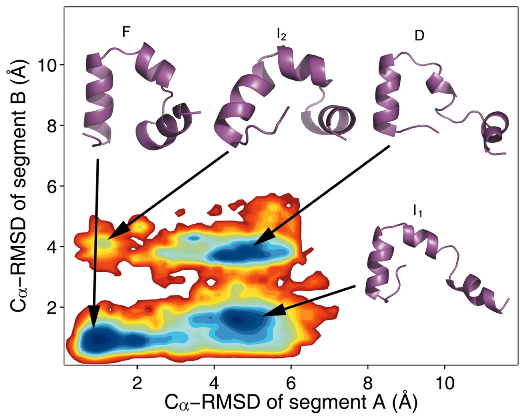


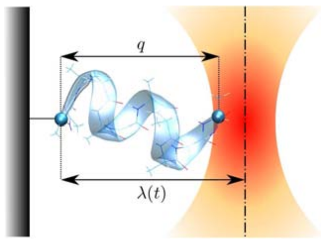


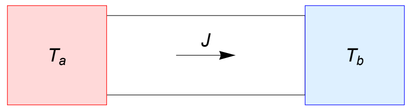
Problèmes

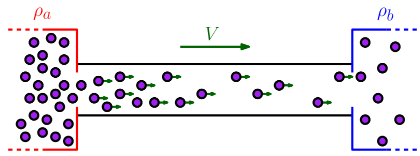
- $P(A_{N,T} = a)$
- Valeurs typiques de $A_{N,T}$ (= états stationnaires)
- Fluctuations autour des valeurs typiques
- Limites: $\quad N \to \infty \qquad T \to \infty \qquad \text{bruit} \to 0$

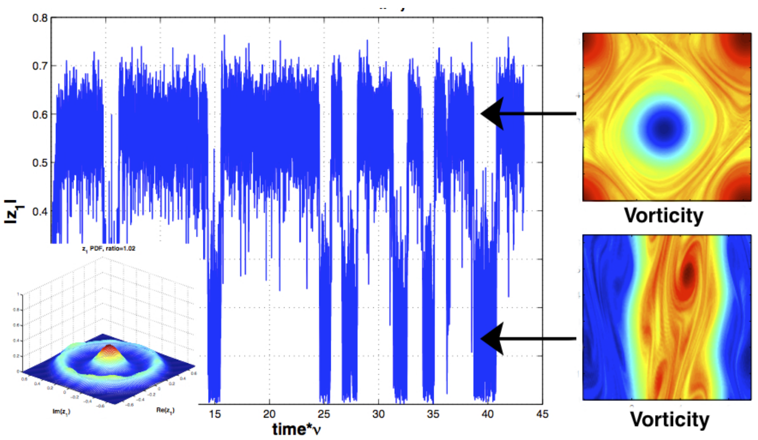
Applications

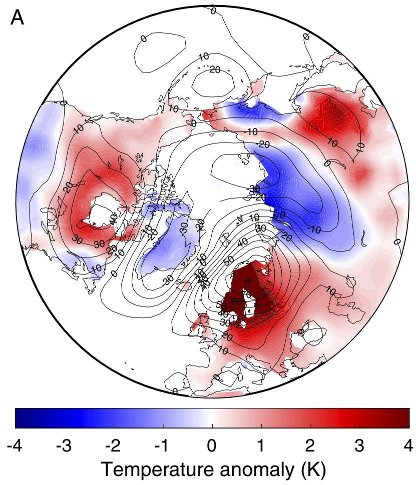












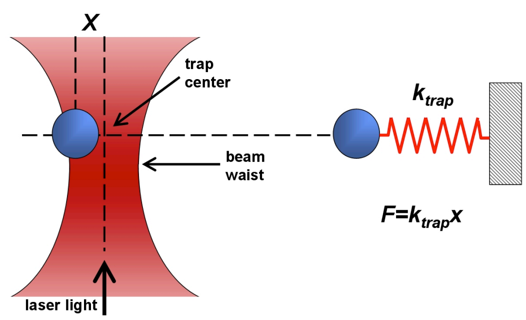
Exemple: Particule brownienne

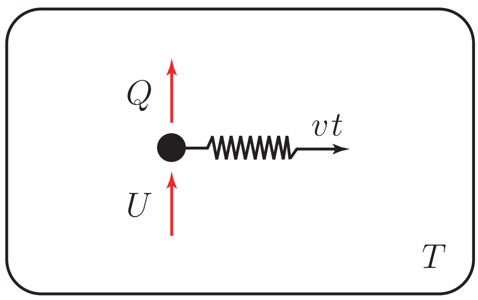
- Bille de verre dans l'eau
- Pinces laser
- Dynamique de Langevin:

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

- Travail fluctuant:

$$\underbrace{W_T}_{\text{travail}} = \underbrace{\Delta U}_{\text{potentiel}} + \underbrace{Q_T}_{\text{chaleur}}$$





PGD

$$P(W_T = w) \approx e^{-TI(w)}$$

Symétrie

$$\frac{P(W_T = w)}{P(W_T = -w)} = e^{Tcw}$$

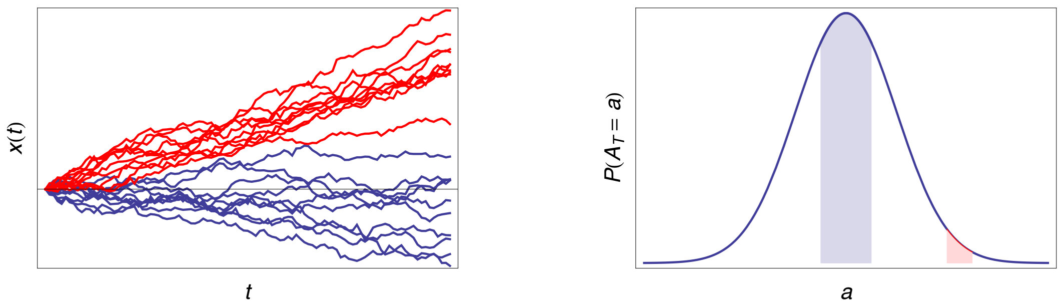
Sommaire

- Variables aléatoires — ensembles — processus stochastique
- Valeurs typiques — état d'équilibre — état stationnaire
- Limites: $N \to \infty$, $T \to \infty$, $\epsilon \to 0$
- Fonction taux $=$ entropie
- Fonction cumulant $=$ énergie libre
- Base mathématique de la physique statistique

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

`http://appliedmaths.sun.ac.za/~htouchette/`

Recherche: Comment les fluctuations sont-elles créées?



- Trajectoires associées à une fluctuation $A_T = a$
- Dynamique de ces trajectoires
- Processus de Markov générant ces trajectoires

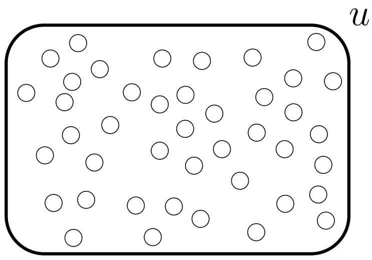
$$\underbrace{X_t \,|\, A_T = a}_{\text{processus conditionné}} \quad \equiv \quad \underbrace{\hat{X}_t}_{\text{process effectif}}$$

Entropy and free energy

- Density of states:

$$\Omega(u) = \#\ \omega \text{ with } U/N = u$$

- Large deviation form: $\Omega(u) \approx e^{Ns(u)}$



Gärtner-Ellis Theorem

$$s(u) = \min_{\beta}\{\beta u - \varphi(\beta)\}$$

- Free energy:

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

- $Z(\beta)$ = partition function = generating function
- $\varphi(\beta)$ = free energy = SCGF
- Basis of Legendre transform in thermodynamics