Capacitary inequalities in discrete setting and application to metastable Markov chains

André Schlichting

Institute for Applied Mathematics, University of Bonn

joint work with M. Slowik (TU Berlin)

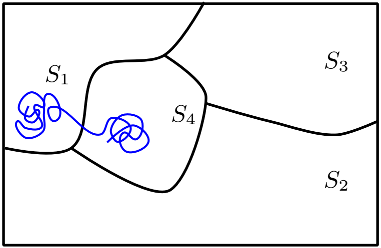
12th German Probability and Statistics Days, Bochum

universitätbonn iam

Metastability: A common phenomenon

The paradigm. Related to the dynamics of first order phase transitions

Change parameters quickly across the line of first order phase transition, the system reveals the existence of multiple time scales:



Short time scales.

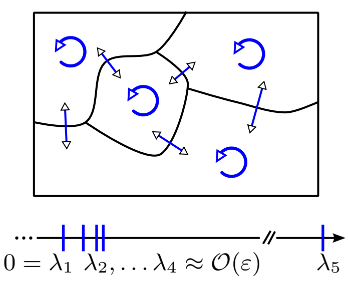
- Existence of disjoint subsets S_i trapping effectively the system
- Quasi-equilibrium ($\hat{=}$ metastable states) is reached within S_i

Larger time scales.

- Rapid transitions between S_i and S_j occur induced by random fluctuations

Spectrum and metastability

Heuristic. Reversible Markov process $\{X_t : t \geq 0\}$, generator L, $\lambda_i \in \mathrm{spec}(-L)$



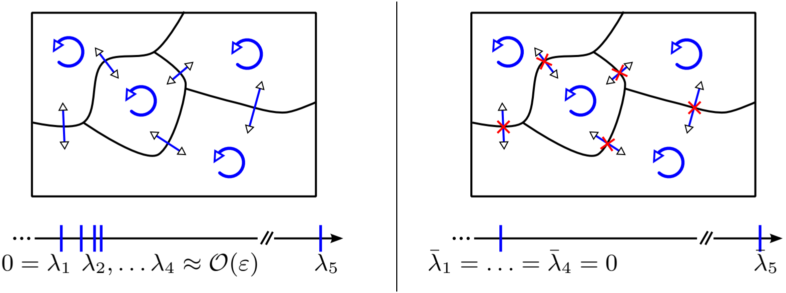
The goal. Understanding of quantitative aspects of dynamical phase transitions:

- expected time of a transition from a metastable to a stable state
- distribution of the exit time from a metastable state
- spectral properties of the generator and mixing times

universitätbonn iam

Spectrum and metastability

Heuristic. Reversible Markov process $\{X_t : t \geq 0\}$, generator L, $\lambda_i \in \mathrm{spec}(-L)$

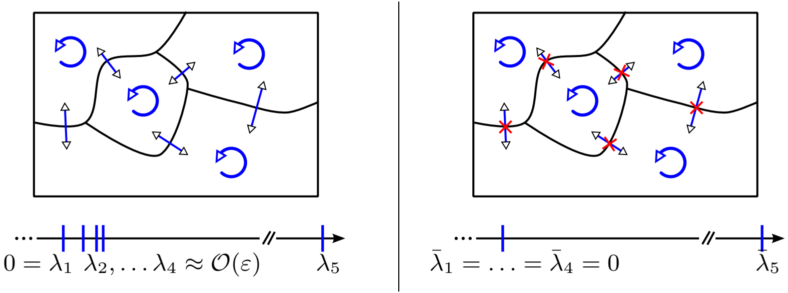


The goal. Understanding of quantitative aspects of dynamical phase transitions:

- expected time of a transition from a metastable to a stable state
- distribution of the exit time from a metastable state
- spectral properties of the generator and mixing times

Spectrum and metastability

Heuristic. Reversible Markov process $\{X_t : t \geq 0\}$, generator L, $\lambda_i \in \mathrm{spec}(-L)$



The goal. Understanding of quantitative aspects of dynamical phase transitions:

- expected time of a transition from a metastable to a stable state
- distribution of the exit time from a metastable state
- spectral properties of the generator and mixing times

Reversible Markov chains

Setting.

- state space $\mathcal{S}$ (finite or countable infinite)
- μ measure on $\mathcal{S}$
- $\big(p(x,y) : x, y \in \mathcal{S}\big)$ stochastic matrix, irreducible (positive recurrent)

Dynamics. Discrete-time Markov chain $X = \{X_t : t \geq 0\}$ on $\mathcal{S}$ with generator

$$(Lf)(x) \;=\; \sum_y p(x,y)\,\big(f(y) - f(x)\big)$$

Assumption: The Markov process X is reversible with respect to μ.

First return time. For any $A \subset \mathcal{S}$, let

$$\tau_A \;=\; \inf\{t > 0 \,:\, X_t \in A\}$$

Setting.

- state space $\mathcal{S}$ (finite or countable infinite)
- μ measure on $\mathcal{S}$
- $\big(p(x,y) : x,y \in \mathcal{S}\big)$ stochastic matrix, irreducible (positive recurrent)

Dynamics. Discrete-time Markov chain $X = \{X_t : t \geq 0\}$ on $\mathcal{S}$ with generator

$$(Lf)(x) \;=\; \sum_y p(x,y)\,\big(f(y) - f(x)\big)$$

Assumption: The Markov process X is reversible with respect to μ.

First return time. For any $A \subset \mathcal{S}$, let

$$\tau_A \;=\; \inf\{t > 0 \;:\; X_t \in A\}$$

Setting.

- state space $\mathcal{S}$ (finite or countable infinite)
- μ measure on $\mathcal{S}$
- $\big(p(x,y) : x, y \in \mathcal{S}\big)$ stochastic matrix, irreducible (positive recurrent)

Dynamics. Discrete-time Markov chain $X = \{X_t : t \geq 0\}$ on $\mathcal{S}$ with generator

$$(Lf)(x) \;=\; \sum_y p(x,y)\,\big(f(y) - f(x)\big)$$

Assumption: The Markov process X is reversible with respect to μ.

First return time. For any $A \subset \mathcal{S}$, let

$$\tau_A \;=\; \inf\{t > 0 \,:\, X_t \in A\}$$

universitätbonn iam

Poincaré and logarithmic Sobolev inequality

$$\mathcal{E}(f,f) \;=\; \frac{1}{2} \sum_{x,y\in\mathcal{S}} \mu(x)\, p(x,y)\, \big(f(x)-f(y)\big)^2$$

Poincaré inequality.

$$\forall\, f:\mathcal{S}\to\mathbb{R}: \qquad \mathrm{var}_\mu[f] \;\le\; C_{\mathrm{PI}}\, \mathcal{E}(f,f). \qquad \mathrm{PI}(C_{\mathrm{PI}})$$

Logarithmic Sobolev inequality.

$$\forall\, f:\mathcal{S}\to\mathbb{R}: \qquad \mathrm{Ent}_\mu[f^2] \;=\; \mathrm{E}_\mu\Big[f^2\,\ln\frac{f^2}{\mathrm{E}_\mu[f^2]}\Big] \;\le\; C_{\mathrm{LSI}}\,\mathcal{E}(f,f). \quad \mathrm{LSI}(C_{\mathrm{LSI}})$$

The goal: Compute for metastable Markov chains in terms of capacities

- the optimal constant C_{PI} in the Poincaré inequality (inverse spectral gap)
- the optimal constant C_{LSI} in the logarithmic Sobolev inequality

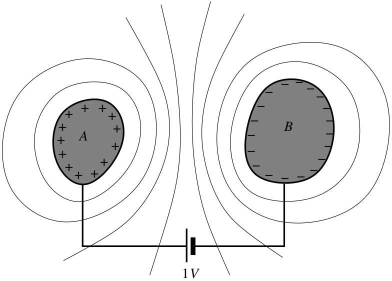
Equilibrium potential. Given $A, B \subset \mathcal{S}$ disjoint

$$\begin{cases} Lh_{A,B} = 0, & \text{on } (A \cup B)^c \\ h_{A,B} = \mathbb{1}_A, & \text{on } A \cup B \end{cases}$$

$$\forall x \in (A \cup B)^c : \quad h_{A,B}(x) = \mathbb{P}_x\big[\tau_A < \tau_B\big]$$

Capacity.

$$\begin{aligned} \mathrm{cap}(A,B) &= \sum_{x \in A} \mu(x)\,(-Lh_{A,B})(x) \\ &= \langle h_{A,B}, -Lh_{A,B} \rangle_\mu \\ &= \mathcal{E}(h_{A,B}, h_{A,B}) \\ &= \sum_{x \in A} \mu(x)\,\mathbb{P}_x\big[\tau_B < \tau_A\big] \end{aligned}$$



Fact.

$$\mathrm{cap}(A,B) = \mathrm{cap}(B,A) \quad \text{and} \quad \mathrm{cap}(A',B) \le \mathrm{cap}(A,B), \quad \forall\, A' \subset A$$

Computation of capacities

Variational principles. Allows to bound capacities from above and from below

Dirichlet principle.

$$\mathrm{cap}(A,B) = \inf_{h \in \mathcal{H}_{A,B}} \frac{1}{2} \sum_{x,y} \mu(x)\, p(x,y)\, \big(h(x) - h(y)\big)^2$$

$\mathcal{H}_{A,B}$: space of functions with boundary constraints; minimizer harmonic function

Thomson principle.

$$\frac{1}{\mathrm{cap}(A,B)} = \inf_{f \in \mathcal{U}_{A,B}} \frac{1}{2} \sum_{x,y} \frac{f(x,y)^2}{\mu(x)\, p(x,y)}$$

$\mathcal{U}_{A,B}$: space of unit AB-flows; maximizer harmonic flow.

Capacitary inequalities

$$\langle h, -Lg \rangle_\mu \;=\; \frac{1}{2} \sum_{x,y \in \mathcal{S}} \mu(x)\, p(x,y)\, \big(h(x) - h(y)\big)\big(g(x) - g(y)\big)$$

Proposition

Let $B \subset \mathcal{S}$ be non-empty. For any $f : \mathcal{S} \to \mathbb{R}$ with $f \equiv 0$ on B set

$$A_t \;:=\; \{x \in \mathcal{S} \,:\, |f(x)| > t\}.$$

Then,

$$\int_0^\infty 2t\, \mathrm{cap}(A_t, B)\, \mathrm{d}t \;\le\; 4\, \mathcal{E}(f,f).$$

Previous and related work

- Maz'ya (1972), operators in divergence form on $\mathbb{R}^d$

Capacitary inequalities

$$\langle h, -Lg \rangle_\mu \;=\; \frac{1}{2} \sum_{x,y \in \mathcal{S}} \mu(x)\, p(x,y)\, \big(h(x) - h(y)\big)\big(g(x) - g(y)\big)$$

Proposition

Let $B \subset \mathcal{S}$ be non-empty. For any $f : \mathcal{S} \to \mathbb{R}$ with $f \equiv 0$ on B set

$$A_t \;:=\; \{x \in \mathcal{S} \,:\, |f(x)| > t\}.$$

Then,

$$\int_0^\infty 2t \;\mathrm{cap}(A_t, B)\; \mathrm{d}t \;\leq\; 4\,\mathcal{E}(f,f).$$

Previous and related work

- Maz'ya (1972), operators in divergence form on $\mathbb{R}^d$

universitätbonn iam

Proposition

Let $B \subset \mathcal{S}$ *be non-empty and* $\nu \in \mathcal{P}_1(\mathcal{S})$*. Then, there exist* $C_1, C_2 \in (0, \infty)$ *satisfying* $C_1 \;\leq\; C_2 \;\leq\; 4C_1$ *such that the following statements are equivalent:*

(i) *For all* $A \subset \mathcal{S} \setminus B$ *it holds*

$$\nu[A] \;\leq\; C_1 \operatorname{cap}(A, B).$$

(ii) *For all* $f : \mathcal{S} \to \mathbb{R}$ *with* $f|_B \equiv 0$ *holds*

$$\|f^2\|_{\ell^1(\nu)} \;\leq\; C_2\, \mathcal{E}(f, f).$$

Consequences

$$\|f\|_{\Phi,\nu} \;:=\; \sup\left\{ \mathrm{E}_\nu[|f|g] \;:\; g \geq 0,\; \mathrm{E}_\nu[\Psi(g)] \leq 1 \right\}$$

Proposition

Let $B \subset \mathcal{S}$ be non-empty and $\nu \in \mathcal{P}_1(\mathcal{S})$. Then, for any Orlicz pair (Φ, Ψ), there exist $C_1, C_2 \in (0, \infty)$ satisfying $C_1 \leq C_2 \leq 4C_1$ such that the following statements are equivalent:

(i) *For all $A \subset \mathcal{S} \setminus B$ it holds*

$$\nu[A]\,\Psi^{-1}\big(1/\nu[A]\big) \;\leq\; C_1 \operatorname{cap}(A, B).$$

(ii) *For all $f : \mathcal{S} \to \mathbb{R}$ with $f|_B \equiv 0$ holds*

$$\|f^2\|_{\Phi,\nu} \;\leq\; C_2\, \mathcal{E}(f, f).$$

Examples:

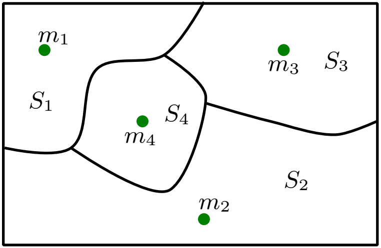
$$\big(\Phi_p(r), \Psi_p(r)\big) \;:=\; \big(\tfrac{1}{p} r^p, \tfrac{1}{p_*} r^{p_*}\big),$$

$$\big(\Phi_{\mathrm{Ent}}(r), \Psi_{\mathrm{Ent}}(r)\big) \;:=\; \big(\mathbf{1}_{[1,\infty)}(r)(r \ln r - r + 1), \mathrm{e}^r - 1\big)$$

Definition

Let $\rho > 0$ and $\mathcal{M} \subset \mathcal{S}$ be finite. $\{X_t : t \geq 0\}$ is ρ-metastable with respect to $\mathcal{M}$ (set of metastable points), if

$$\frac{\max_{m \in \mathcal{M}} \mathbb{P}_m\left[\tau_{\mathcal{M}\setminus m} < \tau_m\right]}{\min_{A \subset \mathcal{S}\setminus\mathcal{M}} \mathbb{P}_{\mu_A}\left[\tau_{\mathcal{M}} < \tau_A\right]} \leq \rho \ll 1.$$



Previous and related definition

- Bovier (2006), reversible Markov chains with finite state space; reversible diffusions
- Bovier, den Hollander (2016): Metastability - a portential theoretic approach
  $\longrightarrow$ see Springer bookshelf

Metastable partition. $\mathcal{S} = \bigcup_{m \in \mathcal{M}} S_m$, the sets S_m, $m \in \mathcal{M}$ are mutually disjoint

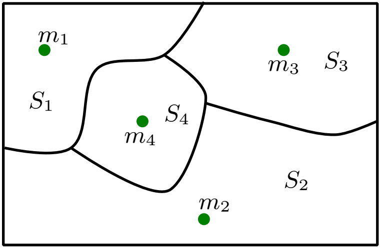
$$S_m \subseteq \left\{x \in \mathcal{S} : \mathbb{P}_x\left[\tau_m < \tau_{\mathcal{M}\setminus m}\right] \geq \max_{m' \in \mathcal{M}\setminus m} \mathbb{P}_x\left[\tau_{m'} < \tau_{\mathcal{M}\setminus m'}\right]\right\}$$

universitätbonn iam

Definition

Let $\rho > 0$ and $\mathcal{M} \subset \mathcal{S}$ be finite. $\{X_t : t \geq 0\}$ is ρ-metastable with respect to $\mathcal{M}$ (set of metastable points), if

$$\frac{\max_{m\in\mathcal{M}} \mathbb{P}_m\left[\tau_{\mathcal{M}\setminus m} < \tau_m\right]}{\min_{A\subset\mathcal{S}\setminus\mathcal{M}} \mathbb{P}_{\mu_A}\left[\tau_{\mathcal{M}} < \tau_A\right]} \leq \rho \ll 1.$$



Previous and related definition

- Bovier (2006), reversible Markov chains with finite state space; reversible diffusions
- Bovier, den Hollander (2016): Metastability - a portential theoretic approach
  $\longrightarrow$ see Springer bookshelf

Metastable partition. $\mathcal{S} = \bigcup_{m\in\mathcal{M}} S_m$, the sets S_m, $m \in \mathcal{M}$ are mutually disjoint

$$S_m \subseteq \left\{x \in \mathcal{S} : \mathbb{P}_x\left[\tau_m < \tau_{\mathcal{M}\setminus m}\right] \geq \max_{m'\in\mathcal{M}\setminus m} \mathbb{P}_x\left[\tau_{m'} < \tau_{\mathcal{M}\setminus m'}\right]\right\}$$

universitätbonn iam

Main result

Theorem

Suppose $\{X_t : t \geq 0\}$ *is a* ρ*-metastable Markov chain with* $\mathcal{M} = \{m_1, m_2\}$*. Then,*

$$C_{\text{PI}} = \frac{\mu[S_1]\,\mu[S_2]}{\text{cap}(m_1, m_2)} \big(1 + O(\sqrt{\rho})\big).$$

Moreover, under further conditions on $\mu[\cdot|S_i]$*, it holds*

$$C_{\text{LSI}} = \frac{\mu[S_1]\,\mu[S_2]}{\Lambda(\mu[S_1], \mu[S_2])} \frac{1}{\text{cap}(m_1, m_2)} \big(1 + O(\sqrt{\rho})\big),$$

where $\Lambda(s,t) = (s-t)/(\ln s - \ln t)$ *denotes the logarithmic mean.*

Previous and related results

- Bovier, Eckhoff, Gayrard, Klein (2002), low lying spectrum, reversible Markov chains
- Bovier, Gayrard, Klein (2005), low lying spectrum, reversible diffusion
- Bianchi, Gaudilliére (2011), spectral gap, reversible Markov chains
- Menz, S. (2014), PI and LSI, reversible diffusions

Main result

Theorem

Suppose $\{X_t : t \geq 0\}$ *is a* ρ*-metastable Markov chain with* $\mathcal{M} = \{m_1, m_2\}$*. Then,*

$$C_{\mathrm{PI}} = \frac{\mu[S_1]\,\mu[S_2]}{\mathrm{cap}(m_1, m_2)} \big(1 + O(\sqrt{\rho})\big).$$

Moreover, under further conditions on $\mu[\cdot|S_i]$*, it holds*

$$C_{\mathrm{LSI}} = \frac{\mu[S_1]\,\mu[S_2]}{\Lambda(\mu[S_1], \mu[S_2])} \frac{1}{\mathrm{cap}(m_1, m_2)} \big(1 + O(\sqrt{\rho})\big),$$

where $\Lambda(s,t) = (s-t)/(\ln s - \ln t)$ *denotes the logarithmic mean.*

Previous and related results

- Bovier, Eckhoff, Gayrard, Klein (2002), low lying spectrum, reversible Markov chains
- Bovier, Gayrard, Klein (2005), low lying spectrum, reversible diffusion
- Bianchi, Gaudilliére (2011), spectral gap, reversible Markov chains
- Menz, S. (2014), PI and LSI, reversible diffusions

Two-scale decomposition

$$\mu_i[\,\cdot\,] \;:=\; \mu[\,\cdot\,|S_i] \qquad \text{and} \qquad \bar{\mu} \;:=\; \mu[S_1]\,\delta_{m_1} \;+\; \mu[S_2]\,\delta_{m_2}$$

Splitting the variance.

$$\operatorname{var}_\mu[f] \;=\; \mu[S_1]\,\underbrace{\operatorname{var}_{\mu_1}[f]}_{\text{local variance}} + \mu[S_2]\,\underbrace{\operatorname{var}_{\mu_2}[f]}_{\text{local variance}} + \mu[S_1]\,\mu[S_2]\,\Big(\underbrace{\mathrm{E}_{\mu_1}[f] - \mathrm{E}_{\mu_2}[f]}_{\text{mean difference}}\Big)^2$$

Splitting the entropy.

$$\operatorname{Ent}_\mu[f^2] \;=\; \mu[S_1]\,\underbrace{\operatorname{Ent}_{\mu_1}[f^2]}_{\text{local entropy}} + \mu[S_2]\,\underbrace{\operatorname{Ent}_{\mu_2}[f^2]}_{\text{local entropy}} + \underbrace{\operatorname{Ent}_{\bar\mu}\big[\,\mathrm{E}_{\mu_\cdot}[f^2]\big]}_{\text{macroscopic entropy}}$$

$$\operatorname{Ent}_{\bar\mu}\big[\,\mathrm{E}_{\mu_\cdot}[f^2]\big] \;\le\; \frac{\mu[S_1]\,\mu[S_2]}{\Lambda(\mu[S_1],\mu[S_2])}\bigg(\operatorname{var}_{\mu_1}[f] + \operatorname{var}_{\mu_2}[f] + \big(\mathrm{E}_{\mu_1}[f] - \mathrm{E}_{\mu_2}[f]\big)^2\bigg)$$

The strategy.

- rough bounds for local quantities,
- sharp bounds for the mean difference

Local variances and mean difference estimate

Fact.

$$\mathbb{P}_{\mu_A}\left[\tau_{m_i} < \tau_A\right] \;\geq\; \frac{1}{|\mathcal{M}|}\,\mathbb{P}_{\mu_A}\left[\tau_{\mathcal{M}} < \tau_A\right] \qquad \forall\, A \subset S_i \setminus \{m_i\}$$

Assumption of metastability. For all $A \subset S_i \setminus \{m_i\}$

$$\mu_i[A] \;\leq\; \frac{\rho|\mathcal{M}|}{\mu[S_i]}\left(\max_{m \in \mathcal{M}\setminus\{m_i\}} \mathbb{P}_m\left[\tau_{\mathcal{M}\setminus\{m\}} < \tau_m\right]\right) \operatorname{cap}(A, m_i)$$

Local variances. $\mathcal{M} = \{m_1, m_2\}$

$$\mu[S_i]\;\operatorname{var}_{\mu_i}[f] \;\leq\; 4\,\rho\,|\mathcal{M}|\,\frac{\mu[S_1]\,\mu[S_2]}{\operatorname{cap}(m_1, m_2)}\,\mathcal{E}(f, f)$$

Mean difference estimate.

$$\mu[S_1]\,\mu[S_2]\,\big(\mathrm{E}_{\mu_1}[f] - \mathrm{E}_{\mu_2}[f]\big)^2 \;\leq\; \frac{\mu[S_1]\,\mu[S_2]}{\operatorname{cap}(m_1, m_2)}\,\mathcal{E}(f, f)\,\Big(1 + O\big(\sqrt{\rho|\mathcal{M}|}\big)\Big)$$

Summary and open problems

What has been done so far.

- Capacitary inequality that allows to establish a local PI and LSI inequality
- Method can be applied beyond the situation of metastable points (e.g. RFCW)
- Generalization to metastable sets under suitable regularity assumptions on the metastable sets

Possible extensions and future challenges

- Continuous high dimensional systems (approximation of stochastic $1d$-Allen-Cahn equation)
- certain types of entropic metastability
- non-reversible systems