

Variations autour de la Descente de gradient à mémoire.

A. Cabot, H. Engler, S. G., L. Miclo, F. Panloup, C. Pellegrini

Institut de Mathématiques de Toulouse
Séminaire de l'équipe M.I.P.

Juin 2011

Plan de l'exposé

Introduction

Étude de l'ODE (SGM')

Modèle diffusif moyenné (DM)

Conclusion

Introduction

Dans cet exposé, on s'intéresse au comportement en temps long de :



$$dx_t = - \left[\frac{1}{k(t)} \int_0^t h(s) \nabla U(x_s) ds \right] dt, \quad (GM)$$

où U est un potentiel coercif strictement positif et $(x_t)_{t \geq 0}$ vit dans \mathbb{R}^d .

- ▶ Ainsi qu'à sa généralisation naturelle :

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

où B_t est un mouvement brownien d -dimensionnel et σ est une matrice de variance/covariance elliptique.

- ▶ h et k sont supposées continues, positives et croissantes. Un cas particulier raisonnable correspond à $k = \int h$.
- ▶ Objectif 1 : Comparer le comportement en temps long des solutions de (GM) et (DM) avec des descentes de gradients/diffusions classiques :

$$d\tilde{x}_t = -\nabla U(\tilde{x}_t) dt \quad \text{et} \quad d\tilde{X}_t = -\nabla U(\tilde{X}_t) dt + \sigma(\tilde{X}_t) d\tilde{B}_t,$$

- ▶ Objectif 2 : Pour le système (3), quel est le comportement de la diffusion avec petit paramètre σ ?

Introduction

Le modèle (GM)

$$dx_t = - \left[\frac{1}{k(t)} \int_0^t h(s) \nabla U(x_s) ds \right] dt, \quad (GM)$$

est équivalent au système

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

Après changement de temps,

$$\dot{\tau}(s) = \sqrt{\frac{k(\tau(s))}{h(\tau(s))}},$$

et en notant $z = x \circ \tau$, z est solution de

$$\ddot{z}(s) + \gamma(s)\dot{z}(s) + \nabla U(z(s)) = 0, \quad (SGM')$$

où

$$\gamma(s) = \left(\frac{\dot{k}h + k\dot{h}}{2h^{3/2}k^{1/2}} \right) \circ \tau(s).$$

Un petit problème de dynamique déterministe

- ▶ On considère le mouvement d'une boule pesante posée sur le graphe d'une fonction U .
- ▶ La boule est soumise à la pesanteur et à une force de frottement opposée à son déplacement, le système dissipatif est solution de l'o.d.e.

$$\ddot{x}_t + \lambda \dot{x}_t + \nabla U(x_t) = 0 \quad (1)$$

et du moment que U est convexe, la solution de (1) converge vers l'unique minimum de U .

- ▶ Plus généralement, (GM') et (SGM') sont une généralisation de (1) avec un amortissement évanescent (ou non)...
- ▶ Pour des facilités d'écriture, on s'intéressera dans le cadre déterministe non diffusif à la formalisation (SGM') :

$$\ddot{z}(s) + \gamma(s)\dot{z}(s) + \nabla U(z(s)) = 0, \quad (SGM')$$

- ▶ Les fonctions mémoires h et k sont "résumées" dans la fonction γ :

$$\gamma = \left(\frac{\dot{k}h + k\dot{h}}{2h^{3/2}k^{1/2}} \right) \circ \tau.$$

Un petit problème de dynamique déterministe (influence de la mémoire)

Cas particulier du système déterministe :

$$U(x) = x^2/2, k(t) = t^\beta, h(t) = t^\alpha$$

- ▶ Si $\beta = \alpha + 2$, alors

$$x(t) = t^{-(\alpha+1)/2} r_\alpha(t)$$

.

- ▶ Si $\beta \neq \alpha + 2$, alors

$$x(t) = t^{-(\beta-1)/2} J_{\alpha,\beta}(t).$$

- ▶ Oscillation très lentes lorsque $t \rightarrow \infty$ avec phénomène de damping autour du minimum de U .
- ▶ De manière plus générale, les systèmes déterministes à mémoire ont une inertie qui implique des oscillations localement.
- ▶ D'un point de vue optimisation : **cette inertie peut permettre de passer certains maxima locaux.**

Plan de l'exposé

Introduction

Étude de l'ODE (SGM')

Existence

Cas convexe

Cas non convexe

Bilan sur la fonction mémoire pour (GM) et (SGM')

Modèle diffusif moyenné (DM)

Conclusion

Notations, hypothèses sur le potentiel U et σ

- ▶ U est une fonction de \mathbb{R}^d dans \mathbb{R}_+^* coercive :

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

- ▶ On suppose de plus que

$$\min_{\mathbb{R}^d} U > 0$$

- ▶ On suppose que U satisfait la **condition de rappel** :

$$\liminf_{x \rightarrow \infty} \langle x, \nabla U(x) \rangle > 0$$

- ▶ U est supposée au moins \mathcal{C}^3 , et on notera D^2U la matrice

$$(D^2U(x))_{i,j} = (\partial_{x_i} \partial_{x_j} U)(x)$$

- ▶ γ est une fonction continue définie de \mathbb{R}_+ dans \mathbb{R}_+ .
- ▶ On pourrait étendre l'étude au cas où U est définie sur un Hilbert.

Propriétés d'existence

On obtient l'existence via une fonction de Lyapunov \mathcal{E} .

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 (SGM') sont définies sur \mathbb{R}_+ et restent bornées.

Le rôle de la fonction mémoire est identifié par le biais de la minoration

$$\forall t > 0 \quad \mathcal{E}(t) - \min U \geq (\mathcal{E}(0) - \min(U))e^{-\int_0^t \gamma(s)ds}.$$

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

$$\int_0^{+\infty} \gamma(s)ds < +\infty,$$

la trajectoire ne peut converger.

Estimées d'énergie - cas convexe -

On se place dans le cadre où U satisfait une inégalité de type convexité (0 est le point où U atteint son minimum) :

$$\exists \theta \quad \forall x \in \mathbb{R}^d \quad U(x) - U(0) \leq \theta \langle \nabla U(x); x \rangle$$

On obtient alors la majoration de la propriété suivante.

Proposition (A. Cabot, H. Engler, S. G.)

Si γ est C^1 et décroissante, alors

$$\int_0^{+\infty} \gamma(s) [\mathcal{E}(s) - \min U] ds < +\infty$$

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.

Convergence des trajectoires - cas convexe - dimension 1

Cas où $U = 0$

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

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

- ▶ Il n'y a convergence que si γ ne décroît pas trop vite vers 0 (cas très lent) :

$$\int_0^t e^{-\int_0^s \gamma(u) du} ds < \infty.$$

- ▶ Bien sûr, le cas très lent implique le cas lent ($\int_0^{+\infty} \gamma = +\infty$).

Généralisation

- ▶ On se place en dimension 1, avec U convexe et $[\alpha, \beta] = \arg \min U$.
- ▶ On suppose que $\int_0^t e^{-\int_0^s \gamma(u) du} ds = +\infty$.

Theorem (A. Cabot, H. Engler, S. G.)

Si $(x(0), \dot{x}(0)) \notin [\alpha, \beta] \times 0$, alors $[\alpha, \beta] \subset \omega(x(0), \dot{x}(0))$ et la trajectoire en converge pas.

Convergence des trajectoires - cas convexe - dimension 1 (suite)

L'hypothèse précédente est presque nécessaire et suffisante pour la non-convergence des trajectoires :

- ▶ On se place en dimension 1 et on suppose de plus que U est "convexe"
- ▶ U possède un plateau de minima

$$[\alpha, \beta] = \arg \min U.$$

- ▶ γ décroît vers 0 telle que

$$\exists \theta < 1 \quad \int_0^\infty e^{-\theta \int_0^s \gamma(u) du} ds < \infty.$$

Theorem (A. Cabot, H. Engler, S. G.)

Sous les hypothèses précédentes, la trajectoire converge.

Remarque : Il manque le cas où $\theta = 1$... mais dans le cas particulier où $\gamma(s) = \frac{c}{s+1}$, on a

$$e^{-\int_0^s \gamma(u) du} ds = (s+1)^{-c},$$

on sait montrer que si $c \in (0; 1]$, la trajectoire ne converge pas alors que si $c > 1$, elle converge. On devrait donc être capable d'aller jusqu'à $\theta = 1$ pour obtenir la convergence.

Convergence des trajectoires - cas convexe - dimension n

- ▶ Cas où

$$\int_0^{\infty} e^{-\int_0^s \gamma(u) du} ds = +\infty.$$

- ▶ On peut étendre les résultats de non convergence précédents à la dimension n dans le cas convexe.
- ▶ Hypothèses pas trop restrictives sur la géométrie de l'argmin de U (hypothèse sur le cône tangent, le cône normal de l'ensemble des minimiseurs).
- ▶ Questions ouvertes :
 - ▶ Résultats de convergence lorsque

$$\int_0^{\infty} e^{-\int_0^s \gamma(u) du} ds < +\infty?$$

- ▶ Résultats de non convergence lorsque

$$\int_0^{\infty} e^{-\int_0^s \gamma(u) du} ds = +\infty,$$

et

$$-N_S(y) \notin \text{int}(T_S(y)) \cup \{0\}?$$

Convergence des trajectoires - cas non convexe

- ▶ On se place dans \mathbb{R}^d .
- ▶ On suppose que U possède m points critiques $U(x_1) < U(x_2) \cdots < U(x_m)$.
- ▶ On se place sous l'hypothèse que $\int_0^\infty \gamma = \infty$:

Proposition

$$\exists !x_i \quad \lim_{t \rightarrow +\infty} \mathcal{E}(x(t)) = U(x_i).$$

De plus, $\liminf |x(t) - x_i| = 0$.

- ▶ Dans le cas particulier où $\gamma(t) \geq c/(t+1)$, avec $c > 0$, on a un résultat plus fort

Theorem (A. Cabot, H. Engler, S.G.)

Pour tout $\epsilon > 0$, il existe un unique x_i tel que

$$\lim_{T \rightarrow +\infty} \frac{1}{T} |\{t \leq T \mid |x(t) - x_i| > \epsilon\}| = 0.$$

On peut récrire l'ergodicité en

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt = x_i.$$

Convergence des trajectoires - cas non convexe unidimensionnel

- ▶ On se place dans \mathbb{R} .
- ▶ On suppose que U possède m points critiques $U(x_1) < U(x_2) \cdots < U(x_m)$ et pour tout $i : U''(x_i) \neq 0$
- ▶ Dans le cas particulier où $\gamma(t) \geq c/(t+1)$, avec $c > 0$, on a

Theorem (A. Cabot, H. Engler, S.G.)

Quels que soient les points d'initialisation, $\lim_{t \rightarrow \infty} x(t) = x^$ existe et appartient à $\{x_1, \dots, x_m\}$. Si on note \mathcal{T} l'ensemble des temps de changement de signe de \dot{x} , alors*

$$|\mathcal{T}| = +\infty \iff x^* \text{ est un minimum (local) de } U.$$

Par ailleurs, l'ensemble des points d'initialisation tel que x^ est un minimum local est un ouvert dense.*

Conclusion sur le rôle de γ

$$\gamma(s) = \left(\frac{\dot{k}h + k\dot{h}}{2h^{3/2}k^{1/2}} \right) \circ \tau(s).$$

- ▶ Dès que h et k sont croissantes et positives strictement, la trajectoire de (GM) existe sur $[0; +\infty[$ et reste bornée.
- ▶ Si $h/k \in L^2(\mathbb{R}_+)$, $\lim \tau < +\infty$. (GM) n'est qu'une portion finie de (SGM') après changement de temps. On a donc convergence de la trajectoire de (GM) mais vers rien de pertinent.
- ▶ Si on considère $h(t) = t^\alpha$ et $k(t) = t^\beta$ alors
 - ▶ Si $\beta - \alpha \leq 2$ et $\alpha + \beta > 0$, alors $\lim \mathcal{E} \circ \tau(t) = \min U$
 - ▶ Si $\beta \leq 1$ et $\alpha \geq 0$, toute solution non stationnaire diverge.
 - ▶ Si $\beta > 1$ et $\beta - \alpha \leq 2$, alors en dimension 1 on a convergence de la trajectoire vers un point critique de U (la plupart du temps un minimum local).
- ▶ Si on considère $h(t) = e^t$ et $k(t) = e^t$, alors $\tau(s) = s$ et $\gamma(s) = s/2$, on retrouve le système (HBF) et tous les résultats de convergence sont vrais.
- ▶ On peut imaginer encore d'autres situations : $h(t) = e^{t^\alpha}$ avec $\alpha < 1$...

Plan de l'exposé

Introduction

Étude de l'ODE (SGM')

Modèle diffusif moyenné (DM)

Définition de la diffusion moyennée

Hypoellipticité

Existence d'un régime stationnaire

Identification de la limite

Convergence lorsque σ tend vers 0

Conclusion

Propriétés d'existence

- ▶ On se donne une fonction k croissante, positive, de classe C^2 . On se limitera ici au cas où $\dot{k}(t) = e^{\lambda t}$.
- ▶ On saurait dire pas mal de choses pour des fonctions mémoires générales.
- ▶ On définit $(B_t)_{t \geq 0}$ un mouvement Brownien standard, et $(X_t)_{t \geq 0}$ est le processus solution de l'eds :

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

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

$$dX_t = -Y_t dt + \sigma(X_t) dB_t \quad \text{et} \quad dY_t = r(t) [\nabla U(X_t) - Y_t] dt,$$

avec $r(t) = \dot{k}(t)/k(t) = \lambda$ (ici). Le processus $(X_t, Y_t)_{t \geq 0}$ est Markovien.

- ▶ Pour plus de simplicité, on considère X dans \mathbb{R} avec $\sigma(x) = \sigma$ constant, même si un grand nombre de résultats s'étendent au cas multi-dimensionnel.

Processus Markovien

- ▶ On définit $(X_t, Y_t)_{t \geq 0}$ solution du processus de descente de gradient moyenné par

$$\begin{cases} dX_t = -Y_t dt + \sigma dB_t \\ dY_t = \lambda(\nabla U(X_t) - Y_t) dt. \end{cases} \quad (2)$$

- ▶ (X_t, Y_t) est un processus Markovien homogène en temps tel que $\forall f \in \mathcal{C}_K^2(\mathbb{R} \times \mathbb{R})$:

$$\mathcal{L}f(x, y) = -y \partial_x f + \lambda(U'(x) - y) \partial_y f + \frac{\sigma^2}{2} \partial_x^2 f.$$

- ▶ Remarque : Cela ressemble aux équations Fokker Planck cinétique ! (Le $\overline{dB_t}$ est placé sur la coordonnée position et non vitesse).
- ▶ L'équation d'évolution est dégénérée sur une coordonnée. Peut-on obtenir malgré tout
 - ▶ Des propriétés d'existence et de régularité des densités de (X_t, Y_t) ?
 - ▶ Des vitesses de convergence à l'équilibre (s'il existe) ?
 - ▶ Des asymptotiques lorsque $\sigma \rightarrow 0$?

Hypoellipticité

- ▶ On se place dans le cas où le potentiel U est C^∞ et ne possède pas de plateau. On définit

$$E_U = \left\{ x \in \mathbb{R} \quad \exists k \geq 1 \quad U^{(k)}(x) \neq 0 \right\}.$$

- ▶ On peut alors déduire du théorème de Hormander que

Proposition (S.G., F. Panloup)

Si $\mathbb{R} \setminus E_U$ n'a que des points isolés, alors (X_t, Y_t) a une densité de probabilité par rapport à la mesure de Lebesgue. On notera p_t la densité de probabilités. On a en plus pour tout $t > 0$: $p_t \in C^\infty(\mathbb{R} \times \mathbb{R})$ et $\forall (x, y) \in \mathbb{R} \times \mathbb{R} \quad p_t(x, y) > 0$.

- ▶ Le processus est qualifié d'« irréductible ».

Existence d'un régime stationnaire

- ▶ On souhaite prouver l'existence d'une densité limite lorsque $t \rightarrow \infty$.
- ▶ Il faut contourner la non coercivité de \mathcal{L} .
- ▶ Hypothèses de rappel fort

$$\exists a \in]0; 1] \quad (\mathbf{H}_a) \begin{cases} \frac{|U'|^2}{U} + D^2U + D^3U = O(U(x) \vee |x|^2)^a \\ (U(x) \vee |x|^2)^a = O(xU'(x)) \end{cases}$$

(\mathbf{H}_a) est vraie pour $U \sim |x|^p, p > 0$ mais plus pour $U(x) \sim \ln(|x| + 1)^\beta$.

- ▶ Utilisation d'une fonction de Lyapunov « légèrement » modifiée pour obtenir l'existence d'un régime stationnaire :

$$\mathcal{V}_m(x, y) = U(x) + \frac{y^2}{2} + m \left(\frac{x^2}{2} - \frac{xy}{\lambda} \right).$$

Sous (\mathbf{H}_a) , on montre l'existence de $(\alpha, \beta) \in \mathbb{R}_+^2$ et $(\rho, \rho') \in \mathbb{R}_+^2$ telles que

$$\mathcal{L}\mathcal{V}_m^\rho \leq \beta - \alpha\mathcal{V}_m^{\rho'}$$

Existence d'un régime stationnaire (suite)

- ▶ Grâce à cette fonction de Lyapunov, on peut déduire le théorème :

Theorem (S.G., F. Panloup)

Sous l'hypothèse (\mathbf{H}_a) , si le processus est hypoelliptique, alors il existe une unique distribution invariante p_∞ solution de

$$y\partial_x p_\infty + \frac{\sigma^2}{2}\partial_x^2 p_\infty + \lambda(y - U'(x))\partial_y p_\infty + p_\infty = 0.$$

- ▶ On peut obtenir des vitesses de convergence "faibles" au sens de la variation totale.

Theorem (S.G., F. Panloup)

Pour tout $p \geq 1$, et sous l'hypothèse (\mathbf{H}_a) , on a

$$\sup_{\{f, |f| \leq \mathcal{V}_m^p\}} |P_t(X_0, Y_0, f) - p_\infty(f)| \leq C_{a,p} \{\mathcal{V}_m(X_0, Y_0)\}^p \begin{cases} \exp(-\delta_{p,r_\infty} t) & \text{si } a = 1 \\ t^{-\frac{p+a-1}{1-a}} & \text{si } a \in (0, 1). \end{cases}$$

Existence d'un régime stationnaire (suite)

- ▶ Vitesse de convergence exponentielle lorsque U est essentiellement surquadratique.
- ▶ Vitesse géométrique dans le cas où U a une croissance plus faible.
- ▶ Pas de borne inférieure.
- ▶ Pas de résultat en norme 2 ou en entropie.
- ▶ La structure de la fonction de Lyapunov ressemble beaucoup à ce qui est fait pour F-P Cinétique avec une norme modifiée par un produit scalaire entre vitesse et position. Ici, on utilise le produit scalaire xy dans la fonction $\mathcal{V}_m \dots$

Identification de la limite - Lien avec FP cinétique

- ▶ En langage EDP, on vient de montrer que la densité p_t de probabilités \mathcal{C}^∞ sur $\mathbb{R} \times \mathbb{R}$ vérifiant

$$\partial_t p_t = \mathcal{L}^*(p_t) := y \partial_x p_t + \frac{\sigma^2}{2} \partial_x^2 p_t + \lambda (y - U'(x)) \partial_y p_t + p_t$$

converge en variation totale, a une vitesse soit exponentielle, soit géométrique vers p_∞ unique solution de $\mathcal{L}^*(\mu) = 0$.

- ▶ Dans les cas Gaussien (et seulement dans ces cas) où $U(x) = x^2/2$, on connaît p_∞ qui est une densité Gaussienne centrée de variance

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

- ▶ Dans le cas gaussien $U(x) = ax^2$ avec $\sigma = 1$ et après changement de variable linéaire, le générateur \mathcal{L}_a obtenu est celui de F-P cinétique.

$$\mathcal{L}_a = y \partial_x - ax \partial_y + \partial_y^2 - y \partial_y$$

Vitesse - Lien avec FP cinétique

On se place dans le cas où $\sigma = 1$. On définit

$$\alpha(P) := \min_{f \in L^2(\mu_a)} - \limsup_{t \rightarrow \infty} \frac{1}{t} \log[P_t(f) - \mu(f)].$$

On définit P_{FPa} le semi-groupe associé à \mathcal{L}_a et Q le semi-groupe (symétrique) associé à

$$\tilde{\mathcal{L}}_a = \partial_x^2 - ax\partial_x + \partial_y^2 - y\partial_y$$

on a alors le résultat

Theorem (S.G., L. Miclo et beaucoup d'autres...)

Dans le cas où $a > 1/4$, $\alpha(P_{FPa}) = 1/2$ et $\alpha(Q) = \min(1, a)$. En particulier, si $a \in [1/4; 1/2]$, alors

$$\alpha(P_{FPa}) > \alpha(Q).$$

On conjecture que ce résultat reste vrai pour $a \in [0; 1/4]$.

Diffusion à petit paramètre

On cherche à identifier le comportement de la diffusion lorsque $\sigma \mapsto 0$. Dans le cas « standard » de la diffusion de Kolmogorov,

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

La distribution stationnaire est le champ de Gibbs

$$\mu_\sigma \propto e^{-U/(2\sigma^2)},$$

et on a par la méthode de Laplace que

$$\mu_\sigma \xrightarrow{\sigma \rightarrow 0} \text{Conv} \{ \delta_{x_i} \mid x_i \in \arg \min U \}.$$

En fait, on montre plus précisément que (μ_σ) satisfait un principe de Grandes Déviations en effectuant un changement de variables $\mu_\sigma = e^{-V/\sigma^2}$. Approche par

- ▶ Solutions de Viscosité + inégalités fonctionnelles Perthame(1990), Barles (1994)
- ▶ Approche probabiliste Wentzell & Freidlin (1969) (description moins précise).

Diffusion moyennée à petit paramètre

On cherche à identifier le comportement de la diffusion lorsque $\sigma \mapsto 0$.

$$\mathcal{L}_\sigma = -y\partial_x + (U'(x) - y)\partial_y + \sigma^2\partial_x^2.$$

On définit $W_\sigma = -\sigma^2 \log \mu_\sigma$, W_σ est solution de

$$H_\sigma(W, DW_\sigma, D^2W_\sigma) = 0,$$

avec

$$H_\sigma(V, DV, D^2V) = -y\partial_x V + (U'(x) - y)\partial_y V + \frac{\{\partial_x V\}^2}{2} + \sigma^2 \left(1 + 1/2\partial_x^2 V\right).$$

Theorem (S.G., F. Panloup, C. Pellegrini)

La suite de mesures (μ_σ) satisfait un principe de Grandes Déviations :

$$\lim_{\sigma \rightarrow 0} \frac{1}{\sigma^2} \log[\mu_\sigma(x, y)] = -W(x, y)$$

où W est solution de l'équation d'Hamilton Jacobi

$$H(W, DW) := -y\partial_x W + (U'(x) - y)\partial_y W + \frac{\{\partial_x W\}^2}{2} = 0$$

Diffusion moyennée à petit paramètre

En notant $b(x, y)$ le vecteur de drift, et en appliquant le principe de programmation dynamique, on peut démontrer que W satisfait l'équation fonctionnelle

$$\forall t > 0 \quad W(x, y) = \inf_{\begin{cases} \dot{z} = -b(z) - \begin{pmatrix} \dot{\varphi} \\ 0 \end{pmatrix} \\ z(0) = (x, y) \end{cases}} \left[\frac{1}{2} \int_0^t \dot{\varphi}^2 + W(z(t)) \right].$$

Si on note z_1, \dots, z_m les points d'équilibres stables du champ de vecteur, on peut même avoir le résultat suivant

Theorem (S.G., F. Panloup, C. Pellegrini)

W vérifie l'équation

$$W(x, y) = \min_{1 \leq i \leq m} \inf_{\begin{cases} \dot{z} = -b(z) - \begin{pmatrix} \dot{\varphi} \\ 0 \end{pmatrix} \\ z(0) = (x, y) \end{cases}} \left[\frac{1}{2} \int_0^\infty \dot{\varphi}^2 + W(z_i) \right]. \quad (3)$$

Diffusion moyennée à petit paramètre

- ▶ Question naturelle : L'équation (3) suffit-elle à caractériser de manière unique le W , fonctionnelle d'action du PGD ?
- ▶ Réponse : en général non. Les propriétés de comparaison pour les solutions de viscosités ne s'appliquent pas ici.
- ▶ On a par contre unicité du W dès lors que l'on sait que (3) est satisfaite avec en plus une connaissance des $W(z_i)$.

On peut obtenir une expression des $W(z_i)$ par le biais des quantités

$$V(z_i, z_j) = \inf_{\begin{cases} \dot{z} = b(z) - \begin{pmatrix} \dot{\phi} \\ 0 \end{pmatrix} \\ z(0) = z_i \quad z(\infty) = z_j \end{cases}} \left[\frac{1}{2} \int_0^t \dot{\phi}^2 \right],$$

et

$$\bar{W}(z_i) = \min_{g \in G(i)} \sum_{(z_k \rightarrow z_l) \in g} V(z_k, z_l).$$

Ici, $G(i)$ désigne l'ensemble des i -graphes, graphes orientés sur $\{z_1, \dots, z_m\}$ avec une seule flèche sortante de chaque z_r ($r \neq i$), et tels que tout état z_r est relié par un chemin à z_i .

Retour au problème de contrôle :

Theorem (A. Freidlin, M. Wentzell)

$$W(z_i) = \bar{W}(z_i) - \min_l \bar{W}(z_l)$$

- ▶ Pour identifier le point critique qui annule W , il faut résoudre le problème du minimiseur de \bar{W} .
- ▶ Ce minimiseur est le point sur lequel se concentre la masse de μ_σ .
- ▶ Certaines simulations numériques incitent à penser que sous des hypothèses de régularité de U , et d'espacement des points critiques, un minimum global x^* de U coïncide avec le minimiseur $(x^*, 0)$ de \bar{W} .

Plan de l'exposé

Introduction

Étude de l'ODE (SGM')

Modèle diffusif moyenné (DM)

Conclusion

Conclusion

- ▶ Bonne connaissance du système déterministe (GM).
 - ▶ Si γ tend trop vite vers 0, il n'y a pas de convergence.
 - ▶ Convergence dans les cas convexes dès que $k(t) = h(t) = e^t$.
- ▶ 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

$$dS_t = -\nabla U(S_t)dt + \sigma(S_t)dB_t.$$

- ▶ Si $k(t) = t^\alpha$, explosion du processus (non exposé aujourd'hui).
 - ▶ Vitesses exponentielles ou polynomiales dans les cas stables.
 - ▶ Vitesses L^2 dans le cas général sous (H_a) ?
- ▶ Diffusion à petits paramètres
 - ▶ Existence d'un PGD
 - ▶ Problème de contrôle à terminer...
- ▶ Élargissements :
 - ▶ Coupler $\sigma_t \rightarrow 0$ avec la vitesse L^2 , application dédiée à l'optimisation.
 - ▶ Généraliser à des mémoires différentes :

$$dX_t = -\frac{1}{Z_t} \left[\int_{a(t)}^{b(t)} \nabla U(X_s) ds \right] dt + \sigma(X_t)dB_t.$$

- ▶ Borne inférieures de vitesses inconnues.
 - ▶ D'un point de vue de la modélisation, étudier des systèmes à mémoire ayant une écriture comparable.

Merci de votre attention !