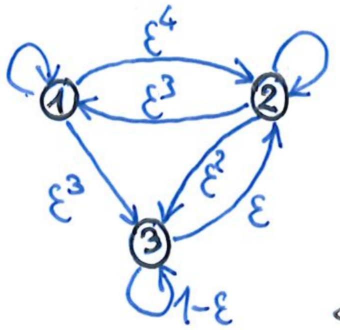
Metastable Markov chains, trace process, and spectral theory

Joint work with Manon Baudel (CERMICS), SIAM J. Math. Anal. 49 (2017)

Example 1



$$P = \begin{pmatrix} 1-\varepsilon^3-\varepsilon^4 & \varepsilon^4 & \varepsilon^3 \\ \varepsilon^3 & 1-\varepsilon^2-\varepsilon^3 & \varepsilon^2 \\ 0 & \varepsilon & 1-\varepsilon \end{pmatrix}$$

$0 \leqslant \varepsilon \leqslant \varepsilon_{max}$

Stat. distrib: $\pi_0 = \dfrac{1}{2(1+\varepsilon+\varepsilon^2)}(1, 1+\varepsilon, \varepsilon+2\varepsilon^2)$

- $\varepsilon = 0$: $P = I$
- $0 < \varepsilon \leqslant \varepsilon_{max}$: irreducible, aperiodic, not reversible

Speed of convergence to π_0?

$$\begin{cases} 1-\lambda_0 = 0 \\ 1-\lambda_1 = 2\varepsilon^3 + O(\varepsilon^5) \\ 1-\lambda_2 = \varepsilon + O(\varepsilon^2) \end{cases} \longrightarrow T \sim (2\varepsilon^3)^{-1}$$

Q: how to easily determine leading term of $1-\lambda_1$?

Approaches:
- analytic (singular perturbation theory)
- probabilistic:
  - clustering of states [Freidlin & Wentzell]
  - speed up time [Betz & Le Roux]

Trace process

$\mathcal{X}$ finite set

$(X_n)_{n\geqslant 0}$ Markov chain on $\mathcal{X}$ (recurrent, aperiodic), matrix P

$A \subset \mathcal{X}$

- process killed upon leaving A: $P_A(x,y) = P(x,y)\,\mathbb{1}_{\{x,y\in A\}}$
- trace process on A: ${}_AP(x,y) = \mathbb{P}^x\{X_{\tau_A^+} = y\}$ $\quad \tau_A^+ = \inf\{n\geqslant 1 : X_n \in A\}$

$$\begin{aligned} {}_AP(x,y) &= \mathbb{P}^x\{\tau_A^+ = 1, X_{\tau_A^+} = y\} + \mathbb{P}^x\{\tau_A^+ \geqslant 2, X_{\tau_A^+} = y\} \\ &= P(x,y) + \sum_{z\in A^c} P(x,z) \sum_{n\geqslant 1} \mathbb{P}^z\{\tau_A^+ = n, X_n = y\} \\ &= P_A(x,y) + \sum_{z,z'\in A^c} P(x,z) \underbrace{\sum_{n\geqslant 1} P_{A^c}^{n-1}(z,z')}_{=[\mathbb{1}-P_{A^c}]^{-1}(z,z')} P(z',y) \end{aligned}$$

Matrix representation:

$$P = \begin{pmatrix} P_A & P_{AA^c} \\ P_{A^cA} & P_{A^c} \end{pmatrix} \Rightarrow \boxed{{}_AP = P_A + P_{AA^c}[\mathbb{1} - P_{A^c}]^{-1} P_{A^cA}}$$

$\mathbb{1} - {}_AP$ = Schur complement of $\mathbb{1} - P_{A^c}$

Application: $\pi_0|_A$ is invariant by ${}_AP$

$$\boxed{\text{Prop: } \pi_0(x)\, \mathbb{P}^x\{\tau_y^+ < \tau_x^+\} = \pi_0(y)\, \mathbb{P}^y\{\tau_x^+ < \tau_y^+\}}$$

proof: $A = \{x, y\}$

$$\pi_0(x) = (\pi_0\, {}_AP)(x) = \pi_0(x)\, \mathbb{P}^x\{X_{\tau_A^+} = x\} + \pi_0(y)\, \mathbb{P}^y\{X_{\tau_A^+} = x\}$$

$$= \pi_0(x)\left[1 - \mathbb{P}^x\{\tau_y^+ < \tau_x^+\}\right] + \pi_0(y)\, \mathbb{P}^y\{\tau_x^+ < \tau_y^+\} \qquad \square$$

Example 1

$A = \{1, 2\}$

$${}_AP = \begin{pmatrix} 1-\varepsilon^3-\varepsilon^4 & \varepsilon^4 \\ \varepsilon^3 & 1-\varepsilon^2-\varepsilon^3 \end{pmatrix} + \begin{pmatrix} \varepsilon^3 \\ \varepsilon^2 \end{pmatrix} \frac{1}{\varepsilon} (0, \varepsilon)$$

$$= \begin{pmatrix} 1-\varepsilon^3-\varepsilon^4 & \varepsilon^3+\varepsilon^4 \\ \varepsilon^3 & 1-\varepsilon^3 \end{pmatrix}$$

$$\begin{tikzcd} 1 \arrow[r, bend left, "\varepsilon^3+\varepsilon^4"] & 2 \arrow[l, bend left, "\varepsilon^3"] \end{tikzcd}$$

$${}_A\pi_0 = \frac{1}{2+\varepsilon}(1, 1+\varepsilon) \qquad \begin{cases} 1 - {}_A\lambda_0 = 0 \\ 1 - {}_A\lambda_1 = 2\varepsilon^3 + \varepsilon^4 \end{cases}$$

Good domains A

$$\begin{cases} p_{in}(A) := \inf_{x \in A^c} \mathbb{P}^x\{X_1 \in A\} \\ p_{out}(A) := \sup_{x \in A} \mathbb{P}^x\{X_1 \in A^c\} \end{cases}$$

$$\boxed{A \text{ is a good domain} \iff p_{out}(A) \ll p_{in}(A)}$$

Example 1

$A = \{1, 2\}$ $\begin{cases} p_{in}(A) = \varepsilon \\ p_{out}(A) = \varepsilon^2 \end{cases} \Rightarrow A$ is a good domain

Main idea: For a good domain A

$$P = \begin{pmatrix} P_A & P_{AA^c} \\ P_{A^cA} & P_{A^c} \end{pmatrix} \quad \text{well-approximated by} \quad \hat{P} = \begin{pmatrix} {}_AP & 0 \\ P_{A^cA} & P_{A^c} \end{pmatrix}$$

$$\|Q\| := \sup_{\|\varphi\|_\infty = 1} \|Q\varphi\|_\infty = \sup_{\|\mu\|_1 = 1} \|\mu Q\|_1 = \sup_{x\in\mathcal{X}} \sum_{y\in\mathcal{X}} |q(x,y)|$$

$$\boxed{\text{Lemma: } \|P - \hat{P}\| \leqslant 2 p_{out}(A)}$$

proof: $$\|P-\hat{P}\| = \sup_{x\in A} \Big[\underbrace{\underbrace{\sum_{y\in A} |P_A(x,y) - {}_AP(x,y)|}_{= {}_AP(x,y) - P_A(x,y)}}_{= {}_AP(x,A) - P_A(x,A) = 1 - P_A(x,A) = P(x,A^c) \leqslant p_{out}(A)} + \underbrace{\sum_{y\in A^c} P(x,y)}_{\leqslant p_{out}(A)} \Big]$$

Fact from spectral theory: (using complex analysis, Riesz projector)

$\hat{\lambda}$ simple eigenvalue of $\hat{P}$ at distance $> \|P-\hat{P}\|$ from other eigenvalues
$\Rightarrow \exists!$ simple ev λ of P at distance $\mathcal{O}(\|P-\hat{P}\|)$ from $\hat{\lambda}$

Consequence: Assume $A^c = \{x\}$

$$\Rightarrow \quad p_{in}(A) = 1 - P(x,x) = 1 - \hat{\lambda}$$

$$\begin{aligned} \Rightarrow \quad 1 - \lambda &= 1 - \hat{\lambda} + \mathcal{O}(\|P-\hat{P}\|) \\ &= 1 - \hat{\lambda} + \mathcal{O}(p_{out}(A)) \\ &= (1-\hat{\lambda}) \Big[1 + \mathcal{O}\Big(\frac{p_{out}(A)}{p_{in}(A)} \Big) \Big] \end{aligned}$$

Example 1: $\hat{\lambda} = 1 - \varepsilon \Rightarrow \lambda_2 = 1 - \varepsilon[1 + \mathcal{O}(\varepsilon)]$

However, $\hat{\lambda} = 1 - 2\varepsilon^3 - \varepsilon^4 \Rightarrow \lambda_1 = 1 - 2\varepsilon^3 - \varepsilon^4 + \mathcal{O}(\varepsilon^2)$

not sufficiently precise...

Solution: Laplace transforms:

$u \in \mathbb{C}$ $\quad \mathbb{E}^x[e^{u\tau_A^+}]$ exists for $|e^{-u}| > 1 - p_{in}(A)$

(follows from $\mathbb{P}^y\{\tau_A^+ > n\} \leqslant (1-p_{in}(A))^n$)

Prop: If $|e^{-u}| > 1 - p_{in}(A)$ then "Feynman-Kac"

$$\begin{cases} (P\phi)(x) = e^{-u}\phi(x) & x \in A^c \\ \phi(x) = \bar{\phi}(x) & x \in A \end{cases}$$

admits as unique solution $\phi(x) = \mathbb{E}^x[e^{u\tau_A}\bar{\phi}(X_{\tau_A})]$
where $\tau_A = \inf\{n \geqslant 0 : X_n \in A\}$

Corollary: If $|e^{-u}| > 1 - p_{in}(A)$ then

$$P\phi = e^{-u}\phi \iff {}_AP^u\phi = e^{-u}\phi \text{ in } A$$

where $\quad {}_AP^u(x,y) = \mathbb{E}^x[e^{u(\tau_A^+-1)} 1\!\!1_{\{X_{\tau_A^+}=y\}}]$

proof of $\Leftarrow$: $\quad x \in A: \ ({}_AP^u\phi)(x) = e^{-u}\phi(x)$

Extend by setting $\quad \phi(x) = \mathbb{E}^x[e^{u\tau_A}\phi(X_{\tau_A})] \quad \forall x \in A^c$

$$\overset{\text{Prop.}}{\Rightarrow} \quad P\phi = e^{-u}\phi \quad \forall x \in A^c$$

$$\begin{aligned} x \in A: \quad P\phi(x) &= \mathbb{E}^x[\phi(X_1)] \\ &= \mathbb{E}^x[1\!\!1_{\{\tau_A^+=1\}}\phi(X_{\tau_A^+})] + \mathbb{E}^x[1\!\!1_{\{\tau_A^+>1\}}\underbrace{\phi(X_1)}_{=\mathbb{E}^{X_1}[e^{u\tau_A}\phi(X_{\tau_A})]}] \\ &= \mathbb{E}^x[e^{u(\tau_A^+-1)}\phi(X_{\tau_A^+})] \quad \square \end{aligned}$$

$\Rightarrow$ remaining ev satisfy $\begin{cases} {}_AP^u\phi = \lambda\phi \\ \lambda = e^{-u} \end{cases}$

In addition,

Prop: $$\|{}_AP^u - {}_AP^0\| \leqslant \frac{|1-e^{-u}| \sup_{x \in A} \mathbb{E}^x[\tau_A^+ - 1]}{1 - |1-e^{-u}| \sup_{x \in A^c} \mathbb{E}^x[\tau_A^+]}$$

Easy to check: $\begin{cases} \sup_{x\in A^c} \mathbb{E}^x[\tau_A^+] \leq \dfrac{1}{p_{in}(A)} \\ \sup_{x\in A} \mathbb{E}^x[\tau_A^+ - 1] \leq \dfrac{p_{out}(A)}{p_{in}(A)} \end{cases}$

$$\Rightarrow \quad \| {}_A P^u - {}_A P^0 \| \leq |1-e^{-u}| \frac{p_{out}(A)}{p_{in}(A) - |1-e^{-u}|}$$

$$\Rightarrow \quad \hat{\lambda} \text{ ev of } {}_A P^0 \text{ perturbs to } \quad \lambda = \hat{\lambda}\left[1 + O\left(\frac{p_{out}(A) + |1-\hat{\lambda}|}{p_{in}(A)}\right)\right]$$

Main result:

A) Non-degenerate case: $\exists\, A_1 \subset A_2 \subset \dots \subset A_N = \mathcal{X}$
$\mathrm{card}(A_{k+1} \setminus A_k) = 1$ relatively good sets
$A_k := \{1, \dots, k\}$

Then $\begin{cases} 1 - \lambda_0 = 1 \\ 1 - \lambda_k = \mathbb{P}^{k+1}\{\tau_{A_k}^+ < \tau_{k+1}^+\}\left[1 + O\left(\dfrac{p_{out}(A_k | A_{k+1})}{p_{in}(A_k | A_{k+1})}\right)\right] \end{cases}$

and eigenvectors well-approx. by $\mathbb{P}^x\{\tau_{k+1} < \tau_{A_k}\}$

B) Degenerate case: c.f. M.B.

Example 2 $\quad X_{n+1} = \Pi(X_n) + \sigma \xi_{n+1}$

$\Pi : \mathbb{R}^d \circlearrowleft$ deterministic map, $(\xi_n)_{n\geq 0}$ i.i.d. r.v. with density

$$\mathbb{P}^x\{X_1 \in A\} = \mathbb{P}\{\sigma\xi_1 \in A - \Pi(X)\} = \int_A h(x,y)\,dy$$

Assumptions: $h = h_\sigma$ Markov kernel on $\mathbb{R}^d$

1) $\Pi \in \mathcal{C}^2$, $\Pi : \mathcal{X} \circlearrowleft$ with $\mathcal{X}$ bdd, finitely many limit pts (hyperbolic)
2) LDP with good rate fct $I(x,y)$, $I(x,y) = 0 \Leftrightarrow y = \Pi(x)$
$\Pi(x^*) = x^* \Rightarrow I$ cont. at (x^*, x^*)
3) Positive Harris recurrence ($\Rightarrow \mathbb{E}^x[\tau_A^+] < \infty$ when $\mathrm{Leb}(A) > 0$)
4) Unif. positivity: $\forall x_i^*$ stable fixed pt, $\exists B_i$ nbh of x_i^* s.t.
$B_1 \cup \dots \cup B_N$ $h_{B_i} =: h_i$ satisfies $\sup_{x\in A} h_i^n(x,y) \leq L \inf_{x\in A} h_i^n(x,y)$

Main result:

A) Non-degenerate case (metastable hierarchy of $x_1^*, \dots, x_N^*$)

QSD on B_{k+1}

$$\begin{cases} 1-\lambda_0 = 0 \\ 1-\lambda_k = \mathbb{P}^{\pi_0^{k+1}}\{\tau^+_{B_1\cup\dots\cup B_k} < \tau^+_{B_{k+1}}\}\left[1+O(e^{-\theta/\sigma^2})\right] & 1 \leq k \leq N-1 \\ |\lambda_k| < 1 - \frac{c}{\log(\sigma^{-1})} & k \geq N \end{cases}$$

and eigenfcts well-approximated by $\mathbb{P}^{\pi_0^{k+1}}\{\tau^+_{B_1\cup\dots\cup B_k} < \tau^+_{B_{k+1}}\}$

B) Degenerate case: c.f. M.B.

Approximation result: $K^0_{\text{tr}}(x, dy) = \sum_{k=0}^{N-1} \lambda_k^0 \, \phi_k^0(x) \, \pi_k^0(dy)$

ev right eigenfct left eigenfct

of $K^0 = {}_{B_1\cup\dots\cup B_N}K$

Thm: $\exists \mu_i, \psi_j$ (signed) s.t. $\begin{cases} \|\mu_i - \pi_0^{B_i}\|_1 \lesssim e^{-\theta/\sigma^2} \\ \|\psi_j - 1_{B_j}\| \lesssim e^{-\theta/\sigma^2} \end{cases}$

s.t. $\mathbb{P}^{\mu_i}\{X_{\tau^{+,nm}_{B_1\cup\dots\cup B_N}} \in B_j\} = \mathbb{P}^i\{Y_n = j\} + \underbrace{O(e^{-\theta/\sigma^2})}_{\text{unif. in } n}$

for some $m = m(\sigma)$

where $(Y_n)_{n\geq 0}$ Markov chain with matrix $P_{ij} = \langle \mu_i | (K^0_{\text{tr}})^m \psi_j \rangle$

$$= \mathbb{P}^{\pi_0^{B_i}}\{X_{\tau^{+,nm}_{B_1\cup\dots\cup B_n}} \in B_j\} \times \left[1+O(e^{-\theta/\sigma^2})\right]$$