Weierstrass Institute for
Applied Analysis and Stochastics

Leibniz Association

The theory of the probabilities of large deviations, and applications in statistical physics

Wolfgang König
TU Berlin and WIAS Berlin

- We look at sequences $(S_n)_{n\in\mathbb{N}}$ of random variables showing an exponential behaviour

$$\mathbb{P}(S_n \approx x) \approx \mathrm{e}^{-nI(x)} \qquad \text{as } n \to \infty$$

  for any x, with $I(x)$ some rate function.

- The event $\{S_n \approx x\}$ is an event of a large deviation (strictly speaking, only if $x \neq \mathbb{E}(S_n)$).
- We make this precise, and build a theory around it.
- We give the main tools of that theory.
- We explain the relation with asymptotics of exponential integrals of the form

$$\mathbb{E}\Big[\mathrm{e}^{nf(S_n)}\Big] \approx \mathrm{e}^{n\sup[f-I]} \qquad \text{as } n \to \infty$$

  and draw conclusions.

- We show how to analyse models from statistical physics with the help of this theory.

- Let $(X_i)_{i\in\mathbb{N}}$ an i.i.d. sequence of real random variables, and consider the mean $S_n = \frac{1}{n}(X_1 + \cdots + X_n)$. Assume that X_1 has all exponential moments finite and expectation zero. Then, for any $x > 0$, the probability of $\{S_n \geq x\}$ converges to zero, according to the law of large numbers.
- What is the decay speed of this probability?

- Let $(X_i)_{i\in\mathbb{N}}$ an i.i.d. sequence of real random variables, and consider the mean $S_n = \frac{1}{n}(X_1 + \cdots + X_n)$. Assume that X_1 has all exponential moments finite and expectation zero. Then, for any $x > 0$, the probability of $\{S_n \geq x\}$ converges to zero, according to the law of large numbers.
- What is the decay speed of this probability?
- It is even exponential, as an application of the Markov inequality (*exponential Chebyshev inequality*) shows for any $y > 0$ and any $n \in \mathbb{N}$:

$$\mathbb{P}(S_n \geq x) = \mathbb{P}(\mathrm{e}^{ynS_n} \geq \mathrm{e}^{yxn}) \leq \mathrm{e}^{-yxn}\mathbb{E}[\mathrm{e}^{ynS_n}] = \mathrm{e}^{-yxn}\mathbb{E}\Big[\prod_{i=1}^n \mathrm{e}^{yX_i}\Big]$$

$$= \mathrm{e}^{-yxn}\mathbb{E}[\mathrm{e}^{yX_1}]^n = \Big(\mathrm{e}^{-yx}\mathbb{E}[\mathrm{e}^{yX_1}]\Big)^n.$$

- Let $(X_i)_{i\in\mathbb{N}}$ an i.i.d. sequence of real random variables, and consider the mean $S_n = \frac{1}{n}(X_1 + \dots + X_n)$. Assume that X_1 has all exponential moments finite and expectation zero. Then, for any $x > 0$, the probability of $\{S_n \geq x\}$ converges to zero, according to the law of large numbers.
- What is the decay speed of this probability?
- It is even exponential, as an application of the Markov inequality (*exponential Chebyshev inequality*) shows for any $y > 0$ and any $n \in \mathbb{N}$:

$$\mathbb{P}(S_n \geq x) = \mathbb{P}(\mathrm{e}^{ynS_n} \geq \mathrm{e}^{yxn}) \leq \mathrm{e}^{-yxn}\mathbb{E}[\mathrm{e}^{ynS_n}] = \mathrm{e}^{-yxn}\mathbb{E}\Big[\prod_{i=1}^{n}\mathrm{e}^{yX_i}\Big]$$

$$= \mathrm{e}^{-yxn}\mathbb{E}[\mathrm{e}^{yX_1}]^n = \Big(\mathrm{e}^{-yx}\mathbb{E}[\mathrm{e}^{yX_1}]\Big)^n.$$

- This may be summarized by saying that

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(S_n \geq x) \leq -I(x), \qquad x \in (0,\infty),$$

with rate function equal to the Legendre transform

$$I(x) = \sup_{y\in\mathbb{R}}[yx - \log\mathbb{E}(\mathrm{e}^{yX_1})].$$

Definition

We say that a sequence $(S_n)_{n\in\mathbb{N}}$ of random variables with values in a metric space $\mathcal{X}$ satisfies a large-deviations principle (LDP) with rate function $I\colon \mathcal{X} \to [0,\infty]$ if the set function $\frac{1}{n}\log \mathbb{P}(S_n \in \cdot)$ converges weakly towards the set function $-\inf_{x\in\cdot} I(x)$, i.e., for any open set $G \subset \mathcal{X}$ and for any closed set $F \subset \mathcal{X}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \in G) \geq -\inf_G I,$$

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \in F) \leq -\inf_F I.$$

- Hence, topology plays an important role in an LDP.
- Often (but not always), I is convex, and $I(x) \geq 0$ with equality if and only if $\mathbb{E}[S_n] = x$.
- I is lower semi-continuous, i.e., the level sets $\{x\colon I(x) \leq \alpha\}$ are closed. If they are even compact, then I is called good. (Many authors include this in the definition.)
- The LDP gives (1) the decay rate of the probability and (2) potentially a formula for deeper analysis.

- random walks $(\mathcal{X} = \mathbb{R})$, CRAMÉR's theorem
- LDPs from exponential moments $\Longrightarrow$ GÄRTNER-ELLIS theorem $\Longrightarrow$ occupation times measures of Brownian motions
- exponential integrals, VARADHAN's lemma $\Longrightarrow$ exponential transforms $\Longrightarrow$ CURIE-WEISS model (ferromagnetic spin system)
- small factor times Brownian motion $(\mathcal{X} = \mathcal{C}[0,1])$ $\Longrightarrow$ SCHILDER's theorem
- empirical measures of i.i.d. sequences $(\mathcal{X} = \mathcal{M}_1(\Gamma))$ $\Longrightarrow$ SANOV's theorem $\Longrightarrow$ Gibbs conditioning principle
- empirical pair measures of Markov chains $(\mathcal{X} = \mathcal{M}_1^{(\mathrm{s})}(\Gamma \times \Gamma))$ $\Longrightarrow$ one-dimensional polymer measures
- continuous functions of LDPs (contraction principle) $\Longrightarrow$ randomly perturbed dynamical systems ($(\mathcal{X} = \mathcal{C}[0,1])$, FREIDLIN-WENTZELL theory)
- empirical stationary fields $(\mathcal{X} = \mathcal{M}_1^{(\mathrm{s})}(\text{point processes}))$ $\Longrightarrow$ thermodynamic limit of many-body systems

Cramér's theorem

The mean $S_n = \frac{1}{n}(X_1 + \cdots + X_n)$ of i.i.d. real random variables $X_1, \dots, X_n$ having all exponential moments finite satisfies, as $n \to \infty$, an LDP with speed n and rate function $I(x) = \sup_{y\in\mathbb{R}}[yx - \log \mathbb{E}(\mathrm{e}^{yX_1})]$.

Cramér's theorem

The mean $S_n = \frac{1}{n}(X_1 + \cdots + X_n)$ of i.i.d. real random variables $X_1, \dots, X_n$ having all exponential moments finite satisfies, as $n \to \infty$, an LDP with speed n and rate function $I(x) = \sup_{y\in\mathbb{R}}[yx - \log \mathbb{E}(\mathrm{e}^{yX_1})]$.

Proof steps:

- The proof of the upper bound for $F = [x, \infty)$ with $x > 0$ was shown above.
- Sets of the form $(-\infty, -x]$ are handled in the same way.

Cramér's theorem

The mean $S_n = \frac{1}{n}(X_1 + \cdots + X_n)$ of i.i.d. real random variables $X_1, \dots, X_n$ having all exponential moments finite satisfies, as $n \to \infty$, an LDP with speed n and rate function $I(x) = \sup_{y\in\mathbb{R}}[yx - \log \mathbb{E}(\mathrm{e}^{yX_1})]$.

Proof steps:

- The proof of the upper bound for $F = [x, \infty)$ with $x > 0$ was shown above.
- Sets of the form $(-\infty, -x]$ are handled in the same way.
- The proof of the corresponding lower bound requires the Cramér transform:

$$\widehat{\mathbb{P}}_a(X_1 \in A) = \frac{1}{Z_a}\mathbb{E}\big[\mathrm{e}^{aX_1}\mathbb{1}\{X_1 \in A\}\big],$$

  and we see that

$$\mathbb{P}(S_n \approx x) = Z_a^n \widehat{\mathbb{E}}_a\left[\mathrm{e}^{-anS_n}\mathbb{1}\{S_n \approx x\}\right] \approx Z_a^n \mathrm{e}^{-axn}\widehat{\mathbb{P}}_a(S_n \approx x).$$

Cramér's theorem

The mean $S_n = \frac{1}{n}(X_1 + \cdots + X_n)$ of i.i.d. real random variables $X_1, \dots, X_n$ having all exponential moments finite satisfies, as $n \to \infty$, an LDP with speed n and rate function $I(x) = \sup_{y\in\mathbb{R}}[yx - \log \mathbb{E}(\mathrm{e}^{yX_1})]$.

Proof steps:

- The proof of the upper bound for $F = [x, \infty)$ with $x > 0$ was shown above.
- Sets of the form $(-\infty, -x]$ are handled in the same way.
- The proof of the corresponding lower bound requires the Cramér transform:

$$\widehat{\mathbb{P}}_a(X_1 \in A) = \frac{1}{Z_a}\mathbb{E}\big[\mathrm{e}^{aX_1}\mathbb{1}\{X_1 \in A\}\big],$$

  and we see that

$$\mathbb{P}(S_n \approx x) = Z_a^n \widehat{\mathbb{E}}_a\left[\mathrm{e}^{-anS_n}\mathbb{1}\{S_n \approx x\}\right] \approx Z_a^n \mathrm{e}^{-axn}\widehat{\mathbb{P}}_a(S_n \approx x).$$

  Picking $a = a_x$ as the maximizer in $I(x)$, then $a_x = \widehat{\mathbb{E}}_{a_x}(S_n) = x$, and we obtain

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \approx xn) = -[a_x x - \log Z_{a_x}] = -I(x).$$

- General sets are handled by using that I is strictly in/decreasing in $[0, \infty)$ / $(-\infty, 0]$.

Far-reaching extension of CRAMÉR's theorem.

We call $(S_n)_{n\in\mathbb{N}}$ exponentially tight if, for any $M > 0$, there is a compact set $K_M \subset \mathcal{X}$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \in K_M^{\mathrm{c}}) \leq -M.$$

GÄRTNER-ELLIS theorem

Let $(S_n)_{n\in\mathbb{N}}$ be an exponentially tight sequence of random variables taking values in a Banach space $\mathcal{X}$. Assume that

$$\Lambda(f) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big[\mathrm{e}^{nf(S_n)}\big], \qquad f \in \mathcal{X}^*,$$

exists and that Λ is lower semicontinuous and Gâteau differentiable (i.e., for all $f, g \in \mathcal{X}^*$ the map $t \mapsto \Lambda(f + tg)$ is differentiable at zero).

Then $(S_n)_{n\in\mathbb{N}}$ satisfies an LDP with rate function equal to the Legendre transform of Λ.

The proof is a (quite technical) extension of the above proof of CRAMÉR's theorem.

Let $B = (B_t)_{t\in[0,\infty)}$ be a Brownian motion in $\mathbb{R}^d$, and let $\mu_t(A) = \frac{1}{t}\int_0^t \mathbb{1}_{\{B_s\in A\}}\,\mathrm{d}s$ denote its normalized occupation times measure.

DONSKER-VARADHAN-GÄRTNER LDP

For any compact nice set $Q \subset \mathbb{R}^d$, the measure μ_t satisfies, as $t\to\infty$, an LDP on the set $\mathcal{M}_1(Q)$ under $\mathbb{P}(\cdot \cap \{B_s \in Q \text{ for any } s\in[0,t]\})$ with scale t and rate function

$$I_Q(\mu) = \frac{1}{2}\int |\nabla f(x)|^2\,\mathrm{d}x,$$

if $f = \frac{\mathrm{d}\mu}{\mathrm{d}x}$ exists and is smooth and satisfies zero boundary condition in Q.

- Indeed,

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}[\mathrm{e}^{t\langle g,\mu_t\rangle}\mathbb{1}_{\{B_{[0,1]}\subset Q\}}] = \lambda_1(g,Q),$$

  the principal eigenvalue of $-\frac{1}{2}\Delta + g$ in Q. The Rayleigh-Ritz formula

$$\lambda_1(g,Q) = \sup_{\|f\|_2=1}\langle(-\tfrac{1}{2}\Delta+g)f,f\rangle = \sup_{\|f\|_2=1}\big(\langle g,f^2\rangle + \tfrac{1}{2}\|\nabla f\|_2^2\big).$$

  shows that it is the Legendre transform of I_Q (substitute $f^2 = \frac{\mathrm{d}\mu}{\mathrm{d}x}$).
- There is an analogous version for continuous-time random walks.

VARADHAN's lemma

If $(S_n)_{n\in\mathbb{N}}$ satisfies an LDP with good rate function I in $\mathcal{X}$, and if $f\colon \mathcal{X}\to\mathbb{R}$ is continuous and bounded, then

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}[\mathrm{e}^{nf(S_n)}]=\sup_{x\in\mathcal{X}}(f(x)-I(x)).$$

This is a substantial extension of the well-known Laplace principle that says that $\int_0^1 \mathrm{e}^{nf(x)}\,\mathrm{d}x$ behaves to first order like $\mathrm{e}^{n\max_{[0,1]} f}$ if $f\colon[0,1]\to\mathbb{R}$ is continuous.

LDP for exponential tilts

If $(S_n)_{n\in\mathbb{N}}$ satisfies an LDP with good rate function I in $\mathcal{X}$, and if $f\colon \mathcal{X}\to\mathbb{R}$ is continuous and bounded, then we define the transformed measure

$$\mathrm{d}\widehat{\mathbb{P}}_n(S_n\in\cdot)=\frac{1}{Z_n}\mathbb{E}\big[\mathrm{e}^{nf(S_n)}\mathbb{1}_{\{S_n\in\cdot\}}\big],\qquad \text{where } Z_n=\mathbb{E}\big[\mathrm{e}^{nf(S_n)}\big].$$

Then the distributions of S_n under $\widehat{\mathbb{P}}_n$ satisfy, as $n\to\infty$, an LDP with rate function

$$I_f(x)=I(x)-f(x)-\inf[I-f].$$

A mean-field model for ferromagnetism:

- configuration space $E = \{-1, 1\}^N$
- energy $H_N(\sigma) = -\frac{1}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j$
- probability $\nu_N(\sigma) = \frac{1}{Z_{N,\beta}} \mathrm{e}^{-\beta H_N(\sigma)} 2^{-N}$.
- mean magnetisation $\overline{\sigma}_N = \frac{1}{N} \sum_{i=1}^N \sigma_i$. Then $-\beta H_N(\sigma) = F(\overline{\sigma})$ with $F(\eta) = \frac{\beta}{2}\eta^2$.
- CRAMÉR $\Longrightarrow$ LDP for $\overline{\sigma}_N$ under $[\frac{1}{2}(\delta_{-1} + \delta_1)]^{\otimes N}$ with rate function

$$I(x) = \sup_{y \in \mathbb{R}} \left[xy - \log(\frac{1}{2}(\mathrm{e}^{-y} + \mathrm{e}^y))\right] = \frac{1+x}{2} \log(1+x) + \frac{1-x}{2} \log(1-x).$$

- Corollary $\Longrightarrow$ LDP for $\overline{\sigma}_N$ under ν_N with rate function $I - F - \inf[I - F]$.
- Minimizer(s) $m_\beta \in [-1, 1]$ are characterised by

$$m_\beta = \frac{\mathrm{e}^{2\beta m_\beta} - 1}{\mathrm{e}^{2\beta m_\beta} + 1}.$$

- Phase transition: $\beta \leq 1 \Longrightarrow m_\beta = 0$ and $\beta > 1 \Longrightarrow m_\beta > 0$.

SCHILDER's theorem

Let $W = (W_t)_{t\in[0,1]}$ be a Brownian motion, then $(\varepsilon W)_{\varepsilon>0}$ satisfies an LDP on $\mathcal{C}[0,1]$ with scale ε^{-2} and rate function $I(\varphi) = \frac{1}{2}\int_0^1 |\varphi'(t)|^2\,\mathrm{d}t$ if φ is absolutely continuous with $\varphi(0) = 0$ (and $I(\varphi) = \infty$ otherwise).

Here is a heuristic proof: for $\varphi \in \mathcal{C}[0,1]$ differentiable with $\varphi(0) = 0$, for large $r \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}(\varepsilon W \approx \varphi) &\approx \mathbb{P}\big(W(i/r) \approx \tfrac{1}{\varepsilon}\varphi(i/r) \text{ for all } i = 0,1,\dots,r\big)\\ &= \prod_{i=1}^r \mathbb{P}\big(W(1/r) \approx \tfrac{1}{\varepsilon}(\varphi(i/r) - \varphi((i-1)/r))\big).\end{aligned}$$

Now use that $W(1/r)$ is normal with variance $1/r$:

$$\begin{aligned}\mathbb{P}(\varepsilon W \approx \varphi) &\approx \prod_{i=1}^r \mathrm{e}^{-\frac{1}{2}r\varepsilon^{-2}(\varphi(i/r)-\varphi((i-1)/r))^2}\\ &= \exp\Big\{-\frac{1}{2}\varepsilon^{-2}\frac{1}{r}\sum_{i=1}^r\Big(\frac{\varphi(i/r)-\varphi((i-1)/r)}{1/r}\Big)^2\Big\}.\end{aligned}$$

Using a RIEMANN sum approximation, we see that $\mathrm{e}^{-\varepsilon^{-2}I(\varphi)}$.

SANOV's theorem

If $(X_i)_{i\in\mathbb{N}}$ is an i.i.d. sequence of random variables with distribution μ on a Polish space Γ, then the empirical measure $S_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ satisfies an LDP on the set $\mathcal{X} = \mathcal{M}_1(\Gamma)$ of probability measures on Γ with rate function equal to the KULLBACK-LEIBLER entropy

$$I(P) = H(P \mid \mu) = \int P(\mathrm{d}x)\, \log \frac{\mathrm{d}P}{\mathrm{d}\mu}(x) = \int \mu(\mathrm{d}x)\, \varphi\Big(\frac{\mathrm{d}P}{\mathrm{d}\mu}(x)\Big),$$

with $\varphi(y) = y\log y$.

Sanov's theorem

If $(X_i)_{i\in\mathbb{N}}$ is an i.i.d. sequence of random variables with distribution μ on a Polish space Γ, then the empirical measure $S_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ satisfies an LDP on the set $\mathcal{X} = \mathcal{M}_1(\Gamma)$ of probability measures on Γ with rate function equal to the Kullback-Leibler entropy

$$I(P) = H(P \mid \mu) = \int P(\mathrm{d}x)\, \log \frac{\mathrm{d}P}{\mathrm{d}\mu}(x) = \int \mu(\mathrm{d}x)\, \varphi\Big(\frac{\mathrm{d}P}{\mathrm{d}\mu}(x)\Big),$$

with $\varphi(y) = y\log y$.

This can be seen as an abstract version of Cramér's theorem for the i.i.d. variables δ_{X_i}:

Entropy $=$ Legendre transform

For any $\nu, \mu \in \mathcal{M}_1(\Gamma)$,

$$H(\nu \mid \mu) = \sup_{f\in\mathcal{C}_\mathrm{b}(\Gamma)} \Big[\int_\Gamma f\,\mathrm{d}\nu - \log\int_\Gamma \mathrm{e}^f\,\mathrm{d}\mu\Big].$$

The minimizer is $f = \log\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$, if it is well-defined.

Here is an application of SANOV's theorem to statistical physics. Assume that $\mathcal{X}$ is finite. We condition $(X_1, \dots, X_n)$ on the event

$$\Big\{\sum_{i=1}^n f(X_i) \in A\Big\} = \{\langle f, S_n\rangle \in A\} = \{S_n \in \Sigma_{A,f}\},$$

for some $A \subset \mathbb{R}$ and some $f\colon \mathcal{X} \to \mathbb{R}$. Assume that

$$\Lambda(\Sigma_{A,f}) \equiv \inf_{\Sigma_{A,f}^\circ} H(\cdot \mid \mu) = \inf_{\overline{\Sigma_{A,f}}} H(\cdot \mid \mu),$$

and denote by $\mathcal{M}(\Sigma_{A,f})$ the set of minimizers. Then

The Gibbs principle

- All the accumulation points of the conditional distribution of S_n given $\{S_n \in \Sigma_{A,f}\}$ lie in $\overline{\mathrm{conv}(\mathcal{M}(\Sigma_{A,f}))}$.
- If $\Sigma_{A,f}$ is convex with non-empty interior, then $\mathcal{M}(\Sigma_{A,f})$ is a singleton, to which this distribution then converges.

Let $(X_i)_{i\in\mathbb{N}_0}$ be a Markov chain on the finite set Γ with transition kernel $P = (p(i,j))_{i,j\in\Gamma}$. Let $\mathcal{X}^{(2)}$ denote the set of probability measures on $\Gamma\times\Gamma$ with equal marginals.

LDP for the empirical pair measures

The empirical pair measure $L_n^{(2)} = \frac{1}{n}\sum_{i=1}^n \delta_{(X_i,X_{i+1})}$ satisfies an LDP on $\mathcal{X}^{(2)}$ with rate function

$$I^{(2)}(\nu) = \sum_{\gamma,\tilde\gamma\in\Gamma} \nu(\gamma,\tilde\gamma)\log\frac{\nu(\gamma,\tilde\gamma)}{\overline{\nu}(\gamma)p(\gamma,\tilde\gamma)}.$$

Let $(X_i)_{i\in\mathbb{N}_0}$ be a Markov chain on the finite set Γ with transition kernel $P = (p(i,j))_{i,j\in\Gamma}$. Let $\mathcal{X}^{(2)}$ denote the set of probability measures on $\Gamma\times\Gamma$ with equal marginals.

LDP for the empirical pair measures

The empirical pair measure $L_n^{(2)} = \frac{1}{n}\sum_{i=1}^n \delta_{(X_i,X_{i+1})}$ satisfies an LDP on $\mathcal{X}^{(2)}$ with rate function

$$I^{(2)}(\nu) = \sum_{\gamma,\tilde\gamma\in\Gamma} \nu(\gamma,\tilde\gamma)\log\frac{\nu(\gamma,\tilde\gamma)}{\overline{\nu}(\gamma)p(\gamma,\tilde\gamma)}.$$

- There is a combinatorial proof. There are versions for Polish spaces Γ, e.g. under the assumption of a strong uniform ergodicity.
- $I^{(2)}(\nu)$ is the entropy of ν with respect to $\overline{\nu}\otimes P$.
- There is an extension to k-tupels, $L_n^{(k)} = \frac{1}{n}\sum_{i=1}^n \delta_{(X_i,\dots,X_{i-1+k})} \in \mathcal{M}_1(\Gamma^k)$. The rate function $I^{(k)}(\nu)$ is the entropy of ν with respect to $\overline{\nu}\otimes P$, where $\overline{\nu}$ is the projection on the first $k-1$ coordinates.
- Using projective limits as $k\to\infty$, one finds, via the DAWSON-GÄRTNER approach, an extension for $k=\infty$, i.e., mixtures of Dirac measures on shifts, see below.

- $(X_n)_{n\in\mathbb{N}_0}$ = simple random walk on $\mathbb{Z}$,$\ell(x) = \sum_{i=1}^n \mathbb{1}_{\{X_i=x\}}$ local times,

$$Y_n = \sum_{i,j=1}^n \mathbb{1}_{\{X_i=X_j\}} = \sum_{x\in\mathbb{Z}} \ell_n(x)^2 \qquad \text{number of self-intersections}$$

- polymer measure $$\mathrm{d}\mathbb{P}_{n,\beta} = \frac{1}{Z_{n,\beta}} \mathrm{e}^{-\beta Y_n}\,\mathrm{d}\mathbb{P}, \qquad \beta\in(0,\infty),$$

- Discrete version of the Ray-Knight theorem $\Longrightarrow$ in some situations, $\ell(x) = m(x) + m(x-1) - 1$ with a Markov chain $(m(x))_{x\in\mathbb{N}_0}$ on $\mathbb{N}$ with transition kernel

$$p(i,j) = 2^{-(i+j-1)}\binom{i+j-1}{i-1}, \qquad i,j\in\mathbb{N}.$$

Hence,

$$\begin{aligned} Z_{n,\beta}(\theta) &:= \mathbb{E}\big[\mathrm{e}^{-\beta Y_n}\mathbb{1}_{\{X_n\approx\theta n\}}\big] \\ &\approx \mathbb{E}\Big[\mathrm{e}^{-\beta\sum_{x=1}^{\theta n}(m(x)+m(x-1)-1)^2}\mathbb{1}_{\{\sum_{x=1}^{\theta n}(m(x)+m(x-1)-1)=n\}}\Big] \\ &\approx \mathbb{E}\Big[\mathrm{e}^{-\beta\theta n\langle L^{(2)}_{\theta n},\varphi^2\rangle}\mathbb{1}_{\{\langle L^{(2)}_{\theta n},\varphi\rangle=1/\theta\}}\Big], \qquad \text{with } \varphi(i,j) = i+j-1. \end{aligned}$$

Hence, the LDP for $L_n^{(2)}$, together with Varadhan's lemma, gives

$$\lim_{n\to\infty}\frac{1}{n}\log Z_{n,\beta}(\theta) = -\chi_\beta(\theta),$$

where

$$\chi_\beta(\theta) = \theta \inf\Big\{\beta\langle\nu,\varphi^2\rangle + I^{(2)}(\nu)\colon \nu\in\mathcal{M}_1^{(s)}(\mathbb{N}^2), \langle\nu,\varphi\rangle = \frac{1}{\theta}\Big\}.$$

The minimizer exists, and is unique; it gives a lot of information about the 'typical' behaviour of the polymer measure. In particular, χ_β is strictly minimal at some *positive* θ_β^*, i.e., the polymer has a positive drift.

(Details: [GREVEN/DEN HOLLANDER (1993)])

An important tool:

Contraction principle

If $(S_n)_{n\in\mathbb{N}}$ satisfies an LDP with rate function I on $\mathcal{X}$, and if $F\colon \mathcal{X}\to\mathcal{Y}$ is a continuous map into another metric space, then also $(F(S_n))_{n\in\mathbb{N}}$ satisfies an LDP with rate function

$$J(y)=\inf\{I(x)\colon x\in\mathcal{X}, F(x)=y\},\qquad y\in\mathcal{Y}.$$

- Markov chains: The (explicit) LDP for empirical pair measures $L_n^{(2)}=\frac{1}{n}\sum_{i=1}^n\delta_{(X_i,X_{i+1})}$ of a Markov chain implies a (less explicit) LDP for the empirical measure $L_n^{(1)}=\frac{1}{n}\sum_{i=1}^n\delta_{X_i}$ of this chain, since the map $\nu\mapsto\overline{\nu}$ (mrginal measure) is continuous. There is in general no better formula than

$$I^{(1)}(\mu)=\inf\left\{I^{(2)}(\nu)\colon \nu\in\mathcal{M}_1^{(\mathrm{s})}(\Gamma\times\Gamma),\overline{\nu}=\mu\right\}.$$

This is an application of the contraction principle to SCHILDER's theorem. It is the starting point of the FREIDLIN-WENTZELL theory.

Let $B = (B_t)_{t\in[0,1]}$ be a d-dimensional Brownian motion, and consider the SDE

$$\mathrm{d}X_t^{(\varepsilon)} = b(X_t^{(\varepsilon)})\mathrm{d}t + \varepsilon \mathrm{d}B_t, \qquad t\in[0,1], \qquad X_0^{(\varepsilon)} = x_0,$$

with $b\colon \mathbb{R}^d \to \mathbb{R}^d$ Lipschitz continuous. That is,

$$X_t^{(\varepsilon)} = x_0 + \int_0^t b(X_s^{(\varepsilon)})\mathrm{d}s + \varepsilon B_t, \qquad t\in[0,1].$$

Hence, $X^{(\varepsilon)}$ is a continuous function of B. Hence, an application of the contraction principle to SCHILDER's theorem gives that $(X^{(\varepsilon)})_{\varepsilon>0}$ satisfies an LDP with scale ε^{-2} and rate function

$$\psi \mapsto \frac{1}{2}\int_0^1 |\psi'(t) - b(\psi(t))|^2 \,\mathrm{d}t, \qquad \text{if } \psi(0) = x_0 \text{ and } \psi \text{ is absolutely continuous.}$$

This is a far-reaching extension of the LDP for k-tuple measures for Markov chains:

- $k = \infty$.
- d-dimensional parameter space instead of $\mathbb{N}$.
- continuous parameter space $\mathbb{R}^d$ instead of $\mathbb{N}^d$
- reference measure is the Poisson point process (PPP) instead of a Markov chain.
- we add marks to the particles.

This is a far-reaching extension of the LDP for k-tuple measures for Markov chains:

- $k = \infty$.
- d-dimensional parameter space instead of $\mathbb{N}$.
- continuous parameter space $\mathbb{R}^d$ instead of $\mathbb{N}^d$
- reference measure is the Poisson point process (PPP) instead of a Markov chain.
- we add marks to the particles.

Let $\omega_{\rm P} = \sum_{i\in I} \delta_{(x_i,m_i)}$ be a marked PPP in $\mathbb{R}^d \times \mathfrak{M}$ with intensity measure $\lambda{\rm Leb} \otimes m$. For a centred box Λ, let $\omega^{(\Lambda)}$ be the Λ-periodic repetition of the restriction of ω to Λ.

empirical stationary field: $$\mathcal{R}_\Lambda(\omega) = \frac{1}{|\Lambda|}\int_\Lambda {\rm d}x\, \delta_{\theta_x(\omega^{(\Lambda)})}$$

This is a stationary marked point processes in $\mathbb{R}^d$.

LDP for the field [GEORGII/ZESSIN (1994)]

As $\Lambda \uparrow \mathbb{R}^d$, the distributions of $\mathcal{R}_\Lambda(\omega_{\rm P})$ satisfy an LDP with rate function

$$I(P) = H(P \mid \omega_{\rm P}) = \lim_{\Lambda\uparrow\mathbb{R}^d} \frac{1}{|\Lambda|} H_\Lambda(P|_\Lambda \mid \omega_{\rm P}|_\Lambda),$$

which is lower semi-continuous and affine.

N independent particles $X_1, \dots, X_N$ in a centred box $\Lambda_N \subset \mathbb{R}^d$ of volume N/ρ with pair interaction

$$V(x_1, \dots, x_N) = \sum_{1 \le i < j \le N} v(|x_i - x_j|), \qquad \text{with } v\colon (0,\infty) \to \mathbb{R} \text{ and } \lim_{r \downarrow 0} v(r) = \infty.$$

Partition function: $$Z_{N,\beta,\Lambda_N} = \frac{1}{N!} \int_{\Lambda_N^N} \mathrm{d}x_1 \dots \mathrm{d}x_N \, \mathrm{e}^{-\beta V(x)}.$$

(Mark-dependent models also within reach in general)

We seek for a formula for the free energy per volume

$$f(\beta, \rho) = -\frac{1}{\beta} \lim_{N \to \infty} \frac{1}{|\Lambda_N|} \log Z_{N,\beta,\Lambda_N}.$$

Strategy:

1. Rewrite Z_{N,β,Λ_N} in terms of a PPP(ρ), $\omega_{\rm P}$.
2. Use that it has N i.i.d. uniform particles, when conditioned on having N particles in Λ_N.
3. Rewrite the energy as $|\Lambda_N|\langle \mathcal{R}_{\Lambda_N}(\omega_{\rm P}), \beta F\rangle$ with suitable F and the conditioning event as $\{\langle \mathcal{R}_{\Lambda_N}(\omega_{\rm P}), \mathcal{N}_U\rangle = \rho\}$.
4. Use the LDP and obtain a variational formula
5. (Try to squeeze some information out ...)

The functionals are (for $\omega = \sum_{i\in I}\delta_{x_i}$, using the unit box $U = [-\frac12, \frac12]^d$),

$$F(\omega) = \frac12 \sum_{i\neq j:\, x_i\in U} v(|x_i - x_j|) \qquad \text{and} \qquad N_U(\omega) = \sum_{i\in I} \mathbb{1}_U(x_i).$$

Hence, we should obtain

$$f(\beta,\rho) = \inf\Big\{\langle P, F\rangle + \frac{1}{\beta} I(P) \colon P \in \mathcal{M}_1^{(s)}(\Omega), \langle P, N_U\rangle = \rho\Big\}.$$