Descente stochastique du second ordre

S. Gadat

Toulouse School of Economics
Joint work with A. Cabot, H. Engler, L. Miclo, F. Panloup, C. Pellegrini.

Toulouse, June, 26 2015

I - 1 Motivations - Optimisation

On s'intéresse à un problème d'optimisation d'une fonction $U : \mathbb{R}^d \longrightarrow \mathbb{R}$
Descente standard :

$$\dot{x}_t = -\nabla U(x_t) \qquad (G)$$

On connaît beaucoup de choses sur cette dynamique :

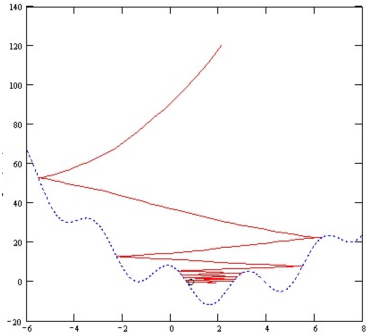
- version continue
- version discrète
- cas fortement convexe
- cas seulement β-convexe
- ...

En particulier, on sait que :

- (G) est piégée dans des minima locaux (non convexité)
- Taux linéaire e^{-Ct} (fortement convexe) ou $1/t$ (β-lisse)

I Motivations - Modélisation & Pièges

Accélérer/éviter pièges ? Idée de physique (Polyak '64, Antipin & Polyak 87) :



- On fait rouler une boule "pesante" sur le graphe de U
- On introduit un effet de friction (amortissement) pour la convergence

Descente *Heavy Ball* :

$$\ddot{x}_t + \gamma_t \dot{x}_t + \nabla U(x_t) = 0 \qquad (HBF)$$

où

$$\lim_{t \to +\infty} \gamma_t = 0.$$

Éviter des minima locaux ? Garder de la mémoire sur les directions passées :

- On introduit deux fonctions h et k positives croissantes
- On produit la descente intégro-différentielle

$$\dot{x}_t = -\frac{1}{k(t)} \int_0^t h(s) \nabla U(x_s) ds \qquad (GM)$$

- Choix typiques :
  - $k(t) = \int_0^t h(s)ds$ ou $\dot{k} = h$
  - $$k(t) = t \quad \text{et} \quad h(t) = 1 \qquad \text{ou} \qquad k(t) = t^{\alpha+1} \quad \text{et} \quad h(t) = (\alpha+1)t^{\alpha}$$
  - $$k(t) = \lambda e^{\lambda t} \qquad \text{et} \qquad h(t) = e^{\lambda t}$$

Remarque : (GM) s'écrit au second ordre :

$$k(t)\ddot{x}(t) + \dot{k}(t)\dot{x}(t) + h(t)\nabla U(x_t) = 0. \qquad (GM')$$

I Motivations - Optimisation des vitesses

Descente accélérée (Nesterov 83) :

$$\begin{cases} x_k = y_{k-1} - \gamma \nabla f(y_{k-1}) \\ y_k = x_k + \frac{k-1}{k+2}(x_k - x_{k-1}). \end{cases} \qquad (N)$$

- Dans les cas fortement convexe ou β-lisse, le schéma du second ordre (N) est meilleur que (G) (il est même optimal dans les deux cas)
- Un petit exercice montre que si γ est petit, alors (N) a une version différentielle du second ordre

$$\ddot{x}_t + \frac{3}{t}\dot{x}_t + \nabla U(x_t) = 0 \qquad (N')$$

II-1 Unification en une seule ODE *via* la représentation HBF

Proposition (CEG, TAMS '09)

$$\dot{x}_t = -\frac{1}{k(t)} \int_0^t h(s) \nabla U(x_s) ds \Longleftrightarrow \ddot{\tilde{x}}_s + a_s \dot{\tilde{x}}_s + \nabla U(\tilde{x}_s) = 0$$

avec $\tilde{x}(s) = x(\tau(s))$ *où*

$$t = \tau(s) \qquad \textit{vérifie} \qquad \dot{\tau}(s) = \sqrt{(k.h^{-1})(\tau(s))} \qquad \textit{et} \qquad a_s = \frac{(\dot{k}h + k\dot{h})}{2h^{3/2}k^{1/2}} \circ \tau(s)$$

Cas quadratique : fonctions de Bessel dès lors que mémoire polynomiale.

Proposition

Si $U(x) = x^2/2$ *et* $h(t) = t^\alpha$ *et* $k(t) = t^\beta$.

$$x_t \sim C t^{-(\alpha+\beta)/4} \cos(\mu t^{(\alpha-\beta+2)/2} + \phi) \quad \textit{lorsque} \quad t \longrightarrow +\infty$$

- Oscillation lorsque $t \to \infty$ avec damping autour de $\arg\min U$.
- De manière plus générale, les systèmes déterministes à mémoire ont une inertie qui implique des oscillations localement.

II-1 Nesterov accéléré et Gradient à mémoire

En particulier :

- $k(t) = \lambda e^{\lambda t}$ et $h(t) = e^{\lambda t}$ correspond à un amortissement constant
$$a_s = \sqrt{\lambda}/2 \qquad \text{et} \qquad \tau(s) = \sqrt{\lambda} s$$
- $k(t) = t^{\alpha+1}$ et $h(t) = (\alpha+1)t^{\alpha}$ correspond à un amortissement évanescent :
$$a_s = \frac{2\alpha+1}{t} \qquad \text{et} \qquad \tau(s) = \frac{s^2}{4(\alpha+1)}.$$

Proposition (CEG, TAMS '09)

Nesterov accéléré est un (GM) $h(t) = 2t$ *et* $k(t) = t^2$ *et* $\tau(s) = \frac{s^2}{8}$.

II-2 Premières propriétés

$$\ddot{x}_t + \gamma_t \dot{x}_t + \nabla U(x_t) = 0$$

Hypothèses

- U est une fonction de $\mathbb{R}^d$ dans $\mathbb{R}_+^\star$ régulière et coercive :

$$\lim_{x \to \infty} U(x) = +\infty$$

- On suppose de plus que $U(0) = \min_{\mathbb{R}^d} U > 0$
- On suppose que U satisfait la condition de rappel :

$$\lim_{x \to \infty} \inf \langle x, \nabla U(x) \rangle > 0$$

- γ est la fonction d'amortissement définie de $\mathbb{R}_+$ dans $\mathbb{R}_+$.

Existence

Proposition

Si on note $\mathcal{E}(t) = U(x_t) + \frac{|\dot{x}_t|^2}{2}$, alors $\dot{\mathcal{E}}(t) = -\gamma_t |\dot{x}(t)|^2$ et les solutions de (GM) sont définies sur $\mathbb{R}_+$ et restent bornées.

$$\forall t > 0 \qquad \mathcal{E}(t) - \min U \geqslant (\mathcal{E}(0) - \min(U)) \exp\left(-2\int_0^t \gamma_s ds\right).$$

Ainsi, même si U est convexe, si γ tend vers 0 trop rapidement :

$$\int_0^{+\infty} \gamma_s ds < +\infty \Longrightarrow \text{la trajectoire ne peut converger.}$$

II-3 Estimées d'énergie - cas convexe -

$$\ddot{x}_t + \gamma_t \dot{x}_t + \nabla U(x_t) = 0$$

Hypothèse de type convexité (un peu plus général que la convexité, $\theta = 1$).

$$\exists \theta \quad \forall x \in \mathbb{R}^d \qquad U(x) \leqslant \theta \langle \nabla U(x); x \rangle$$

Proposition (CEG, TAMS'09.)

Si γ est $\mathcal{C}^1$ et décroissante, alors

$$\int_0^{+\infty} \gamma_s [\mathcal{E}(s) - minU] ds < +\infty$$

Si de plus $\int_0^{+\infty} \gamma_s ds = +\infty$ (cas de décroissance lente), alors

$$\lim \mathcal{E}(t) = \min U.$$

Si enfin $\arg\min U = \{0\}$, la trajectoire converge.

II-3 Estimées d'énergie - cas convexe -

$$\ddot{x}_t + \gamma_t \dot{x}_t + \nabla U(x_t) = 0$$

Rien n'est dit ici sur la vitesse. On peut aller plus loin :

Theorem (Vitesse de convergence de GM - CEG '09)

- *S'il existe* K *t.q.* $\dot{\gamma}_t + K\gamma_t^2 \leqslant 0$ *et* ∇U *Lipschitz, si* $m = min((\theta + 1/2)^{-1}; K)$

$$\mathcal{E}(t) \lesssim \exp\left(-m \int_0^t \gamma_s ds\right)$$

- *S'il existe* $K \leqslant (\theta + 1/2)^{-1}$ *t.q.* $\dot{\gamma}_t + K\gamma_t^2 \geqslant 0$

$$\mathcal{E}(t) \lesssim \gamma_t$$

- Cas de mémoire polynomiale $\gamma_s = \frac{2\alpha+1}{s} \Longleftrightarrow K = \frac{1}{2\alpha+1}$.
- Si $\theta = 1$, fonction convexe, on a $m = K$ dès que $\alpha > 0$. Pour les $\alpha \in \mathbb{N}$, le plus petit damping Nesterov est $\gamma_t = 3/t$.
- On obtient une vitesse en $1/t$ pour le (GM) et (N) dans ce cas. On peut améliorer en $t^{-2/\theta}$ dans ce contexte pour $\theta > 1/2$.

II-4 Convergence des trajectoires - cas convexe

Cas où $U = 0$

- Dans le cas où $U = 0$, on peut calculer explicitement les solutions

$$x_t = x(0) + \dot{x}(0) \int_0^t e^{-\int_0^s \gamma_u du} ds$$

- Convergence si γ ne décroît pas trop vite vers 0 (cas très lent) :

$$\int_0^t e^{-\int_0^s \gamma_u du} ds < \infty.$$

Généralisation

- Dimension d, avec U convexe et $\arg\min U$ *lisse* et non réduit à un point.
- On suppose que $\int_0^t e^{-\int_0^s \gamma_u du} ds = +\infty$.

Theorem (Non convergence CEG, TAMS '09)

Si $(x(0), \dot{x}(0)) \notin \arg\min U \times 0$*, alors la trajectoire ne converge pas.*

Cas de la dimension 1 Potentiel convexe, gradient Lipschitz

Proposition (Convergence des trajectoires, CEG EJDE '09)

Si $\gamma_s = c/s$ *avec* $c > 1$*, les trajectoires convergent en dimension* 1*.*

II-5 Développements

- Résultats positifs de convergence $1D$ dans des cas non convexes.
- Convergence des trajectoires dans $\mathbb{R}^d$ (récent travail de Cabot, 2014).
- Comment traiter le cas des potentiels non convexes généraux ?

III-0 Recuit simulé

Historique :

- Algorithme initial : Kirkpatrick, Gelatt, Vecchi, Sience,'83
- Première preuve : Hajek, M.O.R. '88
- version proba : Holley, Kusuoka, Stroock , J.F.A. '89, Miclo, I.H.P. '92

Quelques rappels sur la méthode du recuit simulé

- U potentiel $\mathcal{C}^2$ de $\mathbb{R}^d$ dans $\mathbb{R}_+$ non convexe (unique minimum absolu)

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

- Si U assez coercif, $(X_t)_{t\geqslant 0}$ a une mesure invariante unique et vérifie une inégalité de type trou spectral/ log sobolev :

$$KL(\mathcal{L}(X_t), \pi_\sigma) \lesssim e^{-Ct}$$

- La mesure invariante est

$$\pi_\sigma \propto e^{-U/\sigma^2}$$

  et si $\sigma \longrightarrow 0$, alors π_σ se concentre sur les minima globaux de U
- Idée : ergodicité de $(X_t)_{t\geqslant 0}$ + schéma de température adapté σ_t.
- On cherche une inégalité différentielle sur

$$D_t := KL(\mathcal{L}(X_t), \pi_{\sigma_t}) \qquad \text{du type} \qquad \dot{D}_t \leqslant -C_t D_t + \epsilon_t$$

- $X_t \longrightarrow \arg\min U$ lorsque $t \longrightarrow +\infty$ si on calibre correctement (C_t, σ_t)

III-1 Diffusion moyennée par sa mémoire

- On se donne une fonction k croissante, positive, de classe $\mathcal{C}^2$.
- On définit $(B_t)_{t\geqslant 0}$ un mouvement Brownien standard, et $(X_t)_{t\geqslant 0}$ est le processus solution de l'eds :

$$dX_t = -\left[\frac{1}{k(t)}\int_0^t \dot{k}(s)\nabla U(X_s)ds\right]dt + \sigma(X_t)dB_t. \qquad (DM)$$

- <u>Problème :</u> $(X_t)_{t\geqslant 0}$ n'a aucune chance d'être Markovien.
- En posant $Y_t = \frac{1}{k(t)}\int_0^t \dot{k}(s)\nabla U(X_s)ds$, on remarque que (X, Y) satisfait

$$\begin{cases} dX_t = -Y_t dt + \sigma(X_t)dB_t \\ dY_t = r(t)(\nabla U(X_t) - Y_t)dt. \end{cases} \qquad (1)$$

avec une fonction importante :

$$r(t) = \frac{\dot{k}(t)}{k(t)}$$

- **Le processus $(X_t, Y_t)_{t\geqslant 0}$ est Markovien** inhomogène en temps tel que :

$$\mathcal{A}f(x,y,t) = -\langle y, \partial_x f\rangle + r(t)\langle \nabla U(x) - y, \partial_y f\rangle + \frac{1}{2}\mathrm{Tr}\left(\sigma^*(x)D_x^2 f\sigma(x)\right) + \partial_t f.$$

III-2 Diffusion moyennée et Fokker Planck cinétique

- Cela ressemble à léquation de Fokker Planck cinétique :

$$\begin{cases} dX_t = Y_t dt \\ dY_t = -[\nabla U(X_t) + Y_t]dt + \sigma dB_t \end{cases} \tag{2}$$

- dB_t est placé sur la coordonnée position et non vitesse (aie !) .
- L'équation d'évolution est dégénérée sur une coordonnée. Peut-on obtenir malgré tout
- Existence et de régularité des densités de (X_t, Y_t) ? (Hypoellipticité)
- Vitesses de convergence ? (contraction du semi-groupe)
- Des asymptotiques lorsque $\sigma \to 0$? (grandes déviations)

FP cinétique est finalement une histoire plus simple (controllabilité, convergence, mesure invariante explicite, . . .)

III-3 Propriétés élémentaires

Proposition (Existence forte du processus)

Étant donné $T > 0$ et un M.B. $(B_t)_{t \geqslant 0}$, il existe une unique solution de (1). De plus, si $\mathbb{E}[U(X_0) + Y_0^2] < \infty$, alors $\sup_{t \leqslant T} \mathbb{E}[U(X_t) + Y_t^2] < \infty$.

Démonstration.

Poser $h(x, y, t) = U(x) + y^2/(2r(t))$ et montrer que

$$\mathcal{A}h(x, y, t) = Tr(\sigma^* D^2 U(x)\sigma) + |y|^2(-1 + \frac{\dot{r}(t)}{2r^2(t)}) \leqslant C_T h.$$

$\square$

Comportement selon la mémoire $r_\infty = \lim r$, nous étudierons 3 cas :

1. $r_\infty \in]0; +\infty[$: stabilité et comportement non standard $k(t) \sim_{t\to\infty} e^{\lambda t}$.
2. $r_\infty = 0$: instabilité du processus. $k(t) \sim_{t\to\infty} t^\beta, \beta > 0$.
3. $r_\infty = +\infty$: stabilité et comportement standard $k(t) \sim_{t\to\infty} e^{t^\alpha}, \alpha > 1$.

Le cas standard est celui de l'EDS :

$$dS_t = -\nabla U(S_t)dt + \sigma d\tilde{B}_t.$$

III-4 Régularité du processus [Hypoellipticité]

- On va démontrer que les lois sont régulières.
- C'est non trivial en raison de la dégénérescence de la diffusion sur Y.

Hypoellipticité : ($\sigma > 0$ et U infiniment dérivable)

$$\mathcal{E}_U = \left\{ x \in \mathbb{R}^d, \forall 1 \leqslant k \leqslant d : \exists (i_1, \ldots i_p) \, \partial_{x_{i_p} \ldots x_{i_1}, x_k} U(x) \neq 0 \right\}$$

Theorem (Hypoellipicité)

Si $\sigma^* \sigma \geqslant \epsilon_0 Id$ *et* $\dim(\mathbb{R}^d - \mathcal{E}_U) \leqslant d - 1$*, alors* (X_t^z, Y_t^z) *admet une densité* $\mathcal{C}^\infty$ *par rapport à la mesure de Lebesgue, notée* $p_t(z, .)$*.*

Démonstration.

Appliquer [H] version inhomogène en temps (Cattiaux, PTRF'02) avec

$$L_D = -\langle y, \partial_x \rangle + r(t) \langle \nabla U(x) - y, \partial_y \rangle, L_\sigma = \frac{\sigma^2}{2} \Delta_x, L_Z = L_D - L_\sigma$$

et montrer que $\dim$ span Lie $\left(\frac{\partial}{\partial t} + L_Z, \sigma_1, \ldots, \sigma_d \right) (x, y, \xi) = 2d + 1$, Les dimensions manquantes proviennent de

$$\left[\partial_{x_i}, \frac{\partial}{\partial t} + L_D \right].$$

□

III-4 Régularité du processus [Hypoellipticité]

- On va démontrer que les lois sont positives pour $t > 0$.
- C'est non trivial en raison de la dégénérescence de la diffusion sur Y.

Contrôlabilité :

$$\begin{cases} \frac{dx_t^\varphi}{dt} = -y_t^\varphi + \dot{\varphi}_t \\ \frac{dy_t^\varphi}{dt} = r(t)[\nabla U(x_t^\varphi) - y_t^\varphi] \end{cases}$$

Theorem (Contrôlabilité)

Si $\lim U(x)/|x| = +\infty$, le système est approximativement contrôlable et localement exactement contrôlable aux min locaux de U ($D^2 U > 0$).

Démonstration.

i) : « Irréductibilité » du processus par la rep . barycentre des $\nabla U(x)$ du y.
ii) : Cond. Kalman et calcul de Malliavin (Delarue & Menozzi, J.F.A. '10) $\square$

On peut alors conclure que la densité $p_t(z, .)$ est $\mathcal{C}^\infty$ sur $\mathbb{R}^{2d}$ et positive.

III-5 Ergodicité du processus [Hypocoercivité]

- **On se place dans le cas $r_\infty > 0$ et σ constant**
- C'est le cas de type mémoire exponentielle.
- Condition de rappel « convexité à l'infini » : $(\mathbf{H_a})$: Il existe $a \in (0, 1]$:

$$\frac{|\nabla U(x)|^2}{U(x)} << (U(x) \vee |x|^2)^a << \langle x, \nabla U(x) \rangle$$

Theorem

Sous $(\mathbf{H_a})$ + condition d'hypoellipticité :

$$\sup_{\{f, |f| \leqslant 1\}} |P_t^{r_\infty}(z_0, f) - p_\infty(f)| \leqslant C_{a,r_\infty}(z_0) \begin{cases} \exp(-\delta_{r_\infty} t) & \text{si } a = 1 \\ t^{-\frac{a}{1-a}} & \text{si } a \in (0, 1). \end{cases}$$

- Si $U(x) \sim |x|^q$ avec $q > 0$, on peut choisir $a = 1 \wedge q/2$.
- C'est un cadre compatible avec la méthode de type recuit simulé.

III-5 Ergodicité du processus [Hypocoercivité]

- **On se place dans le cas $r_\infty = 0$ et σ constant**
- C'est le cas de type mémoire polynomiale $k(t) = t^\alpha$ avec $\alpha > 0$ ou $k(t) = e^{t^\alpha}$ avec $\alpha \in (0, 1)$.

Theorem

Si U est sous-quadratique, alors

$$\lim \sup_{t \longrightarrow +\infty} r(t)\mathbb{E}[X_t^2] = +\infty.$$

On peut même affiner le résultat dans le cas Gaussien $U(x) = \frac{x^2}{2}$.

Theorem

Si $k(t) = (1+t)^\alpha$ avec $\alpha > 1/2$, on a :
i) Le processus $(X_t, Y_t)_{t \geqslant 0}$ est asymptotiquement centré et satisfait

$$\lim_{t \to \infty} \mathbb{E}Y_t^2 = \frac{\alpha}{2\alpha+1}, \qquad \textit{et} \qquad \mathbb{E}X_t^2 \sim \frac{t}{2\alpha+1} \quad \textit{si} \quad t \to +\infty.$$

ii)

$$\left(\sqrt{\frac{2\alpha+1}{t}} X_t, \sqrt{\frac{2\alpha+1}{\alpha}} Y_t\right) \overset{\mathcal{L}}{\Longrightarrow} \mathcal{N}(0, I_2) \quad \textit{si} \quad t \to +\infty.$$

Mémoire polynomiale semble une mauvaise idée, sauf si $\sigma(t) \longrightarrow 0$ vite...

Identification de la limite dans le cas homogène $0 < r_\infty < +\infty$

Proposition (Identification de la loi stationnaire, cas homogène)

Sous les hypo-hypothèses $r(t) = r_\infty \in]0; +\infty[$, p_∞ *est* $\mathcal{C}^\infty$ *et satisfait*

$$\langle y, \partial_x p_\infty \rangle + \frac{1}{2} Tr(\sigma^* D_x^2 p_\infty \sigma) + r_\infty \left(\langle y - \nabla U(x), \partial_y p_\infty \rangle + p_\infty \right) = 0.$$

- Dans le cas gaussien ($U(x) = x^2/2$) et σ constant, p_∞ est une mesure gaussienne centrée de matrice de covariance « tordue »

$$\Sigma^2(r_\infty) = \frac{\sigma^2}{2} \begin{pmatrix} \frac{r_\infty + 1}{r_\infty} & 1 \\ 1 & 1 \end{pmatrix}.$$

- C'est aussi vrai dans le cas inhomogène $r(t) \to r_\infty \in]0; +\infty[$.
- Dans le cas $\sigma = \sigma Id$ constant, p_∞ satisfait

$$\langle y, \partial_x p_\infty \rangle + \frac{\sigma^2}{2} \Delta_x p_\infty + r_\infty \left(\langle y - \nabla U(x), \partial_y p_\infty \rangle + p_\infty \right) = 0.$$

IV-1 Comportement à basse température

- On se restreint au cas de la mémoire exponentielle $k(t) = e^{\lambda t}$.
- On cherche le comportement de μ_σ lorsque $\sigma \longmapsto 0$.

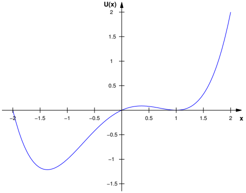
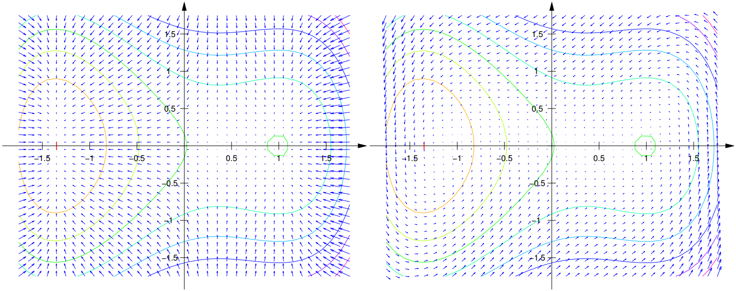


FIGURE : Potentiel $U = x^4/4 - 3x^2/2 + x/5$, Lignes de champ de $U(x) + y^2/2$ pour le cas standard et le cas à mémoire.



IV-1 Comportement à basse température

- Dans le cas « standard » :

$$\mathcal{L}_K = -\nabla U(x)\partial_x + \sigma^2 \partial_x^2,$$

  la distribution invariante est le champ de Gibbs $\nu_\sigma \propto e^{-U/(2\sigma^2)}$.
  Méthode de Laplace $\Rightarrow P.G.D. + \nu_\sigma(\arg\min U) \to_{\sigma\to 0} 1$.
- Pour nous, (hormis le cas gaussien) pas de forme explicite pour $\mu_\sigma \ldots$
- Cela complique singulièrement l'analyse des P.G.D.

$$\begin{cases} dX_t = -Y_t + \sigma dB_t \\ dY_t = \lambda[\nabla U(X_t) - Y_t] \end{cases} \quad \text{vs} \quad \begin{cases} \dot{x} = -y^\varphi + \dot{\varphi} \\ \dot{y} = \lambda[\nabla U(x^\varphi) - y^\varphi] \end{cases}$$

- $\mathbf{z}_\varphi(z,.)$ trajectoire initialisée en z avec un contrôle φ.
- φ dans l'espace de Cameron-Martin $\mathbb{H}$
- Coût de jonction de z_1 à z_2 en temps t (ou quelconque) :

$$I_t(z_1,z_2) = \inf_{\mathbf{z}_\varphi(z_1,t)=z_2,\varphi\in\mathbb{H}} \frac{1}{2}\int_0^t |\dot{\varphi}|^2 \qquad I(z_1,z_2) = \inf_{t\geqslant 0} I_t(z_1,z_2)$$

IV-1 Comportement à basse température

Theorem (GPP, EJP '12)

- *Tension exponentielle de* (μ_σ)
- *Si équilibres* $(z_i)_i$ *discrets avec* $D^2U(z_i)$ *inversible, à extraction près*

$$\lim_{\sigma\to 0} \sigma^2 \log[\mu_\sigma(z)] = -W(z) \quad \textit{avec} \qquad \forall t \geqslant 0 \quad W(z_2) = \inf_{z_1} I_t(z_1, z_2) + W(z_1)$$

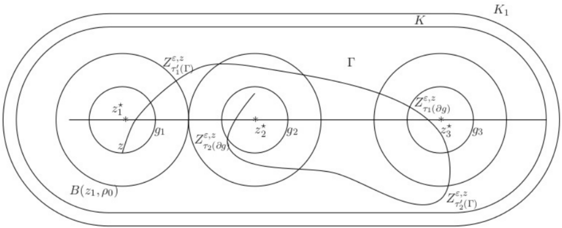
Quelques remarques importantes :

- Formulation HJB ou EDP :

$$-\langle y, \partial_x, W\rangle + \langle \nabla U(x) - y, \partial_y W\rangle + \frac{|\partial_x W|^2}{2} = 0$$

- Inefficience des méthodes EDP pour obtenir l'unicité de la limite.
- Utiliser en plus l'info de la nature « mesure invariante » d'un Markov.

IV-2 Théorie de Freidlin & Wentzell



- μ_σ se représente au travers des transitions de la chaîne squelette
- Utilisation cruciale des voisinages des équilibres $(g_i)_{i=1\ldots L}$ du champs

$$\mu_\sigma(A) = \int_{\partial g} \tilde{\mu}_\sigma^{\partial g}(dz)\mathbb{E}_z \int_0^{\tau_1(\partial g)} \mathbf{1}_{Z_s^{z,\sigma} \in A} ds$$

- Il suffit alors plus simplement de comprendre le mécanisme de transition de la chaîne squelette entre les $(g_i)_{1\leqslant i\leqslant L}$

IV-3 $\{i\}$-graphe et conclusion

On introduit le quasi-potentiel V :

$$V(z_i, z_j) = \inf_{\begin{cases} \varphi \in \mathbb{H}^1 \\ z_\varphi(0) = z_i, z_\varphi(\infty) = z_j \end{cases}} \left[\frac{1}{2} \int_0^{+\infty} \dot{\varphi}^2 \right],$$

V règle l'asymptotique des transitions de la chaîne squelette.
$G(i)$: ensemble des $\{i\}$-graphes : graphe orienté se terminant par $\{i\}$ pour lequel tout nœud possède une unique arête sortante.

$$\mathcal{W}(z_i) := \min_{g \in G(i)} \sum_{(z_k \to z_l) \in g} V(z_k, z_l)$$

Theorem (GPP, EJP'12)

Sous les conditions précédentes, W est unique, continue au vois des $(z_i)_i$, et

$$W(z_i) = \mathcal{W}(z_i) - \min_i \mathcal{W}(z_i).$$

μ_σ se concentre exp vite sur le min global de $\mathcal{W}$.

Conclusion

- Bonne connaissance du système (G).
- Pour la diffusion (DM)
  - $k(t) = e^{\lambda t}$ Comportement stable mais non standard.
  - Si $k(t) = e^{e^t}$ (non exposé aujourd'hui), comportement comparable à la diffusion
    $$dS_t = -\nabla U(S_t)dt + \sigma(S_t)dB_t.$$
- Diffusion à petits paramètres
  - Existence d'un PGD
  - Description précise des minimiseurs
- Élargissements :
- Vrai recuit : coupler $\sigma_t \to 0$ avec vitesse L^2
- Borne inférieure de vitesses inconnues.
- Modélisation, étudier des systèmes à mémoire ayant une écriture comparable (bulles financières, dynamique des populations, ...)
- Passage en algo sto des méthodes d'ordre 2. Vitesse, renormalisation, processus limite ?

Merci de votre attention !