Équations Différentielles Stochastiques Rétrogrades

Philippe Briand

Mars 2001

Introduction

L'objectif de ce cours est de présenter la théorie des équations différentielles stochastiques rétrogrades – EDSR dans toute la suite du cours – et d'en donner des applications dans deux domaines différents : les mathématiques financières d'une part et les équations aux dérivées partielles – EDP en abrégé – d'autre part.

De nombreux mathématiciens ont contribué et contribuent toujours à la théorie des EDSR ; il m'est impossible de tous les citer. Néanmoins, signalons les travaux de E. Pardoux et S. Peng [PP90, PP92, PP94] et l'article de N. El Karoui, S. Peng et M.-C. Quenez [EKPQ97].

Le cours est divisé en deux parties : dans un premier temps, on se consacre à l'étude générale des EDSR puis dans la seconde partie, on étudiera les liens entre les EDSR et les EDP et l'on introduira la notion de solution de viscosité d'une EDP.

Résoudre une EDSR, c'est trouver un couple de processus adaptés par rapport à la filtration du mouvement brownien $(W_t)_{0\leq t\leq T}$, $(Y_t, Z_t)_{t\in[0,T]}$ vérifiant l'équation différentielle stochastique

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t dW_t, \quad 0 \leq t \leq T,$$

avec la condition finale (c'est pour cela que l'on dit rétrograde) $Y_T = \xi$ où ξ est une variable aléatoire de carré intégrable. Comme les EDS, ces équations doivent être comprise au sens intégral i.e.

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T.$$

Les EDSR ont été introduites en 1973 par J.-M. Bismut [Bis73] dans le cas où f est linéaire par rapport aux variables Y et Z. Il a fallu attendre le début des années 90 et le travail de E. Pardoux et S. Peng [PP90] pour avoir le premier résultat d'existence et d'unicité dans le cas où f n'est pas linéaire.

Depuis de nombreux travaux ont été effectués ; la théorie n'a cessé de se développer en raison de ses relations étroites avec les mathématiques financières et les EDP. Donnons un exemple emprunté à chacun des deux thèmes précédents.

En finance, une question importante est de déterminer le prix d'une option – un produit financier. Prenons le cas le plus simple, celui du modèle de Black–Scholes et d'un « call européen ». Le prix de ce produit financier, $(V_t)_{0\leq t\leq T}$ satisfait l'équation :

$$dV_t = (rV_t + \theta Z_t)dt + Z_t dW_t,$$

où r est le taux d'intéret à court terme et θ est le « risk premium », avec la condition finale $V_T = (S_T - K)^+$ où S_t est le prix de l'action sous-jacente et K une constante, un prix fixé à l'avance. Nous voyons que c'est une EDSR linéaire dans ce modèle simple mais qui peut être non-linéaire dans des modèles financiers plus compliqués.

Venons-en au second exemple. Considérons l'EDP suivante

$$\partial_t u(t,x) + \frac{1}{2}\partial^2_{x,x} u(t,x) + f(u(t,x)) = 0, \qquad u(T,x) = g(x).$$

Supposons que cette équation possède une solution régulière, u. Appliquons la formule d'Itô à $u(s, W_s)$; on obtient

$$\begin{aligned} du(s, W_s) &= \{\partial_s u(s, W_s) + \frac{1}{2}\partial^2_{x,x} u(s, W_s)\} ds + \partial_x u(s, W_s) dW_s \\ &= -f(u(s, W_s)) ds + \partial_x u(s, W_s) dW_s. \end{aligned}$$

Nous obtenons encore une EDSR – qui est non-linéaire si f l'est – en posant $Y_s = u(s, W_s)$ et $Z_s = \partial_x u(s, W_s)$ puisque

$$-dY_s = f(Y_s) ds - Z_s dW_s, \quad \text{avec}, \quad Y_T = g(W_T).$$

Le cours commence par des rappels et compléments sur les EDS, puis vient l'étude des EDSR et enfin les applications aux EDP.

Références

[Bis73] J.-M. Bismut, *Conjugate convex functions in optimal stochastic control*, J. Math. Anal. Appl. **44** (1973), 384–404.

[EKPQ97] N. El Karoui, S. Peng, and M.-C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance **7** (1997), no. 1, 1–71.

[PP90] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett. **14** (1990), no. 1, 55–61.

[PP92] ———, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, Stochastic partial differential equations and their applications (Charlotte, NC, 1991) (B. L. Rozovskii and R. B. Sowers, eds.), Lecture Notes in Control and Inform. Sci., vol. 176, Springer, Berlin, 1992, pp. 200–217.

[PP94] ———, *Backward doubly stochastic differential equations and systems of quasilinear SPDE's*, Probab. Theory Related Fields **98** (1994), no. 2, 209–227.

Chapitre 1

Rappels de calcul stochastique

Le but de ce chapitre est de présenter brièvement les résultats de calcul stochastique dont nous aurons besoin dans la suite du cours. Il ne s'agit nullement de développer la théorie générale pour laquelle les ouvrages [KS91] et [RY91] sont des références remarquables. La référence [LL97] est très pédagogique ; le second paragraphe en est fortement inspiré.

1. Définitions

Ici $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace de probabilité complet.

1.1. Généralités.

Définition 1. Soit T un ensemble. On appelle processus stochastique indexé par T et à valeurs dans $\mathbb{R}^d$ une famille $(X_t)_{t\in T}$ d'applications mesurables de $(\Omega, \mathcal{F})$ dans $\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$; pour tout $t \in T$, X_t est une variable aléatoire.

Remarque. Dans ce cours, nous aurons $T = \mathbb{N}$ ce qui correspond aux processus à temps discret, $T = \mathbb{R}_+$ ou $T = [0, a]$ pour les processus à temps continu.

Pour simplifier les énoncés qui suivent sont donnés avec $T = \mathbb{R}_+$; pour alléger l'écriture, nous noterons un processus X plutôt que $(X_t)_{t\geq 0}$.

Définition 2. Un processus X est mesurable si l'application $(t, \omega) \longmapsto X_t(\omega)$ de $\mathbb{R}_+ \times \Omega$ dans $\mathbb{R}^d$ est mesurable par rapport aux tribus $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ et $\mathcal{B}(\mathbb{R}^d)$.

On peut noter qu'un processus stochastique peut être vu comme une fonction aléatoire : à chaque ω, on associe la fonction $t \longmapsto X_t(\omega)$ qui est appelée trajectoire (sample path en anglais).

Définition 3. Soient X et Y deux processus. X est une modification de Y si, pour tout $t \geq 0$, les v.a. X_t et Y_t sont égales $\mathbb{P}$–p.s. : $\forall t \geq 0$, $\mathbb{P}(X_t = Y_t) = 1$.

X et Y sont indistinguables si, $\mathbb{P}$–p.s., les trajectoires de X et de Y sont les mêmes c'est à dire $\mathbb{P}\left(X_t = Y_t,\ \forall t \geq 0\right) = 1$.

La notion d'indistinguabilité est plus forte que la notion de modification. Notons que si X est une modification de Y et si X et Y sont à trajectoires continues à droite (ou à gauche) alors X et Y sont indistinguables.

Comme dans le cas discret, les tribus jouent un rôle très important dans l'étude des processus stochastiques car elles représentent l'information disponible et permettent de traduire les notions de passé, présent et futur.

Nous travaillerons avec une filtration $\{\mathcal{F}_t\}_{t\geq 0}$ de $(\Omega, \mathcal{F})$ c'est à dire une famille croissante de sous-tribus de $\mathcal{F}$ i.e. pour $s \leq t$, $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$. On définit alors $\mathcal{F}_\infty = \sigma\{\cup_t \mathcal{F}_t\}$ ainsi que, pour tout t, $\mathcal{F}_{t^+} = \cap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}$.

Définition 4. On dit qu'une filtration $\{\mathcal{F}_t\}_{t\geq 0}$ est continue à droite si $\mathcal{F}_t = \mathcal{F}_{t^+}$ pour tout t.

On dit qu'elle vérifie les conditions habituelles si elle est continue à droite et si $\mathcal{F}_0$ contient tous les ensembles $\mathbb{P}$–négligeables de $\mathcal{F}$, noté $\mathcal{N}$ dans la suite.

Si l'on se donne un processus X, on introduit la filtration $\mathcal{G}_t = \sigma\{X_s;\ s \leq t\}$. Cette filtration s'appelle la filtration naturelle de X. Mais $\mathcal{G}_0$ ne contient pas nécessairement $\mathcal{N}$. C'est pour cela que l'on introduit souvent la *filtration naturelle augmentée* de X définie par $\mathcal{F}_t^X = \sigma\{\mathcal{N} \cup \mathcal{G}_t\}$. Lorsque nous parlerons de filtration naturelle il s'agira toujours de filtration naturelle augmentée.

Définition 5. Un processus X est adapté par rapport à la filtration $\{\mathcal{F}_t\}_{t\geq 0}$ si, pour tout t, X_t est $\mathcal{F}_t$–mesurable.

Il est inutile de dire qu'un processus est toujours adapté par rapport à sa filtration naturelle.

Remarque. Si $\mathcal{N} \subset \mathcal{F}_0$, et si X est adapté par rapport à $\{\mathcal{F}_t\}_{t\geq 0}$ alors toute modification de X est encore adaptée.

Définition 6. Un processus X est progressivement mesurable par rapport à $\{\mathcal{F}_t\}_{t\geq 0}$ si, pour tout $t \geq 0$, l'application $(s,\omega) \longmapsto X_s(\omega)$ de $[0,t]\times\Omega$ dans $\mathbb{R}^d$ est mesurable par rapport à $\mathcal{B}([0,t])\otimes\mathcal{F}_t$ et $\mathcal{B}(\mathbb{R}^d)$.

Un processus progressivement mesurable est mesurable et adapté. On peut également noter que si X est un processus mesurable et adapté alors il possède une modification progressivement mesurable. Rappelons également le résultat suivant :

Proposition 7. *Si X est un processus stochastique dont les trajectoires sont continues à droite (ou continues à gauche) alors X est mesurable et X est progressivement mesurable s'il est de plus adapté.*

Finissons ces généralités par la notion de temps d'arrêt.

Définition 8. Soit τ une variable aléatoire à valeurs dans $\overline{\mathbb{R}}_+$. τ est un $\{\mathcal{F}_t\}_{t\geq 0}$–temps d'arrêt si, pour tout t, $\{\tau \leq t\} \in \mathcal{F}_t$.

Si τ est un temps d'arrêt, on appelle tribu des évènements antérieurs à τ la tribu définie par

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty,\ A \cap \{\tau \leq t\} \in \mathcal{F}_t,\ \forall t\}.$$

Proposition 9. *Soit τ un temps d'arrêt. Si X est progressivement mesurable, le processus arrêté X^τ – défini par $X_t^\tau = X_{\tau\wedge t}$ – est progressivement mesurable.*

1.2. Mouvement brownien.

Définition 10. On appelle *mouvement brownien standard* un processus stochastique W à valeurs réelles tel que :

1. $\mathbb{P}$–p.s. $t \longmapsto W_t(\omega)$ est continue ;
2. pour $0 \leq s < t$, $W_t - W_s$ est indépendant de la tribu $\sigma\{W_u,\ u \leq s\}$ et de loi gaussienne centrée de variance $t - s$;
3. $W_0 = 0$ $\mathbb{P}$–p.s.

Pour tout $t > 0$, la variable aléatoire W_t suit la loi gaussienne centrée de variance t donc de densité $(2\pi t)^{-1/2}\exp\{-x^2/(2t)\}$. On dit qu'un mouvement brownien (MB dans la suite) part d'un point x si $W_0 = x$.

Remarque. On dit que W est un $\{\mathcal{F}_t\}_{t\geq 0}$–MB si W est un processus continu, adapté à la filtration $\{\mathcal{F}_t\}_{t\geq 0}$, vérifiant :

$$\forall u \in \mathbb{R}, \quad \forall 0 \leq s \leq t, \qquad \mathbb{E}\left(e^{iu(W_t - W_s)} \mid \mathcal{F}_s\right) = \exp\left\{-u^2(t-s)/2\right\}.$$

Proposition 11. *Soit W un MB standard.*

1. *pour tout $s > 0$, $\{W_{t+s} - W_s\}_{t\geq 0}$ est un MB indépendant de $\sigma\{W_u,\, u \leq s\}$;*
2. *$-W$ est aussi un MB ;*
3. *pour tout $c > 0$, $\{cW_{t/c^2}\}_{t\geq 0}$ est un MB ;*
4. *le processus défini par $X_0 = 0$ et $X_t = tW_{1/t}$ est un MB.*

Rappelons qu'une fonction f à valeurs réelles et définie sur $\mathbb{R}_+$ est localement hölderienne d'ordre α si, pour tout $a > 0$, il existe une constante C telle que :

$$\forall (x, y) \in [0, a]^2, \qquad |f(x) - f(y)| \leq C\,|x - y|^\alpha.$$

Théorème 12 (Propriétés des trajectoires)**.** *Si W est un MB, alors presque sûrement, on a :*

1. *$t \longmapsto W_t(\omega)$ n'est à variation finie sur aucun intervalle ;*
2. *$t \longmapsto W_t(\omega)$ est localement hölderienne d'ordre α pour tout $\alpha < 1/2$.*
3. *$t \longmapsto W_t(\omega)$ n'est dérivable en aucun point (ni localement hölderienne d'ordre $\alpha \geq 1/2$).*

Définition 13. On appelle MB standard à valeurs dans $\mathbb{R}^d$, un vecteur $W = \left(W^1, \ldots, W^d\right)$ où les W^i sont des MB réels indépendants.

Proposition 14. *Soit W un MB. La filtration $\left\{\mathcal{F}_t^W\right\}_{t\geq 0}$ vérifie les conditions habituelles et W est un $\left\{\mathcal{F}_t^W\right\}_{t\geq 0}$–MB.*

1.3. Martingales.

Définition 15. Un processus X à valeurs réelles est une surmartingale par rapport à la filtration $\{\mathcal{F}_t\}_{t\geq 0}$ si :

1. pour tout $t \geq 0$, X_t est $\mathcal{F}_t$–mesurable ;
2. pour tout $t \geq 0$, X_t^- est intégrable ;
3. pour $0 \leq s \leq t$, $\mathbb{E}\left(X_t \mid \mathcal{F}_s\right) \leq X_s$.

X est une sousmartingale lorsque $-X$ est une surmartingale ; X est une martingale si X est à la fois une surmartingale et une sous-martingale.

Remarque. Si W est un MB, alors $\left\{W_t^2 - t\right\}_{t\geq 0}$ et $\left\{\exp\left(\sigma W_t - \sigma^2 t/2\right)\right\}_{t\geq 0}$ sont des martingales.

Théorème 16 (Inégalités maximales)**.** *Soit X une martingale (ou une sous-martingale positive) continue à droite. Alors,*

1. *$\forall p \geq 1$, $\forall a > 0$, $\quad a^p\, \mathbb{P}\left(\sup_t |X_t| \geq a\right) \leq \sup_t \mathbb{E}\left[|X_t|^p\right]$;*
2. *$\forall p > 1$, $\quad \mathbb{E}\left[\sup_t |X_t|^p\right] \leq q^p \sup_t \mathbb{E}\left[|X_t|^p\right]$ où $q = p(p-1)^{-1}$.*

Théorème 17 (Théorème d'arrêt)**.** *Si X est une martingale et si σ et τ sont deux temps d'arrêt bornés tels que $\sigma \leq \tau$, alors, $\mathbb{E}(X_\tau | \mathcal{F}_\sigma) = X_\sigma$ $\mathbb{P}$–p.s.*

Rappelons aussi qu'un processus X adapté et intégrable est une martingale si et seulement si, pour tout temps d'arrêt borné τ, $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.

Pour le dernier résultat nous supposons que la filtration $\{\mathcal{F}_t\}_{t\geq 0}$ vérifie les conditions habituelles.

Théorème 18. *Soit X une $\{\mathcal{F}_t\}_{t\geq 0}$–martingale. X possède une modification càdlàg i.e. dont les trajectoires sont continues à droite et possèdent des limites à gauche.*

Indroduisons à présent une classe plus vaste de processus stochastiques : les martingales locales.

Définition 19. Soit X un processus $\{\mathcal{F}_t\}_{t\geq 0}$–adapté, à trajectoires continues à droite. On dit que X est une martingale locale s'il existe une suite croissante de temps d'arrêt $\{\tau_n\}_{n\geq 1}$ telle que $\lim_{n\to\infty} \tau_n = +\infty$ $\mathbb{P}$–p.s et, pour tout n, $X^{\tau_n}\mathbf{1}_{\tau_n>0}$ est une martingale.

Théorème 20. *Soit X une martingale locale continue. Il existe un unique processus croissant et continu, $\langle X, X\rangle$, nul en 0, tel que $X^2 - \langle X, X\rangle$ soit une martingale locale.*

Remarque. Si W est un MB, on a $\langle W, W\rangle_t = t$ car $\left\{W_t^2 - t\right\}_{t\geq 0}$ est une martingale.

Proposition 21. *Soit X une martingale locale continue. Il y a équivalence entre :*

1. *$X_0 \in \mathrm{L}^2$ et $\mathbb{E}[\langle X, X\rangle_\infty] < \infty$;*
2. *X est une martingale bornée dans L^2.*

Rappelons pour finir ce paragraphe les inégalités de Burkholder–Davis–Gundy connues aussi sous le nom d'inégalités BDG.

Théorème 22 (Inégalités BDG)**.** *Soit $p > 0$ un réel. Il existe deux constantes c_p et C_p telles que, pour toute martingale locale continue X, nulle en zéro,*

$$c_p\,\mathbb{E}\left[\langle X, X\rangle_\infty^{p/2}\right] \leq \mathbb{E}\left[\sup_{t\geq 0}|X_t|^p\right] \leq C_p\,\mathbb{E}\left[\langle X, X\rangle_\infty^{p/2}\right].$$

Remarque. En particulier, si $T > 0$,

$$c_p\,\mathbb{E}\left[\langle X, X\rangle_T^{p/2}\right] \leq \mathbb{E}\left[\sup_{0\leq t\leq T}|X_t|^p\right] \leq C_p\,\mathbb{E}\left[\langle X, X\rangle_T^{p/2}\right].$$

2. Calcul d'Itô

Dans tout ce paragraphe, on se donne $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité complet, $\{\mathcal{F}_t\}_{t\geq 0}$ une filtration qui vérifie les conditions habituelles et W un $\{\mathcal{F}_t\}_{t\geq 0}$–MB (on peut prendre $\mathcal{F}_t = \mathcal{F}_t^W$). Les martingales seront toujours càdlàg.

2.1. Intégrale stochastique. L'objectif de ce paragraphe est de définir $\int_0^t H_s\,dW_s$. Ceci n'est pas évident car comme nous l'avons rappelé précédemment les trajectoires du MB ne sont pas à variation finie et donc l'intégrale précédente n'est en aucun cas une intégrale de Lebesgue-Stieljes.

Dans toute la suite, on fixe un réel T strictement positif. Les processus sont définis pour $t \in [0, T]$; on notera X pour $(X_t)_{t\in[0,T]}$.

Définition 23. On appelle processus élémentaire $H = (H_t)_{0\leq t\leq T}$ un processus de la forme :

$$H_t = \phi_0\mathbf{1}_0(t) + \sum_{i=1}^{p}\phi_i\mathbf{1}_{]t_{i-1},t_i]}(t),$$

où $0 = t_0 < t_1 < \ldots < t_p = T$, ϕ_0 est une variable aléatoire $\mathcal{F}_0$–mesurable bornée et, pour $i = 1, \ldots, p$, ϕ_i est une variable aléatoire $\mathcal{F}_{t_{i-1}}$–mesurable et bornée.

Pour un tel processus, on peut définir l'intégrale stochastique par rapport à W comme étant le processus continu $\{I(H)_t\}_{0\leq t\leq T}$ défini par :

$$I(H)_t = \sum_{i=1}^{p} \phi_i (W_{t_i\wedge t} - W_{t_{i-1}\wedge t}),$$

soit encore, si $t \in]t_k, t_{k+1}]$,

$$I(H)_t = \sum_{i=1}^{k} \phi_i (W_{t_i} - W_{t_{i-1}}) + \phi_{k+1}(W_t - W_{t_k}).$$

On note $\int_0^t H_s\, dW_s$ pour $I(H)_t$. On obtient alors directement à l'aide de cette définition le résultat suivant :

Proposition 24. *Si H est un processus élémentaire, alors $\left(\int_0^t H_s\, dW_s\right)_{0\leq t\leq T}$ est une $\{\mathcal{F}_t\}_{t\geq 0}$–martingale continue telle que*

$$\forall t \in [0,T], \qquad \mathbb{E}\left[\left|\int_0^t H_s\, dW_s\right|^2\right] = \mathbb{E}\left[\int_0^t H_s^2\, ds\right].$$

On veut à présent définir l'intégrale stochastique pour une classe plus vaste de processus H. Pour la première extension, on utilise la densité des processus élémentaires dans l'espace vectoriel $\mathcal{M}^2$ suivant :

$$\mathcal{M}^2 = \left\{(H_t)_{0\leq t\leq T},\ \text{progressivement mesurable, } \mathbb{E}\left[\int_0^T H_s^2\, ds\right] < \infty\right\}.$$

On désigne par H^2 l'espace vectoriel des martingales bornées dans L^2 ; le sous-espace de H^2 formé par les martingales qui sont continues est noté H^2_c. On munit H^2 de la norme définie par $\|M\|_{\mathrm{H}^2} = \mathbb{E}[|M_T|^2]^{1/2}$ qui en fait un espace de Hilbert. L'inégalité de Doob montre que cette norme est équivalente à la norme $\mathbb{E}[\sup_t |M_t|^2]^{1/2}$; par suite, H^2_c est un sous-espace fermé. $\mathbb{H}^2$ et $\mathbb{H}^2_c$ désignent les sous-espaces de H^2 et H^2_c constitués des martingales nulles en 0 ; ces deux sous-espaces sont fermés.

On obtient alors le résultat suivant :

Théorème 25. *Il existe une unique application linéaire J de $\mathcal{M}^2$ dans $\mathbb{H}^2_c$ telle que :*

1. *si H est un processus élémentaire, alors $I(H)$ et $J(H)$ sont indistinguables ;*
2. *pour tout t, $\mathbb{E}\left[J(H)_t^2\right] = \mathbb{E}\left[\int_0^t H_s^2\, ds\right]$.*

L'unicité signifie que si J et J' sont deux prolongements vérifiant les propriétés précédentes alors $J(H)$ et $J'(H)$ sont indistinguables.

On note toujours $\int_0^t H_s\, dW_s$ pour $J(H)_t$.

Remarque. Notons M^2 l'ensemble des classes d'équivalence de $\mathcal{M}^2$. M^2 est un espace de Hilbert. L'intégrale stochastique est alors une isométrie de M^2 dans $\mathbb{H}^2_c$.

On obtient les propriétés suivantes :

Proposition 26. *Soit $H \in \mathcal{M}^2$. On a*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t H_s\, dW_s\right|^2\right] \leq 4\mathbb{E}\left[\int_0^T H_s^2\, ds\right],$$

et si τ est un temps d'arrêt,

$$\int_0^\tau H_s\, dW_s = \int_0^T \mathbf{1}_{s\leq\tau} H_s\, dW_s, \quad \mathbb{P}\text{–}p.s.$$

La dernière extension de l'intégrale stochastique dont nous aurons besoin consiste à relaxer l'hypothèse d'intégrabilité portant sur H. On introduit pour cela

$$\mathcal{M}^2_{\text{loc}} = \left\{ (H_t)_{0\le t\le T},\ \text{progressivement mesurable},\ \int_0^T H_s^2\,ds < \infty\ \mathbb{P}\text{–p.s.} \right\}.$$

et on a le résultat suivant :

Proposition 27. *Il existe une unique application linéaire J' de $\mathcal{M}^2_{\text{loc}}$ dans l'ensemble des martingales locales continues telle que :*

1. *si H est un processus élémentaire alors $J'(H)$ et $I(H)$ sont indistinguables ;*
2. *si $(H_n)_n$ est une suite de processus de $\mathcal{M}^2_{\text{loc}}$ telle que $\int_0^T {H_s^n}^2\,ds$ tend vers 0 en probabilité alors $\sup_{0\le t\le T} |J'(H^n)_t|$ tend vers 0 en probabilité.*

On note encore $\int_0^t H_s\,dW_s$ pour $J'(H)_t$.

Remarque. Attention, lorsque $H \in \mathcal{M}^2_{\text{loc}}$, $\left(\int_0^t H_s\,dW_s\right)_{0\le t\le T}$ est seulement une martingale locale et pas nécessairement une martingale.

Proposition 28. *Pour $H \in \mathcal{M}^2_{\text{loc}}$, $\langle \int_0^\cdot H_s\,dW_s\rangle_t = \int_0^t H_s^2\,ds$.*

2.2. Processus d'Itô. Nous introduisons à présent une classe de processus qui sera très utile dans la suite.

Définition 29. On appelle processus d'Itô un processus X à valeurs réelles tel que :

$$\mathbb{P}\text{–p.s.} \qquad \forall 0 \le t \le T, \qquad X_t = X_0 + \int_0^t K_s\,ds + \int_0^t H_s\,dW_s,$$

où X_0 est $\mathcal{F}_0$–mesurable, K et H sont deux processus progressivement mesurables vérifiant les conditions, $\mathbb{P}$–p.s. :

$$\int_0^T |K_s|\,ds < \infty \quad \text{et} \quad \int_0^T |H_s|^2\,ds < \infty.$$

On peut montrer que si un processus d'Itô est une martingale locale continue alors $K_t = 0$ $m \otimes \mathbb{P}$–p.p. On en déduit alors que la décomposition d'un processus d'Itô est unique au sens où si

$$X_t = X_0 + \int_0^t K_s\,ds + \int_0^t H_s\,dW_s = X_0' + \int_0^t K_s'\,ds + \int_0^t H_s'\,dW_s$$

alors $X_0 = X_0'$ $\mathbb{P}$–p.s. et $H_t' = H_t$, $K_t = K_t'$ $m \otimes \mathbb{P}$–p.p.

Si X et Y sont deux processus d'Itô,

$$X_t = X_0 + \int_0^t K_s\,ds + \int_0^t H_s\,dW_s \quad \text{et} \quad Y_t = X_0' + \int_0^t K_s'\,ds + \int_0^t H_s'\,dW_s$$

on pose $\langle X, Y\rangle_t = \int_0^t H_s H_s'\,ds$ et $dX_t = K_t\,dt + H_t\,dW_t$. On a alors la

Proposition 30 (Intégration par parties)**.** *Si X et Y sont deux processus d'Itô, alors*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s\,dY_s + \int_0^t Y_s\,dX_s + \langle X, Y\rangle_t.$$

Théorème 31 (Formule d'Itô)**.** *Soient $(t,x) \longmapsto f(t,x)$ une fonction réelle deux fois différentiable en x et une fois différentiable en t et X un processus de Itô. On a :*

$$f(t, X_t) = f(0, X_0) + \int_0^t f_s'(s, X_s)\,ds + \int_0^t f_x'(s, X_s)\,dX_s + \frac{1}{2}\int_0^t f_{xx}''(s, X_s)\,d\langle X, X\rangle_s.$$

Nous finissons ce paragraphe en étendant la formule précédente au cas d'un mouvement brownien d–dimensionnel et d'un processus d'Itô n–dimensionnel. Les hypothèses sur les coefficients sont celles de la Définition 29.

Théorème 32. *Soit X un processus d'Itô à valeurs dans $\mathbb{R}^n$: pour $i = 1, \ldots, n$,*

$$X_t^i = X_0^i + \int_0^t K_s^i \, ds + \sum_{k=1}^d \int_0^t H_s^{i,k} \, dW_s^k.$$

Si f est deux fois différentiable en x et une fois en t on a :

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \partial_s f(s, X_s) \, ds + \sum_{i=1}^n \int_0^t \partial_{x_i} f(s, X_s) \, dX_s^i \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \int_0^t \partial^2_{x_i, x_j} f(s, X_s) \, d\langle X^i, X^j \rangle_s, \end{aligned}$$

avec $dX_s^i = K_s^i \, ds + \sum_{k=1}^d H_s^{i,k} \, dW_s^k$ et $d\langle X^i, X^j \rangle_s = \sum_{k=1}^d H_s^{i,k} H_s^{j,k} \, ds$.

Le résultat est plus simple à retenir sous forme vectorielle. Pour cela, on note X le vecteur colonne de $\mathbb{R}^n$ de coordonnées X^i, K le vecteur de $\mathbb{R}^n$ de coordonnées K^i et W le vecteur de $\mathbb{R}^d$ de coordonnée W^j. On introduit alors la matrice de taille $n \times d$, $H = (H^{i,j})_{1 \le i \le n, \, 1 \le j \le d}$. Avec ces notations, on a :

$$X_t = X_0 + \int_0^t K_s \, ds + \int_0^t H_s \, dW_s,$$

où $H_s \, dW_s$ est un produit matrice–vecteur colonne. La formule d'Itô s'écrit sous la forme, notant $x \cdot y$ le produit scalaire dans $\mathbb{R}^n$ et H^* la transposée de H

$$f(t, X_t) = f(0, X_0) + \int_0^t \partial_s f(s, X_s) \, ds + \int_0^t \nabla f(s, X_s) \cdot dX_s + \frac{1}{2} \int_0^t \operatorname{trace}\left(\mathrm{D}^2 f(s, X_s) H_s H_s^*\right) ds,$$

soit encore

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \left(\partial_s f(s, X_s) + \nabla f(s, X_s) \cdot K_s\right) ds \\ &\quad + \frac{1}{2} \int_0^t \operatorname{trace}\left(\mathrm{D}^2 f(s, X_s) H_s H_s^*\right) ds + \int_0^t \mathrm{D} f(s, X_s) H_s \, dW_s. \end{aligned}$$

3. Résultats importants

Rappelons tout d'abord la caractérisation du mouvement brownien en termes de martingales due à Paul Lévy.

Théorème 33. *Soit X une $\{\mathcal{F}_t\}_{t \ge 0}$–martingale locale continue, nulle en 0. On suppose que, pour tout $i, j \in \{1, \ldots, d\}$, $\langle X^i, X^j \rangle_t = \delta_{i,j} \, t$. Alors X est un $\{\mathcal{F}_t\}_{t \ge 0}$–mouvement brownien dans $\mathbb{R}^d$.*

Démonstration. Plaçons-nous sur $[0, T]$ et remarquons que X est une martingale de carré intégrable si la condition de crochet est satisfaite. Nous devons montrer que

$$\forall 0 \le s \le t \le T, \quad \forall u \in \mathbb{R}^d, \qquad \mathbb{E}\left(e^{iu \cdot (X_t - X_s)} \mid \mathcal{F}_s\right) = \exp\left\{-|u|^2 (t-s)/2\right\}.$$

Appliquons la formule d'Itô à la fonction $x \longmapsto e^{iu \cdot x}$ entre s et t pour obtenir

$$e^{iu \cdot X_t} = e^{iu \cdot X_s} + \int_s^t i e^{iu \cdot X_r} u \cdot dX_r - \frac{|u|^2}{2} \int_s^t e^{iu \cdot X_r} \, dr.$$

Comme X est une martingale de carré intégrable, l'intégrale stochastique précédente est une martingale ; si on multiplie l'égalité par $e^{-iu\cdot X_s}\mathbf{1}_A$ où $A \in \mathcal{F}_s$, on obtient en prenant l'espérance,

$$\mathbb{E}\left[e^{iu\cdot(X_t-X_s)}\mathbf{1}_A\right] = \mathbb{P}(A) - \frac{|u|^2}{2}\int_s^t \mathbb{E}\left[e^{iu\cdot(X_r-X_s)}\mathbf{1}_A\right] dr,$$

et par suite

$$\mathbb{E}\left[e^{iu\cdot(X_t-X_s)}\mathbf{1}_A\right] = \mathbb{P}(A)\exp\left\{-|u|^2(t-s)/2\right\},$$

ce qui conclut la preuve de ce résultat. □

Nous allons utiliser cette caractérisation du mouvement brownien pour démontrer que toute martingale dans la filtration brownienne est une intégrale stochastique par rapport au mouvement brownien. Insistons lourdement sur le fait que nous travaillons avec la filtration naturelle du mouvement brownien W et reprenons la notation $\mathcal{F}_t^W$ pour ne pas l'oublier.

Théorème 34 (Représentation des martingales browniennes)**.** *Soit M une martingale (càdlàg) de carré intégrable pour la filtration $\left\{\mathcal{F}_t^W\right\}_{t\in[0,T]}$. Alors il existe un unique processus $(H_t)_{t\in[0,T]}$, appartenant à $\mathrm{M}^2(\mathbb{R}^k)$, tel que*

$$\mathbb{P}\text{–}p.s. \qquad \forall t \in [0,T], \quad M_t = M_0 + \int_0^t H_s \cdot dW_s.$$

Il est important de remarquer que ce résultat implique que, dans la filtration brownienne, les martingales sont continues.

Démonstration. Nous nous plaçons en dimension un pour simplifier l'écriture. Rappelons que H^2 est l'espace des martingales de carré intégrable, H^2_c le sous-espace formé des martingales continues. La norme utilisée est $\|M\|^2 = \mathbb{E}\left[M_T^2\right]$ qui fait des deux espaces précédents deux espaces de Hilbert. $\mathbb{H}^2$ et $\mathbb{H}^2_c$ sont les sous-espaces de H^2 et H^2_c formés par les martingales nulles en 0. On note J l'application intégrale stochastique : $J : \mathrm{M}^2 \longrightarrow \mathbb{H}^2_c$.

On doit montrer que $\mathbb{H}^2 = J\left(\mathrm{M}^2\right)$. Pour cela remarquons tout d'abord que si $M \in \mathbb{H}^2$ il existe un unique couple de martingales (X,Y) tel que $M = X + Y$ avec $X \in J\left(\mathrm{M}^2\right)$ et $Y \in \mathbb{H}^2$ vérifiant $\langle Y, W\rangle = 0$. Cette dernière condition signifie que XY est une martingale pour toute $X \in J\left(\mathrm{M}^2\right)$. L'unicité est évidente. Pour l'existence, notons que $J\left(\mathrm{M}^2\right)_T$ l'ensemble des conditions terminales des intégrales browniennes, est un fermé de $\mathrm{L}^2\left(\mathcal{F}_T^W\right)$. Soit F l'orthogonal. On considère la décomposition orthogonale $M_T = X_T + Y_T$ et on définit $Y_t = \mathbb{E}\left(Y_T \mid \mathcal{F}_t^W\right)$. On doit montrer que le crochet de Y et W est nul ce qui revient à montrer que YW est une martingale. Si τ est un temps d'arrêt,

$$\mathbb{E}[Y_\tau W_\tau] = \mathbb{E}\left[W_\tau \mathbb{E}\left(Y_T \mid \mathcal{F}_\tau^W\right)\right] = \mathbb{E}\left[W_\tau Y_T\right] = \mathbb{E}\left[W_T^\tau Y_T\right] ;$$

la dernière espérance est nulle car W^τ appartient à $J\left(\mathrm{M}^2\right)$ et donc Y est orthogonale à W.

Comme $J\left(\mathrm{M}^2\right)$ est fermé il suffit de démontrer que $G \subset J\left(\mathrm{M}^2\right)$ pour un sous-ensemble dense de $\mathbb{H}^2$. Les martingales bornées sont denses. Mais on peut remarquer que les martingales bornées telles que Y est également bornée sont denses. En effet, si M est une martingale bornée, notons (X,Y) sa décomposition. On introduit la suite de temps d'arrêt $\sigma_n = \inf\{t,\ |X_t| \geq n\}$. Comme X est continue, X^{σ_n} et par suite $Y^{\sigma_n} = M^{\sigma_n} - X^{\sigma_n}$ sont des martingales bornées. On a bien évidemment $M_T^{\sigma_n} \longrightarrow M_T$ dans L^2. Montrons que la décomposition de M^{σ_n} est $(X^{\sigma_n}, Y^{\sigma_n})$. Il s'agit de montrer que W est orthogonal à Y^{σ_n} puisque X^{σ_n} s'écrit comme une intégrale stochastique. Or $\langle Y^{\sigma_n}, W\rangle = \langle Y, W\rangle^{\sigma_n}$ ce qui donne le résultat.

On suppose donc que X, Y, et M sont bornées par α. Soit $D = 1 + Y_T/(2\alpha)$. On a $D \geq 1/2$ et $\mathbb{E}[D] = 1$. On considère $\mathbb{P}^*$ la probabilité sur $\mathcal{F}_T^W$ de densité D par rapport à $\mathbb{P}$. Soit $X \in J\left(\mathrm{M}^2\right)$; montrons que X est une $\mathbb{P}^*$–martingale. Soit donc τ un temps d'arrêt. On a, comme $\mathbb{E}[X_\tau] = 0$,

$$\mathbb{E}^*[X_\tau] = \mathbb{E}[DX_\tau] = \mathbb{E}[Y_T X_\tau]/(2\alpha) = \mathbb{E}[Y_\tau X_\tau]/(2\alpha) = \langle Y, X\rangle_\tau/(2\alpha) = 0,$$

ce qui montre que X est une $\mathbb{P}^*$–martingale. Comme $W_t^2 - t = 2\int_0^t W_r\,dW_r$, le calcul précédent montre que W_t et $W_t^2 - t$ sont des $\mathbb{P}^*$–martingales continues. Donc d'après le théorème de Paul Lévy, sous $\mathbb{P}$ comme sous $\mathbb{P}^*$, W est un mouvement brownien. Comme D est mesurable par rapport à $\mathcal{F}_T^W$, on en déduit que $D = 1$ ce qui donne $Y_T = 0$ et donc $Y = 0$. □

On déduit de ce résultat que si ξ est une variable aléatoire de carré intégrable, $\mathcal{F}_T^W$–mesurable, il existe un unique processus $(H_t)_{t\in[0,T]} \in \mathrm{M}^2(\mathbb{R}^d)$ tel que

$$\xi = \mathbb{E}[\xi] + \int_0^T H_s \cdot dW_s.$$

Théorème 35 (Girsanov)**.** *Soit $(h_t)_{0\le t\le T}$ un processus progressivement mesurable, à valeurs dans $\mathbb{R}^d$ tel que $\mathbb{P}$–p.s. $\int_0^T |h_s|^2\,ds < +\infty$. On suppose que le processus $(D_t)_{0\le t\le T}$ défini par*

$$D_t = \exp\left\{\int_0^t h_s \cdot dW_s - \frac{1}{2}\int_0^t |h_s|^2\,ds\right\}$$

est une martingale. Soit $\mathbb{P}^*$ la mesure de densité D_T par rapport à $\mathbb{P}$ sur $\mathcal{F}_T$. Introduisons le processus $B_t = W_t - \int_0^t h_s\,ds$. Alors, sous la probabilité $\mathbb{P}^*$, B est un mouvement brownien standard.

Remarque. Lorsque $\mathbb{E}\left[\exp\left\{1/2\int_0^T |h_s|^2\,ds\right\}\right] < +\infty$, D est une martingale. Ce critère est connu sous le nom de condition de Novikov.

Nous finissons ce chapitre préliminaire par le critère de Kolmogorov. Fixons quelques notations. Soit X un processus à valeurs dans un espace de Banach indicé par un paramètre d–dimensionnel. On note $\|\cdot\|$ la norme de l'espace de Banach et $|\cdot|$ celle de $\mathbb{R}^d$. Rappelons qu'une fonction f à valeurs dans un Banach et définie sur $\mathbb{R}^d$ est localement hölderienne d'ordre α si, pour tout $a > 0$, il existe une constante C telle que :

$$\forall |x|, |y| \le a, \qquad \|f(x) - f(y)\| \le C\,|x-y|^\alpha.$$

Théorème 36 (Critère de Kolmogorov)**.** *Soit $(X_t)_{t\in[0,1]^d}$ un processus à valeurs dans un Banach. Supposons qu'il existe trois constantes strictement positives, γ, ε, C, telles que :*

$$\forall s,t \in [0,1]^d, \qquad \mathbb{E}\left[\|X_t - X_s\|^\gamma\right] \le C\,|t-s|^{d+\varepsilon}.$$

Alors il existe une modification Y de X telle que

$$\forall \alpha \in [0, \varepsilon/\gamma[, \qquad \mathbb{E}\left[\left(\sup_{s\ne t} \|Y_t - Y_s\|/|t-s|^\alpha\right)^\gamma\right] < +\infty.$$

En particulier, les trajectoires de Y sont hölderiennes d'ordre α.

Remarque. Notons que $[0,1]^d$ n'est pas fondamental. Si $d = 2$, on peut avoir par exemple $\mathbb{R}_+ \times \mathbb{R}$, mais les trajectoires de Y sont alors localement hölderiennes.

Références

[KS91] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, 2nd ed., Grad. Texts in Math., vol. 113, Springer-Verlag, New York, 1991.

[LL97] D. Lamberton and B. Lapeyre, *Introduction au calcul stochastique appliqué à la finance*, second ed., Ellipses Édition Marketing, Paris, 1997.

[RY91] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Grundlehren Math. Wiss., vol. 293, Springer-Verlag, Berlin Heidelberg New York, 1991.

Chapitre 2

Équations différentielles stochastiques rétrogrades

L'objectif de ce chapitre est d'introduire la notion d'équations différentielles stochastiques rétrogrades, EDSR en abrégé, et de préciser la terminologie employée dans ce contexte. Nous montrerons le résultat classique d'existence et d'unicité ; des exemples d'EDSR sont donnés à la fin du chapitre.

1. Vocabulaire et notations

1.1. Présentation du problème. Soient $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ un espace probabilisé filtré et variable aléatoire ξ mesurable par rapport à $\mathcal{F}_T$. On voudrait résoudre l'équation différentielle suivante :

$$\frac{-dY_t}{dt} = f(Y_t), \quad t \in [0,T], \qquad \text{avec,} \quad Y_T = \xi,$$

en imposant que, pour tout instant t, Y_t ne dépende pas du futur après t c'est à dire que le processus Y soit adapté par rapport à la filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

Prenons l'exemple le plus simple à savoir $f \equiv 0$. Le candidat naturel est $Y_t = \xi$ qui n'est pas adapté si ξ n'est pas déterministe. La meilleure approximation – disons dans L^2 – adaptée est la martingale $Y_t = \mathbb{E}(\xi \mid \mathcal{F}_t)$. Si on travaille avec la filtration naturelle d'un mouvement brownien, le théorème de représentation des martingales browniennes permet de construire un processus Z de carré intégrable et adapté tel que :

$$Y_t = \mathbb{E}(\xi \mid \mathcal{F}_t) = \mathbb{E}[\xi] + \int_0^t Z_s \, dW_s.$$

Un calcul élémentaire montre alors que

$$Y_t = \xi - \int_t^T Z_s \, dW_s, \qquad i.e. \quad -dY_t = -Z_t \, dW_t, \quad \text{avec,} \quad Y_T = \xi.$$

On voit donc apparaître sur l'exemple le plus simple une seconde inconnue qui est le processus Z dont le rôle est de rendre le processus Y adapté.

Par conséquent, comme une seconde variable apparaît, pour obtenir la plus grande généralité, on permet à f de dépendre du processus Z ; l'équation devient donc :

$$-dY_t = f(t, Y_t, Z_t) \, dt - Z_t \, dW_t, \qquad \text{avec,} \quad Y_T = \xi.$$

1.2. Notations. Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité complet et W un MB d–dimensionnel sur cet espace. On notera $\{\mathcal{F}_t\}_{t\geq 0}$ la filtration naturelle du MB W. On travaillera avec deux espaces de processus :

– on notera tout d'abord $\mathcal{S}^2(\mathbb{R}^k)$ l'espace vectoriel formé des processus Y, progressivement mesurables, à valeurs dans $\mathbb{R}^k$, tels que :

$$\|Y\|_{\mathcal{S}^2}^2 := \mathbb{E}\left[\sup_{0\le t\le T}|Y_t|^2\right] < \infty,$$

et $\mathcal{S}_c^2(\mathbb{R}^k)$ le sous-espace formé par les processus continus. Deux processus indistinguables seront toujours identifiés et nous garderons les mêmes notations pour les espaces quotients.
– et ensuite $\mathcal{M}^2(\mathbb{R}^{k\times d})$ celui formé par les processus Z, progressivement mesurables, à valeurs dans $\mathbb{R}^{k\times d}$, tels que :

$$\|Z\|_{\mathcal{M}^2}^2 := \mathbb{E}\left[\int_0^T \|Z_t\|^2 dt\right] < \infty,$$

où si $z \in \mathbb{R}^{k\times d}$, $\|z\|^2 = \mathrm{trace}(zz^*)$. $\mathrm{M}^2(\mathbb{R}^{k\times d})$ désigne l'ensemble des classes d'équivalence de $\mathcal{M}^2(\mathbb{R}^{k\times d})$.

$\mathbb{R}^k$ et $\mathbb{R}^{k\times d}$ seront souvent omis ; les espaces $\mathcal{S}^2$, $\mathcal{S}_c^2$ et M^2 sont des espaces de Banach pour les normes définies précédemment. Nous désignerons $\mathcal{B}^2$ l'espace de Banach $\mathcal{S}_c^2(\mathbb{R}^k) \times \mathrm{M}^2(\mathbb{R}^{k\times d})$.

Dans tout ce chapitre, ainsi que dans le suivant, nous nous donnons une application aléatoire f définie sur $[0,T]\times\Omega\times\mathbb{R}^k\times\mathbb{R}^{k\times d}$ à valeurs dans $\mathbb{R}^k$ telle que, pour tout $(y,z)\in\mathbb{R}^k\times\mathbb{R}^{k\times d}$, le processus $\{f(t,y,z)\}_{0\le t\le T}$ soit progressivement mesurable. On considère également une variable aléatoire ξ, mesurable par rapport à $\mathcal{F}_T$ et à valeurs dans $\mathbb{R}^k$.

Dans ce contexte, on veut résoudre l'équation différentielle stochastique rétrograde (EDSR en abrégé) suivante :

$$-dY_t = f(t,Y_t,Z_t)\,dt - Z_t\,dW_t, \quad 0\le t\le T, \qquad Y_T=\xi,$$

ou, de façon équivalente, sous forme intégrale,

$$Y_t = \xi + \int_t^T f(r,Y_r,Z_r)\,dr - \int_t^T Z_r\,dW_r, \qquad 0\le t\le T. \tag{1}$$

La fonction f s'appelle le générateur de l'EDSR et ξ la condition terminale. Sans plus tarder, précisons ce que l'on entend par solution de l'EDSR (1).

Définition 1. Une solution de l'EDSR (1) est un couple de processus $\{(Y_t,Z_t)\}_{0\le t\le T}$ vérifiant :

1. Y et Z sont progressivement mesurables à valeurs respectivement dans $\mathbb{R}^k$ et $\mathbb{R}^{k\times d}$;
2. $\mathbb{P}$–p.s. $\int_0^T \left\{|f(r,Y_r,Z_r)| + \|Z_r\|^2\right\} dr < \infty$;
3. $\mathbb{P}$–p.s., on a :

$$Y_t = \xi + \int_t^T f(r,Y_r,Z_r)\,dr - \int_t^T Z_r\,dW_r, \qquad 0\le t\le T.$$

Remarque. Il est important de retenir les deux points suivants : tout d'abord, les intégrales de l'équation (1) étant bien définies, Y est une semi-martingale continue ; ensuite, comme le processus Y est progressivement mesurable, il est adapté et donc en particulier Y_0 est une quantité déterministe.

Avant de donner un premier théorème d'existence et d'unicité, nous allons montrer, que sous une hypothèse relativement faible sur le générateur f, le processus Y appartient à $\mathcal{S}^2$.

Proposition 2. *Supposons qu'il existe un processus $\{f_t\}_{0\le t\le T}$, positif, appartenant à $\mathrm{M}^2(\mathbb{R})$ et une constante positive λ tels que*

$$\forall (t,y,z)\in[0,T]\times\mathbb{R}^k\times\mathbb{R}^{k\times d}, \qquad |f(t,y,z)| \le f_t + \lambda\left(|y| + \|z\|\right).$$

Si $\{(Y_t,Z_t)\}_{0\le t\le T}$ est une solution de l'EDSR (1) telle que $Z\in\mathrm{M}^2$ alors Y appartient à $\mathcal{S}_c^2$.

Démonstration. Le résultat se déduit principalement du lemme de Gronwall et du fait que Y_0 est déterministe. En effet, on a, pour tout $t \in [0,T]$,

$$Y_t = Y_0 - \int_0^t f(r, Y_r, Z_r)\, dr + \int_0^t Z_r\, dW_r,$$

et par suite, utilisant l'hypothèse sur f,

$$|Y_t| \leq |Y_0| + \int_0^T (f_r + \lambda \|Z_r\|)\, dr + \sup_{0 \leq t \leq T} \left| \int_0^t Z_r\, dW_r \right| + \lambda \int_0^t |Y_r|\, dr.$$

Posons

$$\zeta = |Y_0| + \int_0^T (f_r + \lambda \|Z_r\|)\, dr + \sup_{0 \leq t \leq T} \left| \int_0^t Z_r\, dW_r \right|.$$

Par hypothèse, Z appartient à M^2 et donc, via l'inégalité de Doob, le troisième terme est de carré intégrable ; il en est de même pour $\{f_t\}_{0 \leq t \leq T}$, et Y_0 est déterministe donc de carré intégrable ; il s'en suit que ζ est une variable aléatoire de carré intégrable.

Comme Y est un processus continu – cf. remarque précédente, le lemme de Gronwall fournit l'inégalité $\sup_{0 \leq t \leq T} |Y_t| \leq \zeta e^{\lambda T}$ qui montre que Y appartient à $\mathcal{S}^2$. □

Remarque. Le résultat est encore valable lorsque $\|f_\cdot\|_1$ est une variable aléatoire de carré intégrable.

Finissons par un résultat d'intégrabilité qui nous servira à plusieurs reprise.

Lemme 3. *Soient $Y \in \mathcal{S}^2(\mathbb{R}^k)$ et $Z \in \mathrm{M}^2(\mathbb{R}^{k \times d})$. Alors $\left\{ \int_0^t Y_s \cdot Z_s\, dW_s,\ t \in [0,T] \right\}$ est une martingale uniformément intégrable.*

Démonstration. Les inégalités BDG donnent

$$\begin{aligned}
\mathbb{E}\left[\sup_{0 \leq t \leq T} \left| \int_0^t Y_r \cdot Z_r\, dW_r \right| \right] &\leq C\, \mathbb{E}\left[\Big(\int_0^T |Y_r|^2 \|Z_r\|^2\, dr \Big)^{1/2} \right] \\
&\leq C\, \mathbb{E}\left[\sup_{0 \leq t \leq T} |Y_t| \Big(\int_0^T \|Z_r\|^2\, dr \Big)^{1/2} \right],
\end{aligned}$$

et par suite, comme $ab \leq a^2/2 + b^2/2$,

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} \left| \int_0^t Y_r \cdot Z_r\, dW_r \right| \right] \leq C' \left(\mathbb{E}\left[\sup_{0 \leq t \leq T} |Y_t|^2 \right] + \mathbb{E}\left[\int_0^T \|Z_r\|^2\, dr \right] \right).$$

Or cette dernière quantité est finie par hypothèse ; d'où le résultat. □

2. Le cas Lipschitz

2.1. Le résultat de Pardoux–Peng. Dans ce paragraphe, nous allons montrer un premier résultat d'existence et d'unicité qui sera généralisé au chapitre suivant. Ce résultat est dû à E. Pardoux et S. Peng [PP90] ; c'est le premier résultat d'existence et d'unicité pour les EDSR dans le cas où le générateur est non-linéaire.

Rappelons pour la dernière fois que f est définie sur $[0,T] \times \Omega \times \mathbb{R}^k \times \mathbb{R}^{k \times d}$ à valeurs dans $\mathbb{R}^k$, telle que, pour tout $(y,z) \in \mathbb{R}^k \times \mathbb{R}^{k \times d}$, le processus $\{f(t,y,z)\}_{0 \leq t \leq T}$ soit progressivement mesurable. On considère également ξ une variable aléatoire, $\mathcal{F}_T$–mesurable, à valeurs dans $\mathbb{R}^k$.

Voici les hypothèses sous lesquelles nous allons travailler.

(L) Il existe une constante λ telle que $\mathbb{P}$–p.s.,

1. condition de Lipschitz en (y,z) : pour tout t, y, y', z, z',
$$|f(t,y,z) - f(t,y',z')| \leq \lambda\left(|y-y'| + \|z-z'\|\right);$$
2. condition d'intégrabilité :
$$\mathbb{E}\left[|\xi|^2 + \int_0^T |f(r,0,0)|^2\, dr\right] < \infty.$$

Nous commençons par un cas très simple, celui où f ne dépend ni de y ni de z i.e. on se donne ξ de carré intégrable et un processus $\{F_t\}_{0\leq t\leq T}$ dans $\mathrm{M}^2(\mathbb{R}^k)$ et on veut trouver une solution de l'EDSR
$$Y_t = \xi + \int_t^T F_r\, dr - \int_t^T Z_r\, dW_r, \qquad 0 \leq t \leq T. \tag{2}$$

Lemme 4. *Soient $\xi \in \mathrm{L}^2(\mathcal{F}_T)$ et $\{F_t\}_{0\leq t\leq T} \in \mathrm{M}^2(\mathbb{R}^k)$. L'EDSR (2) possède une unique solution (Y,Z) telle que $Z \in \mathrm{M}^2$.*

Démonstration. Supposons dans un premier temps que (Y,Z) soit une solution vérifiant $Z \in \mathrm{M}^2$. Si on prend l'espérance conditionnelle sachant $\mathcal{F}_t$, on a nécessairement,
$$Y_t = \mathbb{E}\left(\xi + \int_t^T F_r\, dr \,\Big|\, \mathcal{F}_t\right).$$

On définit donc Y à l'aide de la formule précédente et il reste à trouver Z. Remarquons que, d'après le théorème de Fubini, comme F est progressivement mesurable, $\int_0^t F_r\, dr$ est un processus adapté à la filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$; en fait dans $\mathcal{S}^2_{\mathrm{c}}$ puisque F est de carré intégrable. On a alors, pour tout $t \in [0,T]$,
$$Y_t = \mathbb{E}\left(\xi + \int_0^T F_r\, dr \,\Big|\, \mathcal{F}_t\right) - \int_0^t F_r\, dr := M_t - \int_0^t F_r\, dr.$$

M est une martingale brownienne ; via le théorème de représentation des martingales browniennes on construit un processus Z appartenant à M^2 tel que
$$Y_t = M_t - \int_0^t F_r\, dr = M_0 + \int_0^t Z_r\, dW_r - \int_0^t F_r\, dr.$$

On vérifie facilement que (Y,Z) ainsi construit est une solution de l'EDSR étudiée puisque comme $Y_T = \xi$,
$$\begin{aligned} Y_t - \xi &= M_0 + \int_0^t Z_r\, dW_r - \int_0^t F_r\, dr - \left(M_0 + \int_0^T Z_r\, dW_r - \int_0^T F_r\, dr\right) \\ &= \int_t^T F_r\, dr - \int_t^T Z_r\, dW_r. \end{aligned}$$

L'unicité est évidente pour les solutions vérifiant $Z \in \mathrm{M}^2$. □

Nous montrons à présent le théorème de Pardoux et Peng.

Théorème 5. Pardoux–Peng 90. *Sous l'hypothèse (L), l'EDSR (1) possède une unique solution (Y,Z) telle que $Z \in \mathrm{M}^2$.*

Démonstration. Nous utilisons un argument de point fixe dans l'espace de Banach $\mathcal{B}^2$ en construisant une application Ψ de $\mathcal{B}^2$ dans lui-même de sorte que $(Y, Z) \in \mathcal{B}^2$ est solution de l'EDSR (1) si et seulement si c'est un point fixe de Ψ.

Pour (U, V) élément de $\mathcal{B}^2$, on définit $(Y, Z) = \Psi(U, V)$ comme étant la solution de l'EDSR :

$$Y_t = \xi + \int_t^T f(r, U_r, V_r)\, dr - \int_t^T Z_r\, dW_r, \qquad 0 \le t \le T.$$

Remarquons que cette dernière EDSR possède une unique solution qui est dans $\mathcal{B}^2$. En effet, posons $F_r = f(r, U_r, V_r)$. Ce processus appartient à M^2 puisque, f étant Lipschitz,

$$|F_r| \le |f(r, 0, 0)| + \lambda |U_r| + \lambda \|V_r\|,$$

et ces trois derniers processus sont de carré intégrable. Par suite, nous pouvons appliquer le Lemme 4 pour obtenir une unique solution (Y, Z) telle que $Z \in \mathrm{M}^2$. (Y, Z) appartient à $\mathcal{B}^2$: l'intégralité de Z est obtenue par construction et, d'après la Proposition 2, Y appartient à $\mathcal{S}^2_{\mathrm{c}}$. L'application Ψ de $\mathcal{B}^2$ dans lui-même est donc bien définie.

Soient (U, V) et (U', V') deux éléments de $\mathcal{B}^2$ et $(Y, Z) = \Psi(U, V)$, $(Y', Z') = \Psi(U', V')$. Notons $y = Y - Y'$ et $z = Z - Z'$. On a, $y_T = 0$ et

$$dy_t = -\left\{ f(t, U_t, V_t) - f(t, U'_t, V'_t) \right\} dt + z_t\, dW_t.$$

On applique la formule de Itô à $e^{\alpha t}|y_t|^2$ pour obtenir :

$$d\left(e^{\alpha t}|y_t|^2\right) = \alpha e^{\alpha t}|y_t|^2\, dt - 2e^{\alpha t} y_t \cdot \left\{ f(t, U_t, V_t) - f(t, U'_t, V'_t) \right\} dt + 2e^{\alpha t} y_t \cdot z_t\, dW_t + e^{\alpha t}\|z_t\|^2\, dt.$$

Par conséquent, intégrant entre t et T, on obtient

$$\begin{aligned} e^{\alpha t}|y_t|^2 + \int_t^T e^{\alpha r}\|z_r\|^2\, dr \quad &= \quad \int_t^T e^{\alpha r} \left(-\alpha |y_r|^2 + 2y_r \cdot \left\{ f(r, U_r, V_r) - f(r, U'_r, V'_r) \right\} \right) dr \\ &\quad - \int_t^T 2e^{\alpha r} y_r \cdot z_r\, dW_r, \end{aligned}$$

et, comme f est Lipschitz, il vient, notant u et v pour $U - U'$ et $V - V'$ respectivement,

$$e^{\alpha t}|y_t|^2 + \int_t^T e^{\alpha r}\|z_r\|^2\, dr \le \int_t^T e^{\alpha r}\big(-\alpha |y_r|^2 + 2\lambda |y_r|\, |u_r| + 2\lambda |y_r|\, \|v_r\| \big)\, dr - \int_t^T 2e^{\alpha r} y_r \cdot z_r\, dW_r.$$

Pour tout $\varepsilon > 0$, on a $2ab \le a^2/\varepsilon + \varepsilon b^2$, et donc, l'inégalité précédente donne

$$\begin{aligned} e^{\alpha t}|y_t|^2 + \int_t^T e^{\alpha r}\|z_r\|^2\, dr \quad &\le \quad \int_t^T e^{\alpha r} \left(-\alpha + 2\lambda^2/\varepsilon \right) |y_r|^2\, dr - \int_t^T 2e^{\alpha r} y_r \cdot z_r\, dW_r \\ &\quad + \varepsilon \int_t^T e^{\alpha r} \left(|u_r|^2 + \|v_r\|^2 \right) dr, \end{aligned}$$

et prenant $\alpha = 2\lambda^2/\varepsilon$, on a, notant $R_\varepsilon = \varepsilon \displaystyle\int_0^T e^{\alpha r} \left(|u_r|^2 + \|v_r\|^2 \right) dr$,

$$\forall t \in [0, T], \qquad e^{\alpha t}|y_t|^2 + \int_t^T e^{\alpha r}\|z_r\|^2\, dr \le R_\varepsilon - 2\int_t^T e^{\alpha r} y_r \cdot z_r\, dW_r. \tag{3}$$

D'après le Lemme 3, la martingale locale $\left\{ \int_0^t e^{\alpha r} y_r \cdot z_r\, dW_r \right\}_{t \in [0,T]}$ est en réalité une martingale nulle en 0 puisque Y, Y' appartiennent à $\mathcal{S}^2$ et Z, Z' appartiennent à M^2.

En particulier, prenant l'espérance – ce qui fait partir l'intégrale stochastique via la remarque précédente –, on obtient facilement, pour $t = 0$,

$$\mathbb{E}\left[\int_0^T e^{\alpha r}\|z_r\|^2\,dr\right] \leq \mathbb{E}\left[R_\varepsilon\right]. \tag{4}$$

Revenant à l'inégalité (3), les inégalités BDG fournissent – avec C universelle –,

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2\right] &\leq \mathbb{E}\left[R_\varepsilon\right] + C\,\mathbb{E}\left[\left(\int_0^T e^{2\alpha r}|y_r|^2\,\|z_r\|^2\,dr\right)^{1/2}\right] \\ &\leq \mathbb{E}\left[R_\varepsilon\right] + C\,\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t/2}|y_t|\left(\int_0^T e^{\alpha r}\|z_r\|^2\,dr\right)^{1/2}\right],\end{aligned}$$

puis, comme $ab \leq a^2/2 + b^2/2$,

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2\right] \leq \mathbb{E}\left[R_\varepsilon\right] + \frac{1}{2}\,\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2\right] + \frac{C^2}{2}\,\mathbb{E}\left[\int_0^T e^{\alpha r}\|z_r\|^2\,dr\right].$$

Prenant en considération l'inégalité (4), on obtient finalement

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2 + \int_0^T e^{\alpha r}\|z_r\|^2\,dr\right] \leq \left(3 + C^2\right)\mathbb{E}\left[R_\varepsilon\right],$$

et par suite, revenant à la définition de R_ε,

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2 + \int_0^T e^{\alpha r}\|z_r\|^2\,dr\right] \leq \varepsilon\left(3 + C^2\right)(1\vee T)\,\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|u_t|^2 + \int_0^T e^{\alpha r}\|v_r\|^2\,dr\right].$$

Prenons ε tel que $\varepsilon(3+C^2)(1\vee T) = 1/2$, de sorte que l'application Ψ est alors une contraction stricte de $\mathcal{B}^2$ dans lui-même si on le munit de la norme

$$\|(U,V)\|_\alpha = \mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|U_t|^2 + \int_0^T e^{\alpha r}\|V_r\|^2\,dr\right]^{1/2},$$

qui en fait un espace de Banach – cette dernière norme étant équivalente à la norme usuelle correspondant au cas $\alpha = 0$.

Ψ possède donc un unique point fixe, ce qui assure l'existence et l'unicité d'une solution de l'EDSR (1) dans $\mathcal{B}^2$.

On obtient ensuite une unique solution vérifiant $Z \in \mathrm{M}^2$ puisque la Proposition 2 implique qu'un telle solution appartient à $\mathcal{B}^2$. □

Remarque. À partir de maintenant et sans plus insister, l'expression « la solution de l'EDSR » signifiera la solution de l'EDSR vérifiant $Z \in \mathrm{M}^2$.

2.2. Le rôle de Z. Nous allons voir que le rôle de Z, plus précisément celui du terme $\int_t^T Z_r\,dW_r$ est de rendre le processus Y adapté et que lorsque ceci n'est pas nécessaire Z est nul.

Proposition 6. *Soit (Y,Z) la solution de l'EDSR (1) et soit τ un temps d'arrêt majoré par T. On suppose, outre l'hypothèse (L), que ξ est $\mathcal{F}_\tau$–mesurable et que $f(t,y,z) = 0$ dès que $t \geq \tau$.*

Alors $Y_t = Y_{t\wedge\tau}$ et $Z_t = 0$ si $t \geq \tau$.

Démonstration. On a, $\mathbb{P}$–p.s.,

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)\, dr - \int_t^T Z_r\, dW_r, \quad 0 \le t \le T,$$

et donc, pour $t = \tau$, comme $f(t, y, z) = 0$ dès que $t \ge \tau$,

$$Y_\tau = \xi + \int_\tau^T f(r, Y_r, Z_r)\, dr - \int_\tau^T Z_r\, dW_r = \xi - \int_\tau^T Z_r\, dW_r.$$

Il vient alors $Y_\tau = \mathbb{E}(\xi|\mathcal{F}_\tau) = \xi$ et par suite $\int_\tau^T Z_r\, dW_r = 0$ d'où l'on tire que

$$\mathbb{E}\left[\Big(\int_\tau^T Z_r\, dW_r\Big)^2\right] = \mathbb{E}\left[\int_\tau^T \|Z_r\|^2\, dr\right] = 0,$$

et finalement que $Z_r \mathbf{1}_{r \ge \tau} = 0$.

Il s'en suit immédiatement que, si $t \ge \tau$, $Y_t = Y_\tau$, puisque par hypothèse,

$$Y_\tau = Y_t + \int_\tau^t f(r, Y_r, Z_r)\, dr - \int_\tau^t Z_r\, dW_r = Y_t + 0 - 0,$$

ce qui termine la preuve. □

Notons que dans le cas où ξ et f sont déterministes alors Z est nul et Y est la solution de l'équation différentielle

$$\frac{dY_t}{dt} = f(t, Y_t, 0), \qquad Y_T = \xi.$$

2.3. Une estimation à priori. Nous finissons ce paragraphe en donnant une première estimation sur les EDSR : il s'agit en fait d'étudier la dépendance de la solution de l'EDSR par rapport aux données qui sont ξ et le processus $\{f(t, 0, 0)\}_{0 \le t \le T}$. Cette étude sera reprise au chapitre suivant.

Proposition 7. *Supposons que (ξ, f) vérifie (L). Soit (Y, Z) la solution de l'EDSR (1) telle que $Z \in \mathrm{M}^2$. Alors, il existe une constante C_u universelle telle que, pour tout $\beta \ge 1 + 2\lambda + 2\lambda^2$,*

$$\mathbb{E}\left[\sup_{0 \le t \le T} e^{\beta t}|Y_t|^2 + \int_0^T e^{\beta t}\|Z_t\|^2\, dt\right] \le C_u\, \mathbb{E}\left[e^{\beta T}|\xi|^2 + \int_0^T e^{\beta t}|f(t, 0, 0)|^2\, dt\right].$$

Démonstration. On applique la formule de Itô à $e^{\beta t}|Y_t|^2$ pour obtenir :

$$e^{\beta t}|Y_t|^2 + \int_t^T e^{\beta r}\|Z_r\|^2\, dr = e^{\beta T}|\xi|^2 + \int_t^T e^{\beta r}\left(-\beta|Y_r|^2 + 2Y_r \cdot f(r, Y_r, Z_r)\right) dr - \int_t^T 2e^{\beta r} Y_r \cdot Z_r\, dW_r.$$

Comme f est λ–Lipschitz, on a, pour tout (t, y, z),

$$2y \cdot f(t, y, z) \le 2|y|\, |f(t, 0, 0)| + 2\lambda|y|^2 + 2\lambda|y|\, \|z\|,$$

et donc utilisant le fait que $2ab \le \varepsilon a^2 + b^2/\varepsilon$ pour $\varepsilon = 1$ puis 2,

$$2y \cdot f(t, y, z) \le (1 + 2\lambda + 2\lambda^2)|y|^2 + |f(t, 0, 0)|^2 + \|z\|^2/2.$$

Pour $\beta \ge 1 + 2\lambda + 2\lambda^2$ on obtient, pour tout $t \in [0, T]$,

$$e^{\beta t}|Y_t|^2 + \frac{1}{2}\int_t^T e^{\beta r}\|Z_r\|^2\, dr \le e^{\beta T}|\xi|^2 + \int_0^T e^{\beta r}|f(r, 0, 0)|^2\, dr - 2\int_t^T e^{\beta r} Y_r \cdot Z_r\, dW_r. \quad (5)$$

La martingale locale $\left\{\int_0^t e^{\beta r} Y_r \cdot Z_r\, dW_r,\ t \in [0,T]\right\}$ est une martingale – cf. Lemme 3. En particulier, prenant l'espérance, on obtient facilement, pour $t = 0$,

$$\mathbb{E}\left[\int_0^T e^{\beta r}\|Z_r\|^2\, dr\right] \le 2\,\mathbb{E}\left[e^{\beta T}|\xi|^2 + \int_0^T e^{\beta r}|f(r,0,0)|^2\, dr\right].$$

Revenant à l'inégalité (5), les inégalités BDG fournissent – avec C universelle –,

$$\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t}|Y_t|^2\right] \le \mathbb{E}\left[e^{\beta T}|\xi|^2 + \int_0^T e^{\beta r}|f(r,0,0)|^2\, dr\right] + C\,\mathbb{E}\left[\left(\int_0^T e^{2\beta r}|Y_r|^2\,\|Z_r\|^2\, dr\right)^{1/2}\right].$$

D'autre part,

$$\begin{aligned} C\,\mathbb{E}\left[\left(\int_0^T e^{2\beta r}|Y_r|^2\,\|Z_r\|^2\, dr\right)^{1/2}\right] &\le C\,\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t/2}|Y_t|\left(\int_0^T e^{\beta r}\|Z_r\|^2\, dr\right)^{1/2}\right] \\ &\le \frac{1}{2}\,\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t}|Y_t|^2\right] + \frac{C^2}{2}\,\mathbb{E}\left[\int_0^T e^{\beta r}\|Z_r\|^2\, dr\right]. \end{aligned}$$

Il vient alors,

$$\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t}|Y_t|^2\right] \le 2\,\mathbb{E}\left[e^{\beta T}|\xi|^2 + \int_0^T e^{\beta r}|f(r,0,0)|^2\, dr\right] + C^2\,\mathbb{E}\left[\int_0^T e^{\beta r}\|Z_r\|^2\, dr\right],$$

et finalement on obtient

$$\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t}|Y_t|^2 + \int_0^T e^{\beta r}\|Z_r\|^2\, dr\right] \le 2\left(2 + C^2\right)\,\mathbb{E}\left[e^{\beta T}|\xi|^2 + \int_0^T e^{\beta r}|f(r,0,0)|^2\, dr\right],$$

ce qui termine la preuve de la proposition prenant $C_u = 2\left(2 + C^2\right)$. □

3. Théorème de comparaison

3.1. EDSR linéaires. Dans ce paragraphe nous étudions le cas particulier des EDSR linéaires pour lesquelles nous allons donner une formule plus ou moins explicite.

On se place dans le cas $k = 1$; Y est donc réel et Z est une matrice de taille $1 \times d$ c'est à dire un vecteur ligne de dimension d.

Proposition 8. *Soit $\{(a_t, b_t)\}_{t\in[0,T]}$ un processus à valeurs dans $\mathbb{R}\times\mathbb{R}^d$, progressivement mesurable et borné. Soient $\{c_t\}_{t\in[0,T]}$ un élément de $\mathrm{M}^2(\mathbb{R})$ et ξ une variable aléatoire, $\mathcal{F}_T$–mesurable, de carré intégrable, à valeurs réelles.*

L'EDSR linéaire

$$Y_t = \xi + \int_t^T \{a_r Y_r + Z_r b_r + c_r\}\, dr - \int_t^T Z_r\, dW_r,$$

possède une unique solution qui vérifie :

$$\forall t \in [0,T], \qquad Y_t = \Gamma_t^{-1}\,\mathbb{E}\left(\xi\Gamma_T + \int_t^T c_r\Gamma_r\, dr \,\middle|\, \mathcal{F}_t\right),$$

avec, pour tout $t \in [0,T]$,

$$\Gamma_t = \exp\left\{\int_0^t b_r \cdot dW_r - \frac{1}{2}\int_0^t |b_r|^2\, dr + \int_0^t a_r\, dr\right\}.$$

Démonstration. Commençons par remarquer que le processus Γ vérifie :

$$d\Gamma_t = \Gamma_t \left(a_t\, dt + b_t \cdot dW_t\right), \qquad \Gamma_0 = 1.$$

D'autre part, comme b est borné, l'inégalité de Doob montre que Γ appartient à $\mathcal{S}^2$.

De plus, les hypothèses de cette proposition assure l'existence d'une unique solution (Y, Z) à l'EDSR linéaire ; il suffit de poser $f(t, y, z) = a_t y + z b_t + c_t$ et de vérifier que (L) est satisfaite. Y appartient à $\mathcal{S}^2$ par la Proposition 2.

La formule d'intégration par parties donne

$$d\Gamma_t Y_t = \Gamma_t\, dY_t + Y_t\, d\Gamma_t + d\langle \Gamma, Y\rangle_t = -\Gamma_t c_t\, dt + \Gamma_t Z_t\, dW_t + \Gamma_t Y_t b_t \cdot dW_t,$$

ce qui montre que le processus $\Gamma_t Y_t + \int_0^t c_r \Gamma_r\, dr$ est une martingale locale qui est en fait une martingale car $c \in \mathrm{M}^2$ et Γ, Y sont dans $\mathcal{S}^2$.

Par suite,

$$\Gamma_t Y_t + \int_0^t c_r \Gamma_r\, dr = \mathbb{E}\left(\Gamma_T Y_T + \int_0^T c_r \Gamma_r\, dr \,\Big|\, \mathcal{F}_t\right),$$

ce qui donne la formule annoncée. □

Remarque. Notons que si $\xi \geq 0$ et si $c_t \geq 0$ alors la solution de l'EDSR linéaire vérifie $Y_t \geq 0$. Cette remarque va nous permettre d'obtenir le théorème de comparaison au paragraphe suivant.

Pour illustrer ce résultat prenons le cas où a et c sont nuls. On a alors

$$Y_t = \mathbb{E}\left(\xi \exp\left\{\int_t^T b_r \cdot dW_r - \frac{1}{2}\int_t^T |b_r|^2\, dr\right\} \,\Big|\, \mathcal{F}_t\right) = \mathbb{E}^*\left(\xi \mid \mathcal{F}_t\right),$$

où $\mathbb{P}^*$ est la mesure de densité par rapport à $\mathbb{P}$

$$L_T = \exp\left\{\int_0^T b_r \cdot dW_r - \frac{1}{2}\int_0^T |b_r|^2\, dr\right\}.$$

Une autre façon de voir cela, plus dans l'esprit « probabilité risque neutre », est de regarder l'EDSR sous $\mathbb{P}^*$. En effet, sous $\mathbb{P}^*$, $B_t = W_t - \int_0^t b_r\, dr$ est un MB – c'est le théorème de Girsanov. Or l'équation peut s'écrire

$$-dY_t = Z_t b_t\, dt - Z_t\, dW_t = -Z_t\, dB_t, \qquad Y_T = \xi.$$

Donc, sous $\mathbb{P}^*$, Y est une martingale, ce qui montre aussi la formule.

On retrouve ainsi les changements de mesures de probabilité du type « transformation de Girsanov ».

3.2. Théorème de comparaison. Ce paragraphe est consacré au « théorème de comparaison » qui permet de comparer les solutions de deux EDSR (dans $\mathbb{R}$) dès que l'on sait comparer les conditions terminales et les générateurs. Ce théorème est dû à l'origine à S. Peng [Pen92].

Théorème 9. *Supposons que $k = 1$ et que (ξ, f), (ξ', f') vérifient l'hypothèse (L). On note (Y, Z) et (Y', Z') les solutions des EDSR correspondantes. On suppose également que $\mathbb{P}$–p.s. $\xi \leq \xi'$ et que $f(t, Y_t, Z_t) \leq f'(t, Y_t, Z_t)$ $m \otimes \mathbb{P}$–p.p. (m mesure de Lebesgue). Alors,*

$$\mathbb{P}-p.s., \qquad \forall t \in [0, T], \qquad Y_t \leq Y'_t.$$

Si de plus, $Y_0 = Y'_0$, alors $\mathbb{P}$–p.s., $Y_t = Y'_t$, $0 \leq t \leq T$ et $f(t, Y_t, Z_t) = f'(t, Y_t, Z_t)$ $m \otimes \mathbb{P}$–p.p. En particulier, dès que $\mathbb{P}(\xi < \xi') > 0$ ou $f(t, Y_t, Z_t) < f'(t, Y_t, Z_t)$ sur un ensemble de $m \otimes \mathbb{P}$–mesure strictement positive alors $Y_0 < Y'_0$.

Démonstration. La preuve s'effectue par linéarisation ce qui permet de se ramener aux EDSR linéaires. On cherche une équation satisfaite par $U = Y' - Y$; on a notant $V = Z' - Z$ et $\zeta = \xi' - \xi$,

$$U_t = \zeta + \int_t^T (f'(r, Y'_r, Z'_r) - f(r, Y_r, Z_r))\, dr - \int_t^T V_r\, dW_r.$$

On découpe l'accroissement des f en trois morceaux en écrivant

$$\begin{aligned} f'(r, Y'_r, Z'_r) - f(r, Y_r, Z_r) &= f'(r, Y'_r, Z'_r) - f'(r, Y_r, Z'_r) + f'(r, Y_r, Z'_r) - f'(r, Y_r, Z_r) \\ &\quad + f'(r, Y_r, Z_r) - f(r, Y_r, Z_r) \quad \text{(qui est positif ici).} \end{aligned}$$

On introduit deux processus a et b : a est à valeurs réelles et b est un vecteur (colonne) de dimension d. On pose :

$$a_r = \frac{f'(r, Y'_r, Z'_r) - f'(r, Y_r, Z'_r)}{U_r}, \quad \text{si} \quad U_r \neq 0, \qquad \text{et} \quad a_r = 0 \quad \text{sinon.}$$

Pour définir b, on doit introduire une autre notation : pour $0 \leq i \leq d$, $Z_r^{(i)}$ est la ligne dont les $d - i$ dernières composantes sont celles de Z'_r et les i premières celles de Z_r. Pour $1 \leq i \leq d$, on pose

$$b_r^i = \frac{f'\left(r, Y_r, Z_r^{(i-1)}\right) - f'\left(r, Y_r, Z_r^{(i)}\right)}{V_r^i}, \quad \text{si} \quad V_r^i \neq 0, \qquad \text{et} \quad b_r^i = 0 \quad \text{sinon.}$$

Remarquons que, puisque f' est Lipschitz, ces deux processus sont progressivement mesurables et bornés. Avec ces notations, on a,

$$U_t = \zeta + \int_t^T (a_r U_r + V_r b_r + c_r)\, dr - \int_t^T V_r\, dW_r,$$

où $c_r = f'(r, Y_r, Z_r) - f(r, Y_r, Z_r)$. Par hypothèse, on a $\zeta \geq 0$ et $c_r \geq 0$. Utilisant la formule « explicite » pour les EDSR linéaires – Proposition 8 –, on a, pour $t \in [0, T]$,

$$U_t = {\Gamma_t}^{-1}\, \mathbb{E}\left(\zeta \Gamma_T + \int_t^T c_r \Gamma_r\, dr \,\middle|\, \mathcal{F}_t\right),$$

avec, pour $0 \leq r \leq T$,

$$\Gamma_r = \exp\left\{\int_0^r b_u \cdot dW_u - \frac{1}{2}\int_0^r |b_u|^2\, du + \int_0^r a_u\, du\right\}.$$

Comme déjà mentionné lors de la remarque suivant la Proposition 8, cette formule montre que $U_t \geq 0$, dès que $\zeta \geq 0$ et $c_r \geq 0$.

Pour la seconde partie du résultat, si de plus $U_0 = 0$ on a :

$$0 = \mathbb{E}\left(\zeta \Gamma_T + \int_0^T c_r \Gamma_r\, dr\right),$$

et la variable aléatoire intégrée est positive. Par conséquent, elle est nulle $\mathbb{P}$–p.s. ce qui termine la preuve de ce théorème en remarquant que dans ce cas $\zeta = 0$ et $c_r = 0$. □

Remarque. On peut supposer que $f(t, Y'_t, Z'_t) \leq f'(t, Y'_t, Z'_t)$ au lieu de $f(t, Y_t, Z_t) \leq f(t, Y_t, Z_t)$ pour obtenir le résultat précédent. Il suffit de faire une linéarisation en partant de l'écriture

$$\begin{aligned} f'(r, Y'_r, Z'_r) - f(r, Y_r, Z_r) &= f'(r, Y'_r, Z'_r) - f(r, Y'_r, Z'_r) \quad \text{(supposé positif ici)} \\ &\quad + f(r, Y'_r, Z'_r) - f(r, Y_r, Z'_r) + f(r, Y_r, Z'_r) - f(r, Y_r, Z_r). \end{aligned}$$

3.3. Modèle de Black–Scholes. Le modèle de Black–Scholes conduit à un exemple d'EDSR linéaire. Prenons le cas le plus simple. On a se place dans un marché financier sur lequel on a une action dont le prix d'une part est régi par l'EDS

$$dS_t = S_t(\mu\, dt + \sigma\, dW_t), \qquad S_0 = x,$$

où $\mu \in \mathbb{R}$, $\sigma > 0$; le paramètre σ s'appelle la volatilité. On a, pour tout $t \geq 0$,

$$S_t = x \exp\left\{\sigma W_t + (\mu - \sigma^2/2)t\right\}.$$

Parallèlement à cette action, on a un placement sans risque – disons un livret de Caisse d'épargne pour fixer les idées – dont le taux de rendement est constant, égal à r ; le prix d'une part est donné part

$$dE_t = rE_t\, dt, \quad E_0 = y, \qquad \text{i.e.} \quad E_t = ye^{rt}.$$

Une stratégie est la donnée d'un couple de processus, $(p_t, q_t)_{t\geq 0}$, adapté par rapport à la filtration du MB W ; q_t représente le nombre de parts d'actif sans risque et p_t celui d'actif risqué i.e. le nombre d'actions détenues dans le portefeuille à l'instant t. La valeur du portefeuille est donc, à l'instant t,

$$V_t = q_t E_t + p_t S_t.$$

On ne considère que des stratégies autofinancées ce qui se traduit par le fait que l'évolution de la valeur du portefeuille est décrite par

$$dV_t = q_t\, dE_t + p_t\, dS_t = rq_t E_t\, dt + p_t S_t(\mu\, dt + \sigma\, dW_t).$$

Or $q_t E_t = V_t - p_t S_t$ et donc, on a, notant $\pi_t = p_t S_t$ – la somme d'argent détenue en actions –

$$dV_t = rV_t\, dt + \pi_t \sigma(\mu - r)/\sigma\, dt + \pi_t \sigma\, dW_t,$$

soit encore notant $Z_t = \pi_t \sigma$ et $\theta = (\mu - r)/\sigma$ le « risk premium »,

$$dV_t = rV_t\, dt + \theta Z_t\, dt + Z_t\, dW_t.$$

Un problème fréquent en finance consiste à donner un prix aux options. Une option européenne d'achat, un « call », de maturité T et de prix d'exercice K est un contrat qui donne le droit mais non l'obligation à son détenteur d'acheter une part de l'action au prix d'exercice K à la date T. Le vendeur de l'option s'engage donc à payer à son détenteur la somme $(S_T - K)^+$ qui représente le profit que permet l'exercice de l'option. Plus généralement on peut imaginer un actif contingent dont le bénéfice est une variable aléatoire positive ξ qui dépend de $(S_t)_{0\leq t\leq T}$. À quel prix v vendre l'option ? Le vendeur doit s'assurer qu'en vendant l'option à ce prix à la date $t = 0$, il disposera de la somme ξ à la date $t = T$.

Pour trouver v, l'idée fondamentale est celle de duplication : le vendeur vend l'option au prix v et investit cette somme dans le marché en suivant la stratégie $(Z_t)_{0\leq t\leq T}$ à trouver ! La valeur de son portefeuille est régie par l'EDS

$$dV_t = rV_t\, dt + \theta Z_t\, dt + Z_t\, dW_t, \qquad V_0 = v.$$

Le problème est alors de trouver v et $\{Z_t\}_{0\leq t\leq T}$ de sorte que la solution de l'EDS précédente vérifie $V_T = \xi$; on dit que dans ce cas que v est le prix équitable. En d'autres termes, peut-on trouver $\{(V_t, Z_t)\}_{0\leq t\leq T}$ adapté tels que :

$$dV_t = rV_t\, dt + \theta Z_t\, dt + Z_t\, dW_t, \qquad V_T = \xi,$$

auquel cas il suffit de vendre l'option au prix $v = V_0$. Il s'agit dans ce cas de résoudre une EDSR qui est linéaire.

Supposons maintenant que « le régulateur du marché » veuille éviter la revente instantanée de l'action. Il peut soit interdire formellement cette transaction soit pénaliser les investisseurs qui se livrent à cette pratique – par exemple en leur faisant verser une somme proportionnelle à $\beta\pi_t^- = \gamma Z_t^-$ ($\gamma > 0$). Dans ce dernier cas, dupliquer un actif contingent revient à résoudre l'EDSR,

$$dV_t = (rV_t + \theta Z_t - \gamma Z_t^-)\,dt + Z_t\,dW_t, \qquad V_T = \xi.$$

L'EDSR n'est plus linéaire mais vérifie néanmoins l'hypothèse (L). Un autre exemple d'EDSR non-linéaire intervenant en finance est le suivant :

$$dV_t = (rV_t + \theta Z_t)\,dt + Z_t\,dW_t - (R-r)(V_t - Z_t/\sigma)^-\,dt, \qquad V_T = \xi.$$

On est amené à résoudre cette dernière équation pour dupliquer un actif contingent lorsque le taux d'emprunt est $R > r$.

Signalons que dans tous les exemples précédents, les stratégies sont admissibles i.e. $V_t \geq 0$. Cela résulte du théorème de comparaison puisque $\xi \geq 0$ et $f(t, 0, 0) \geq 0$.

Références

[Pen92] S. Peng, *Stochastic Hamilton-Jacobi-Bellman equations*, SIAM J. Control Optim. **30** (1992), no. 2, 284–304.

[PP90] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett. **14** (1990), no. 1, 55–61.

Chapitre 3

EDSR monotones

L'objectif de ce chapitre est d'établir un résultat d'existence et d'unicité pour les EDSR s'affranchissant partiellement de la condition de Lipschitz en y ; nous établissons également diverses estimations pour les solutions des EDSR.

1. Position du problème

Nous allons affaiblir la condition de Lipschitz de f dans la variable y pour la remplacer par une condition de monotonie. Ce type d'hypothèse est apparu dans l'article de S. Peng [Pen91] pour traiter le cas des EDSR avec temps terminal aléatoire c'est à dire des EDSR pour lesquelles on impose la condition $Y_\tau = \xi$ avec τ temps d'arrêt. Le résultat que nous présentons ici est dû à R. Darling et E. Pardoux [DP97]. L'hypothèse de monotonie est très employée : elle permet, comme déjà dit, de traiter les EDSR avec temps final aléatoire, voir par exemple [Pen91, BH98] et d'autre part d'affaiblir l'hypothèse de croissance sur f en y, voir [BC00] et surtout le résultat de E. Pardoux [Par99].

Considérons W un mouvement brownien d–dimensionnel défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ complet. Soit $f : [0,T] \times \Omega \times \mathbb{R}^k \times \mathbb{R}^{k\times d} \longrightarrow \mathbb{R}^k$ une fonction aléatoire telle que pour tout (y,z) le processus $\{f(t,y,z)\}_{0\leq t\leq T}$ soit progressivement mesurable et soit ξ une variable aléatoire $\mathcal{F}_T$–mesurable ($\mathcal{F}_t$ est, comme au chapitre précédent, la filtration augmentée de W). Voici les hypothèses de ce chapitre :

(My) Il existe des constantes $K \geq 0$, $\mu \in \mathbb{R}$, $C \geq 0$ et un processus progressivement mesurable $\{f_t\}_{0\leq t\leq T}$, positif, tels que, $\mathbb{P}$–p.s.

1. condition de Lipschitz en z :
$$\forall (t,y),\ \forall (z,z'), \qquad |f(t,y,z) - f(t,y,z')| \leq K\|z - z'\| \,;$$
2. monotonie en y : pour tout t, z,
$$\forall (y,y'), \qquad (y-y')\cdot(f(t,y,z) - f(t,y',z)) \leq \mu |y-y'|^2 \,;$$
3. croissance linéaire en (y,z) :
$$\forall (t,y,z), \qquad |f(t,y,z)| \leq f_t + C|y| + K\|z\| \,;$$
4. continuité en y : pour tout (t,z), $y \longmapsto f(t,y,z)$ est continue ;
5. ξ est $\mathcal{F}_T$–mesurable et
$$\mathbb{E}\left[|\xi|^2 + \int_0^T f_t^2\,dt\right] < \infty.$$

Remarque. On peut signaler tout d'abord que si $y \longmapsto f(y)$ est K–Lipschitz alors f est monotone avec $\mu = K$. Par contre, la fonction réelle $y \longmapsto -\sqrt{y^+}$ est monotone avec 0 pour constante μ mais n'est pas Lipschitz.

Notre objectif est à présent d'étudier l'existence et l'unicité des solutions de l'EDSR

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)\, dr - \int_t^T Z_r\, dW_r, \qquad 0 \le t \le T, \tag{1}$$

lorsque (ξ, f) vérifie l'hypothèse (My).

Remarquons ensuite que sous l'hypothèse (My), si (Y, Z) est solution de l'EDSR (1) alors Y appartient à $\mathcal{S}^2$ dès que Z appartient à M^2. C'est une conséquence directe de la Proposition 2.2.

2. Estimations à priori

Dans ce paragraphe nous allons donner plusieurs estimations « à priori » concernant les solutions des EDSR. Ces estimations traduisent la dépendance de la solution d'une EDSR par rapport aux paramètres d'entrée i.e. le générateur et la condition terminale. Nous présentons ici des estimations qui font intervenir des constantes universelles (plutôt que des constantes dépendant de T) ce qui peut s'avérer utile dans certains cas. Ce type d'estimations semble être apparu à la suite d'un travail de R. Buckdahn, M. Quicampoix et A. Rascanu [BQR97]. D'autres estimations à priori étaient présentes dans les versions préliminaires de l'article de N. El Karoui, S. Peng et M.-C. Quenez [EKPQ97]. Rappelons que nous notons $\mathcal{B}^2$ l'espace de Banach $\mathcal{S}^2_c \times \mathrm{M}^2$.

Proposition 1. *Soit $\xi \in \mathrm{L}^2(\mathcal{F}_T)$. Supposons qu'il existe des constantes μ, K et un processus $\{f_t\}_{0\le t\le T}$ appartenant à $\mathrm{M}^2(\mathbb{R}^+)$ tels que, $\mathbb{P}$–p.s.*

$$\forall (t, y, z), \qquad y \cdot f(t, y, z) \le |y|\, f_t + \mu |y|^2 + K|y|\, \|z\|.$$

Soit $(Y, Z) \in \mathcal{B}^2$ solution de l'EDSR (1). Alors, il existe une constante universelle C_u telle que

$$\mathbb{E}\left[\sup_{0\le t\le T} e^{\alpha t}|Y_t|^2 + \int_0^T e^{\alpha t}\|Z_t\|^2\, dt\right] \le C_u\, \mathbb{E}\left[e^{\alpha T}|\xi|^2 + \Big(\int_0^T e^{\alpha t/2} f_t\, dt\Big)^2\right],$$

avec $\alpha = 2(\mu + K^2)$. De plus, on a, pour tout $t \in [0, T]$, notant $\gamma = 1 + 2\mu + K^2$,

$$|Y_t|^2 \le \mathbb{E}\left(e^{\gamma(T-t)}|\xi|^2 + \int_t^T e^{\gamma(r-t)} f_r^2\, dr \,\Big|\, \mathcal{F}_t\right).$$

Remarque. Si f vérifie l'hypothèse (My), on peut prendre $f_t = |f(t, 0, 0)|$. On peut également noter, pour le second point, que dès que ξ et f_t sont bornés par une constante M,

$$\sup\nolimits_{0\le t\le T} |Y_t|^2 \le M^2\, C(\gamma, T).$$

On peut toujours prendre $C(\gamma, T) = (1 + T)e^{\gamma^+ T}$ et si $\gamma \ge 1$, $C(\gamma, T) = 2e^{\gamma T}$ convient aussi.

Cette remarque nous servira pour établir l'existence des solutions sous (MY) ; elle est due initialement à S. Peng [Pen92].

Démonstration. On applique la formule de Itô à $e^{at}|Y_t|^2$, a étant un réel, pour obtenir :

$$e^{at}|Y_t|^2 + \int_t^T e^{ar}\|Z_r\|^2\, dr = e^{aT}|\xi|^2 + \int_t^T e^{ar}\left(-a|Y_r|^2 + 2Y_r \cdot f(r, Y_r, Z_r)\right) dr - \int_t^T 2e^{ar} Y_r \cdot Z_r\, dW_r.$$

Par hypothèse, on a, pour tout (t, y, z),

$$2y \cdot f(t, y, z) \le 2|y|\, f_t + 2\mu|y|^2 + 2K|y|\, \|z\|,$$

et on utilise le fait que $2ab \leq \varepsilon a^2 + b^2/\varepsilon$ pour tout $\varepsilon > 0$.

Pour la deuxième partie de la proposition, on utilise l'inégalité

$$2y \cdot f(t,y,z) \leq (1 + 2\mu + K^2)|y|^2 + f_t^2 + \|z\|^2.$$

Si on prend $a = \gamma = 1 + 2\mu + K^2$ on obtient, pour tout $t \in [0,T]$,

$$e^{\gamma t}|Y_t|^2 \leq e^{\gamma T}|\xi|^2 + \int_t^T e^{\gamma r} f_r^2 \, dr - 2\int_t^T e^{\gamma r} Y_r \cdot Z_r \, dW_r.$$

Un argument utilisé plusieurs fois déjà montre que la martingale locale $\left\{\int_0^t e^{\gamma r} Y_r \cdot Z_r \, dW_r\right\}_{t\in[0,T]}$ est en réalité une martingale uniformément intégrable. Il reste à conditionner par rapport à $\mathcal{F}_t$ pour obtenir la seconde partie de la proposition.

Pour la première partie, on commence par écrire

$$2y \cdot f(t,y,z) \leq 2(\mu + K^2)|y|^2 + 2|y|\, f_t + \|z\|^2/2,$$

puis on prend $a = \alpha = 2(\mu + K^2)$ pour obtenir

$$e^{\alpha t}|Y_t|^2 + \frac{1}{2}\int_t^T e^{\alpha r}\|Z_r\|^2 \, dr \leq e^{\alpha T}|\xi|^2 + 2\int_0^T e^{\alpha r}|Y_r|\, f_r \, dr - 2\int_t^T e^{\alpha r} Y_r \cdot Z_r \, dW_r. \qquad (2)$$

En particulier, prenant l'espérance – ce qui fait partir l'intégrale stochastique via la remarque précédente –, on obtient facilement, pour $t = 0$,

$$\mathbb{E}\left[\int_0^T e^{\alpha r}\|Z_r\|^2 \, dr\right] \leq 2\,\mathbb{E}\left[e^{\alpha T}|\xi|^2 + 2\int_0^T e^{\alpha r}|Y_r|\, f_r \, dr\right].$$

Revenant à l'inégalité (2), les inégalités BDG fournissent – avec C universelle –,

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|Y_t|^2\right] \leq \mathbb{E}\left[e^{\alpha T}|\xi|^2 + 2\int_0^T e^{\alpha r}|Y_r|\, f_r \, dr\right] + C\,\mathbb{E}\left[\Big(\int_0^T e^{2\alpha r}|Y_r|^2\, \|Z_r\|^2 \, dr\Big)^{1/2}\right].$$

D'autre part,

$$\begin{aligned} C\,\mathbb{E}\left[\Big(\int_0^T e^{2\alpha r}|Y_r|^2\, \|Z_r\|^2 \, dr\Big)^{1/2}\right] &\leq C\,\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t/2}|Y_t|\,\Big(\int_0^T e^{\alpha r}\|Z_r\|^2 \, dr\Big)^{1/2}\right] \\ &\leq \frac{1}{2}\,\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|Y_t|^2\right] + \frac{C^2}{2}\,\mathbb{E}\left[\int_0^T e^{\alpha r}\|Z_r\|^2 \, dr\right]. \end{aligned}$$

Il vient alors,

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|Y_t|^2\right] \leq 2\,\mathbb{E}\left[e^{\alpha T}|\xi|^2 + 2\int_0^T e^{\alpha r}|Y_r|\, f_r \, dr\right] + C^2\,\mathbb{E}\left[\int_0^T e^{\alpha r}\|Z_r\|^2 \, dr\right],$$

et finalement on obtient

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|Y_t|^2 + \int_0^T e^{\alpha r}\|Z_r\|^2 \, dr\right] \leq 2(2 + C^2)\,\mathbb{E}\left[e^{\alpha T}|\xi|^2 + 2\int_0^T e^{\alpha r}|Y_r|\, f_r \, dr\right].$$

Pour finir la preuve de la proposition, il suffit de remarquer que

$$\begin{aligned} 4(2 + C^2)\,\mathbb{E}\left[\int_0^T e^{\alpha r}|Y_r|\, f_r \, dr\right] &\leq 4(2 + C^2)\,\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t/2}|Y_t| \int_0^T e^{\alpha r/2} f_r \, dr\right] \\ &\leq \frac{1}{2}\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|Y_t|^2\right] + 8(2 + C^2)^2\,\mathbb{E}\left[\Big(\int_0^T e^{\alpha r/2} f_r \, dr\Big)^2\right], \end{aligned}$$

pour obtenir au total

$$\mathbb{E}\left[\sup_{0\le t\le T} e^{\alpha t}|Y_t|^2 + \int_0^T e^{\alpha r}\|Z_r\|^2\,dr\right] \le 16(2+C^2)^2\,\mathbb{E}\left[e^{\alpha T}|\xi|^2 + \Big(\int_0^T e^{\alpha r/2} f_r\,dr\Big)^2\right].$$

□

La proposition précédente possède le corollaire important suivant.

Corollaire 2. *Soient (ξ, f) et (ξ', f') vérifiant (My) avec des constantes (μ, K, C), (μ', K', C') ; soient (Y, Z) et (Y', Z') des solutions des EDSR associées telles que Z, Z' appartiennent à M^2. Il existe une constante universelle C_u telle que*

$$\mathbb{E}\left[\sup_{0\le t\le T} e^{\alpha t}|\delta Y_t|^2 + \int_0^T e^{\alpha t}\|\delta Z_t\|^2\,dt\right] \le C_u\,\mathbb{E}\left[e^{\alpha T}|\delta\xi|^2 + \Big(\int_0^T e^{\alpha t/2}\,|\delta f(t, Y'_t, Z'_t)|\,dt\Big)^2\right],$$

avec $\alpha = 2(\mu + K^2)$ et les notations classiques $\delta Y = Y - Y'$, $\delta Z = Z - Z'$, $\delta\xi = \xi - \xi'$ de même que $\delta f(t, y, z) = f(t, y, z) - f'(t, y, z)$.

Démonstration. Remarquons tout d'abord que sous (My), dès que $Z \in \mathrm{M}^2$, (Y, Z) solution de l'EDSR de paramètres ξ et f appartient à $\mathcal{B}^2$. Notons que $(\delta Y, \delta Z)$ est solution de l'EDSR de paramètres $\delta\xi$ et $g(t, y, z) = f(t, y + Y'_t, z + Z'_t) - f'(t, Y'_t, Z'_t)$. Or, d'après l'hypothèse (My),

$$\begin{aligned} y\cdot g(t,y,z) \quad &= \quad y\cdot\left(f(t, y+Y'_t, z+Z'_t) - f(t, Y'_t, z+Z'_t)\right) + y\cdot\left(f(t, Y'_t, z+Z'_t) - f(t, Y'_t, Z'_t)\right) \\ &\quad + y\cdot\left(f(t, Y'_t, Z'_t) - f'(t, Y'_t, Z'_t)\right), \end{aligned}$$

d'où l'on déduit que $y\cdot g(t, y, z) \le \mu|y|^2 + K|y|\,\|z\| + |y|\,|\delta f(t, Y'_t, Z'_t)|$. Il suffit d'appliquer la proposition précédente pour conclure. □

Remarque. Un bon exercice consiste à montrer que sous l'hypothèse de la proposition précédente, on a, pour une autre constante universelle C_u,

$$\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t}|Y_t|^2 + \int_0^T e^{\beta t}\|Z_t\|^2 dt\right] \le C_u\,\mathbb{E}\left[e^{\beta T}|\xi|^2 + \int_0^T e^{\beta t} f_t^2\,dt\right],$$

avec $\beta = 1 + 2(\mu + K^2)$. Il suffit de s'inspirer de la Proposition 2.7. La remarque vaut aussi pour le Corollaire 2.

3. Existence et unicité sous (My)

Nous allons à présent nous intéresser à l'existence et l'unicité de solutions de notre EDSR sous (My). La démonstration de ce résultat est plus délicate ; elle se décompose en deux étapes : tout d'abord nous traitons le cas d'un générateur f indépendant de z, puis nous utilisons un argument de point fixer pour traiter le cas général. La première étape est de loin la plus difficile.

3.1. Le résultat. Dans un premier temps nous admettons le résultat lorsque f ne dépend que de y et nous en déduisons le théorème général.

Proposition 3. *Sous (My), pour tout processus $V \in \mathrm{M}^2$, l'EDSR*

$$Y_t = \xi + \int_t^T f(r, Y_r, V_r)dr - \int_t^T Z_r dW_r, \qquad 0 \le t \le T,$$

possède une unique solution telle que $Z \in \mathrm{M}^2$.

Muni de cette proposition, nous démontrons le résultat principal de ce chapitre.

Théorème 4. *Sous (My), l'EDSR (1) possède une unique solution telle que* $Z \in \mathrm{M}^2$.

Démonstration. L'unicité est évidente : il suffit en effet d'appliquer le Corollaire 2. Pour l'existence nous utilisons un argument de point fixe basé sur la proposition précédente. Pour (U, V) élément de $\mathcal{B}^2$, notons $(Y, Z) = \Psi(U, V)$ la solution de l'EDSR de la proposition précédente. Ψ est une application du Banach $\mathcal{B}^2$ dans lui même. Si $(Y', Z') = \Psi(U', V')$, les estimations à priori – cf. Corollaire 2 – donnent, si $\alpha = 2\mu$,

$$\mathbb{E}\left[\sup_{0\le t\le T}|\delta Y_t|^2 + \int_0^T \|\delta Z_t\|^2\,dt\right] \le C_u e^{|\alpha|T}\,\mathbb{E}\left[\Big(\int_0^T |f(t,Y_t,V_t) - f(t,Y_t,V_t')|\,dt\Big)^2\right],$$

où δ a sa signification habituelle. L'inégalité de Hölder conduit à la majoration, comme f est K–Lipschitz en z,

$$\mathbb{E}\left[\sup_{0\le t\le T}|\delta Y_t|^2 + \int_0^T \|\delta Z_t\|^2\,dt\right] \le C_u e^{|\alpha|T} K^2 T\,\mathbb{E}\left[\int_0^T \|\delta V_t\|^2\,dt\right] \;;$$

cette dernière inégalité montre que Ψ est une contraction stricte si T est suffisamment petit et fournit donc le résultat sous cette restriction. Dans le cas général, il suffit de subdiviser, l'intervalle de temps $[0, T]$ en un nombre fini d'intervalles de longueur convenable : on résout d'abord sur $[T - \eta, T]$ puis sur $[T - 2\eta, T - \eta]$, ... □

3.2. Preuve de la Proposition 3. L'unicité a déjà été étudiée lors du théorème précédent. Passons maintenant aux choses sérieuses : l'existence. Soit donc $V \in \mathrm{M}^2$; notons $h(t, y) = f(t, y, V_t)$. On doit construire une solution pour l'EDSR

$$Y_t = \xi + \int_t^T h(r, Y_r)\,dr - \int_t^T Z_r\,dW_r, \qquad 0 \le t \le T. \tag{3}$$

La fonction h vérifie (My) puisque :
- $|h(t, y)| \le h_t + C|y|$; où $h_t = f_t + K\|V_t\|$ appartient à M^2 ;
- $(y - y')\cdot(h(t, y) - h(t, y')) \le \mu|y - y'|^2$;
- $y \longmapsto h(t, y)$ est continue ;
- ξ est de carré intégrable.

Notons tout d'abord que l'on peut supposer que $\mu = 0$ sans perte de généralité puisque (Y, Z) est solution de l'EDSR (3) si et seulement si $(Y'_t, Z'_t) = (e^{\mu t}Y_t, e^{\mu t}Z_t)$ est solution de l'EDSR de paramètres (ξ', h') où $\xi' = e^{\mu T}\xi$ et $h'(t, y) = e^{\mu t}h(t, e^{-\mu t}y) - \mu y$. Or (ξ', h') vérifie (My) avec $\mu' = 0$, $h'_t = e^{\mu t}h_t$ et $C' = C + |\mu|$. On suppose donc que $\mu = 0$.

Étape 1 : ξ et $\sup_t h_t$ bornés par un réel M. On considère alors une fonction $\rho : \mathbb{R}^k \longrightarrow \mathbb{R}_+$, de classe $\mathcal{C}^\infty$ dont le support est la boule unité et telle que $\int \rho(u)du = 1$. Soit $n \in \mathbb{N}^*$. On note $\rho_n(u) = n^k\rho(nu)$. Soit d'autre part $\theta_n : \mathbb{R}^k \longrightarrow [0, 1], \mathcal{C}^\infty$, telle que $\theta_n(y) = 1$ si $|y| \le n$ et $\theta_n(y) = 0$ dès que $|y| \ge n + 1$. On définit enfin, pour $n \ge 1$,

$$h_n(t, y) = \rho_n * \theta_n h(t, \cdot)(y) = \int_{\mathbb{R}^k} \rho_n(y - u)\theta_n(u)h(t, u)du = \int_{\mathbb{R}^k} \rho_n(u)\theta_n(y - u)h(t, y - u)du.$$

Étudions d'abord les propriétés de h_n. La première chose à remarquer est que $y \longmapsto h_n(t, y)$ est de classe $\mathcal{C}^\infty$ à support compact : en fait, $h_n(t, y) = 0$ dès que $|y| > n + 2$ (pour tout (t, ω)). De plus, d'après l'hypothèse de croissance de h, on a

$$|h_n(t, y)| \le \int \rho_n(u)|h(t, y - u)|\,du \le \int \rho_n(u)\,(h_t + C|y| + C|u|)\,du,$$

et donc comme le support de ρ est la boule unité

$$|h_n(t,y)| \leq (M+C) + C|y|. \tag{4}$$

En particulier, $h_n(t,0)$ est un processus borné par $M+C$.

Montrons que la fonction h_n est Lipschitz en y uniformément par rapport à (t,ω) en estimant $\sup_y \|\nabla h_n(t,y)\|$. Si $|y| > n+2$, $\nabla h_n(t,y) = 0$. Pour $|y| \leq n+2$, on a

$$\|\nabla h_n(t,y)\| \leq \left\| \int \nabla\rho_n(u) \otimes \theta_n(y-u) h(t,y-u)\, du \right\| \leq \int |\nabla\rho_n(u)| \left(h_t + C|y| + C|u|\right) du,$$

et par suite, comme le support de ρ est inclus dans la boule unité, on a pour $|y| \leq n+2$,

$$\|\nabla h_n(t,y)\| \leq n\left\{(M+C) + C|y|\right\} \int |\nabla\rho(u)|\, du \leq C'\, n^2.$$

On peut donc appliquer le Théorème 2.5, pour obtenir une unique solution (Y^n, Z^n) dans $\mathcal{B}^2$ de l'EDSR

$$Y_t^n = \xi + \int_t^T h_n(r, Y_r^n)\, dr - \int_t^T Z_r^n\, dW_r, \qquad 0 \leq t \leq T.$$

Comme ξ est bornée par M, la majoration (4) permet d'appliquer la remarque suivant la Proposition 1, pour obtenir $\sup_n \sup_t |Y_t^n|^2 \leq 2(M+C)^2 \exp\{(1+2C)T\}$. En fait, on peut obtenir une meilleure estimation en remarquant que $y \cdot h_n(t,y) \leq (M+C)|y|$. En effet, on a

$$y \cdot h_n(t,y) = \int \rho_n(u)\theta_n(y-u) y \cdot h(t,y-u)\, du,$$

et on écrit $y \cdot h(t,y-u) = y \cdot (h(t,y-u) - h(t,-u)) + y \cdot h(t,-u) \leq 0 + |y|(M + C|u|)$. Il suffit alors d'intégrer en tenant compte du fait que le support de ρ_n est inclus dans la boule unité. La remarque de la page 2 conduit à

$$\sup_n \sup_t |Y_t^n|^2 \leq 2(M+C)^2 \exp\{T\}. \tag{5}$$

Cette dernière propriété va jouer un rôle important : la fonction h_n n'est pas nécessairement monotone dans tout l'espace (ce que l'on voulait faire au départ) mais h_n est monotone dans la boule de centre 0 et de rayon $n-1$. En effet, si $|y| \leq n-1$,

$$h_n(t,y) = \int_{|u|\leq 1} \rho_n(u)\theta_n(y-u) h(t,y-u)\, du = \int_{|u|\leq 1} \rho_n(u) h(t,y-u)\, du,$$

puisque $\theta_n(x) = 1$ si $|x| \leq n$. Par suite, si $|y| \leq n-1$ et $|y'| \leq n-1$, alors, comme h est monotone,

$$(y-y') \cdot (h_n(t,y) - h_n(t,y')) = \int \rho_n(u)(y-y') \cdot (h(t,y-u) - h(t,y'-u))\, du \leq 0.$$

Nous allons à présent montrer que la suite (Y^n, Z^n) converge dans $\mathcal{B}^2$. Pour cela prenons, $m \geq n \geq 1+a$ où $a = \sqrt{2}(M+C)\exp(T/2)$. La formule d'Itô donne, si on note $\delta Y = Y^m - Y^n$, $\delta Z = Z^m - Z^n$,

$$|\delta Y_t|^2 + \int_t^T \|\delta Z_r\|^2\, dr = 2\int_t^T \delta Y_r \cdot \{h_m(r, Y_r^m) - h_n(r, Y_r^n)\}\, dr - 2\int_t^T \delta Y_r \cdot \delta Z_r\, dW_r.$$

On ne peut pas appliquer directement les estimations à priori car les fonctions h_m et h_n ne sont pas globalement monotones. Toutefois, h_m est monotone dans la boule de rayon $m-1$; comme $m-1 \geq a$, Y_r^m et Y_r^n appartiennent à cette boule d'après l'estimation (5). Par suite,

$$\begin{aligned}
\delta Y_r.\{h_m(r,Y_r^m) - h_n(r,Y_r^n)\} &\leq \delta Y_r.\{h_m(r,Y_r^m) - h_m(r,Y_r^n)\} + \delta Y_r.\{h_m(r,Y_r^n) - h_n(r,Y_r^n)\} \\
&\leq 0 + 2a \sup_{|y|\leq a} |h_m(t,y) - h_n(t,y)|.
\end{aligned}$$

Il vient alors,

$$|\delta Y_t|^2 + \int_t^T \|\delta Z_r\|^2\,dr = 4a\int_0^T \sup_{|y|\le a} |h_m(t,y) - h_n(t,y)|\,dr - 2\int_t^T \delta Y_r\cdot\delta Z_r\,dW_r.$$

Utilisant la même démarche que pour les estimations à priori – on établit l'estimation pour le terme en Z puis on emploie les inégalités BDG –, on obtient pour une constante C,

$$\mathbb{E}\left[\sup_{0\le t\le T}|\delta Y_t|^2 + \int_0^T \|\delta Z_r\|^2\,dr\right] \le C\,\mathbb{E}\left[\int_0^T \sup_{|y|\le a}|h_m(t,y)-h_n(t,y)|\,dr\right].$$

Comme la fonction $y \longmapsto h(t,y)$ est continue, $h_n(t,\cdot)$ converge vers $h(t,\cdot)$ uniformément sur les compacts $\lambda\otimes\mathbb{P}$–presque partout. De plus, d'après la majoration (4), on a

$$\sup_{|y|\le a}|h_m(t,y)-h_n(t,y)| \le 2(M+C+Ca).$$

Le théorème de convergence dominée de Lebesgue montre que la suite (Y^n,Z^n) est de Cauchy dans l'espace de Banach $\mathcal{B}^2$.

Soit (Y,Z) la limite de cette suite ; montrons que (Y,Z) est bien solution de l'EDSR (3). Pour cela, on va passer à la limite terme à terme dans l'EDSR dirigée par h_n. Y_t^n converge vers Y_t dans L^2 et $\int_t^T Z_r^n\,dW_r$ converge vers $\int_t^T Z_r\,dW_r$ également dans L^2 puisque

$$\mathbb{E}\left[|Y_t^n - Y_t|^2\right] \le \mathbb{E}\left[\sup_t |Y_t^n - Y_t|^2\right], \quad \mathbb{E}\left[\left|\int_t^T (Z_r^n - Z_r)\,dW_r\right|^2\right] \le 4\,\mathbb{E}\left[\int_0^T \|Z_r^n - Z_r\|^2\,dr\right].$$

Pour finir, notons que

$$\begin{aligned}&\mathbb{E}\left[\sup_t \left|\int_t^T \{h_n(r,Y_r^n) - h(r,Y_r)\}\,dr\right|^2\right]\\ &\le\ 2T\,\mathbb{E}\left[\int_0^T |h_n(r,Y_r^n)-h(r,Y_r^n)|^2\ dr\right] + 2T\,\mathbb{E}\left[\int_0^T |h(r,Y_r^n)-h(r,Y_r)|^2\ dr\right]\ ;\end{aligned}$$

le premier terme tend vers 0 car $|h_n(r,Y_r^n) - h(r,Y_r^n)| \le \sup_{|y|\le a}|h_n(r,y)-h(r,y)|$. Le second terme tend vers 0 car, comme la fonction $h(t,\cdot)$ est continue $h(t,Y_t^n) \longrightarrow h(t,Y_t)$. Dans chacun des cas les dominations sont immédiates. Ceci termine la première étape.

Étape 2 : cas général. Pour tout entier $p\ge 1$, définissons $\xi^p = \xi\mathbf{1}_{|\xi|\le p}$, $h^p(t,y) = h(t,y)\mathbf{1}_{h_t\le p}$. Le couple (ξ^p,h^p) vérifie les hypothèses de l'étape précédente. h^p est monotone, $y\longmapsto h^p(t,y)$ est continue et

$$|\xi^p|\le p, \qquad |h^p(t,y)| \le p + C|y|.$$

L'EDSR de paramètres (ξ^p,h^p) possède donc une unique solution dans $\mathcal{B}^2$, (Y^p,Z^p). Par monotonie, $y.h^p(t,y) \le |y|\,|h^p(t,0)| \le p|y|$. La remarque suivant la Proposition 1 montre alors que $\sup_t |Y^p|^2 \le 2p^2e^T$. Montrons que la suite (Y^p,Z^p) est de Cauchy dans $\mathcal{B}^2$. Soit donc $q\ge p$. On utilise les estimations à priori – cf. Corollaire 2 – pour obtenir

$$\mathbb{E}\left[\sup_{0\le t\le T}|\delta Y_t|^2 + \int_0^T \|\delta Z_t\|^2\,dt\right] \le C_u\,\mathbb{E}\left[\left(\int_0^T \mathbf{1}_{h_t>p}\ |h(t,Y_t^p)|\,dt\right)^2\right].$$

Or $|h(t,Y_t^p)| \le h_t + C|Y_t^p| \le h_t + C\sqrt{2}pe^{T/2}$. Par suite,

$$\mathbf{1}_{h_t>p}\ |h(t,Y_t^p)| \le \left(1 + C\sqrt{2}e^{T/2}\right)\ h_t\mathbf{1}_{h_t>p}.$$

L'inégalité de Hölder fournit alors

$$\mathbb{E}\left[\sup_{0\le t\le T}|\delta Y_t|^2+\int_0^T\|\delta Z_t\|^2\,dt\right]\le C_u T\left(1+C\sqrt{2}e^{T/2}\right)^2\,\mathbb{E}\left[\int_0^T \mathbf{1}_{h_t>p}h_t^2\,dt\right]\ ;$$

le dernier majorant tend vers 0 puisque h_t appartient à M^2.

Pour terminer la preuve de cette proposition, il reste à montrer que la limite de la suite (Y^p, Z^p) est la solution de l'EDSR (3). On a

$$Y_t^p=\xi^p+\int_t^T h^p(r,Y_r^p)dr-\int_t^T Z_r^p dW_r,\qquad 0\le t\le T,$$

et comme dans l'étape précédente, $\left(\xi^p, Y_t^p, \int_t^T Z_r^p dW_r\right)$ converge vers $\left(\xi, Y_t, \int_t^T Z_r dW_r\right)$ dans L^2. Reste à étudier la convergence de $\int_t^T h^p(r,Y_r^p)\,dr$ vers $\int_t^T h(r,Y_r)\,dr$. Procédant de même que lors de la première étape, on a

$$\begin{aligned}&\mathbb{E}\left[\sup_t\left|\int_t^T\{h^p(r,Y_r^p)-h(r,Y_r)\}\,dr\right|^2\right]\\ &\le\ 2T\,\mathbb{E}\left[\int_0^T|h^p(r,Y_r^p)-h(r,Y_r^p)|^2\ dr\right]+2T\,\mathbb{E}\left[\int_0^T|h(r,Y_r^p)-h(r,Y_r)|^2\ dr\right]\ ;\end{aligned}$$

Le premier terme est majoré par

$$2T\left(1+C\sqrt{2}e^{T/2}\right)^2\,\mathbb{E}\left[\int_0^T \mathbf{1}_{h_t>p}h_t^2\,dt\right]$$

qui tend vers 0 comme déjà dit. Pour le second terme, la continuité de $h(t,\cdot)$ montre que l'intégrant $|h(r,Y_r^p)-h(r,Y_r)|$ tend vers 0 $\lambda\otimes\mathbb{P}$–presque partout. On a de plus

$$|h(r,Y_r^p)|^2\le 2\left(h_r^2+C^2|Y_r^p|^2\right)$$

ce qui donne l'équi–intégrabilité puisque h_r appartient à M^2 et Y^p converge vers Y dans M^2.

On obtient donc en passant à la limite terme à terme dans L^2 dans l'EDSR dirigée par (ξ^p, h^p),

$$Y_t=\xi+\int_t^T h(r,Y_r)\,dr-\int_t^T Z_r\,dW_r,$$

ce qui achève la démonstration.

Références

[BC00] Ph. Briand and R. Carmona, *BSDEs with polynomial growth generators*, J. Appl. Math. Stochastic Anal. **13** (2000), no. 3, 207–238.

[BH98] Ph. Briand and Y. Hu, *Stability of BSDEs with random terminal time and homogenization of semilinear elliptic PDEs*, J. Funct. Anal. **155** (1998), no. 2, 455–494.

[BQR97] R. Buckdahn, M. Quincampoix, and A. Rascanu, *Propriété de viabilité pour des équations différentielles stochastiques rétrogrades et applications à des équations aux dérivées partielles*, C. R. Acad. Sci. Paris Sér. I Math. **325** (1997), no. 11, 1159–1162.

[DP97] R. W. R. Darling and E. Pardoux, *Backwards SDE with random terminal time and applications to semilinear elliptic PDE*, Ann. Probab. **25** (1997), no. 3, 1135–1159.

[EKPQ97] N. El Karoui, S. Peng, and M.-C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance **7** (1997), no. 1, 1–71.

[Par99] E. Pardoux, *BSDEs, weak convergence and homogenization of semilinear PDEs*, Nonlinear analysis, differential equations and control (Montreal, QC, 1998), Kluwer Acad. Publ., Dordrecht, 1999, pp. 503–549.

[Pen91] S. Peng, *Probabilistic interpretation for systems of quasilinear parabolic partial differential equations*, Stochastics Stochastics Rep. **37** (1991), no. 1-2, 61–74.

[Pen92] ——, *Stochastic Hamilton-Jacobi-Bellman equations*, SIAM J. Control Optim. **30** (1992), no. 2, 284–304.

Chapitre 4

Équations différentielles stochastiques

Dans ce chapitre, nous rappelons tout d'abord le théorème d'existence et d'unicité sur les équations différentielles stochastiques – EDS en abrégé dans la suite du cours – puis nous nous intéresserons au flot stochastique engendré par une EDS. Nous finirons par la propriété de Markov.

1. Le résultat classique d'Itô

1.1. Définitions et notations. On se place toujours sur un espace de probabilité complet, disons $(\Omega, \mathcal{F}, \mathbb{P})$ et on se donne W un MB d–dimensionnel sur cet espace. On considère également une variable aléatoire Z, de carré intégrable, et indépendante du MB W. On considère la filtration définie, pour tout t positif, par $\mathcal{F}_t = \sigma\big\{\sigma\{Z,\ W_s; s \leq t\} \cup \mathcal{N}\big\}$.

Soit T un réel strictement positif. On considère deux fonctions $b : [0,T] \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ et $\sigma : [0,T] \times \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times d}$ qui sont mesurables. On cherche à résoudre l'équation différentielle stochastique :

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad \text{avec}, \quad X_0 = Z;$$

comme nous allons le voir par la suite cette équation est à interpréter au sens d'une équation intégrale, à savoir :

$$X_t = Z + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r, \quad 0 \leq t \leq T. \tag{1}$$

Le coefficient b s'appelle la dérive tandis que la matrice $\sigma\sigma^t$ s'appelle la matrice de diffusion.

Précisons tout d'abord ce que nous entendons par une solution de l'EDS (1).

Définition 1. Une solution (forte) de l'EDS (1), X, est un processus continu tel que :

1. X est progressivement mesurable ;
2. $\mathbb{P}$–p.s. $\displaystyle\int_0^T \{|b(r,X_r)| + \|\sigma(r,X_r)\|^2\}dr < \infty$, où $\|\sigma\| = \mathrm{trace}(\sigma\sigma^*)$;
3. $\mathbb{P}$–p.s., on a : $\displaystyle X_t = Z + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r, \quad 0 \leq t \leq T.$

On notera $\mathcal{S}^2$ l'espace de Banach constitué des processus X, progressivement mesurables, tels que $\mathbb{E}\big[\sup_{0\leq t\leq T} |X_t|^2\big] < \infty$ muni de la norme $\|X\| := \mathbb{E}\big[\sup_{0\leq t\leq T} |X_t|^2\big]^{1/2}$, et $\mathcal{S}^2_c$ le sous espace de $\mathcal{S}^2$ formé des processus continus. Notez que deux processus indistinguables sont identifiés et par abus d'écriture l'espace quotient est noté de la même façon.

Nous finissons ce paragraphe en rappelant un résultat élémentaire assez utile pour la suite.

Lemme 2. Lemme de Gronwall. *Soit $g : [0,T] \longrightarrow \mathbb{R}$ une fonction continue telle que, pour tout t,*

$$g(t) \leq a + b\int_0^t g(s)ds, \qquad a \in \mathbb{R},\ b \geq 0.$$

Alors, pour tout t, $g(t) \leq a\exp(bt)$.

Remarque. Pour simplifier les calculs déjà nombreux, les démonstrations seront effectuées dans le cas $n = d = 1$.

1.2. Existence et unicité. Nous allons établir le résultat suivant, dû à Itô :

Théorème 3. *Soient b et σ deux fonctions boréliennes. On suppose qu'il existe une constante K telle que, pour tout $t \in [0,T]$, $x, y \in \mathbb{R}^n$,*

1. *condition de Lipschitz en espace, uniforme en temps :*

$$|b(t,x) - b(t,y)| + \|\sigma(t,x) - \sigma(t,y)\| \leq K|x-y|;$$

2. *croissance linéaire : $|b(t,x)| + \|\sigma(t,x)\| \leq K(1+|x|)$;*
3. *$\mathbb{E}[|Z|^2] < \infty$.*

Alors l'EDS (1) possède une unique solution (à l'indistinguabilité près). Cette solution appartient à $\mathcal{S}^2$ et donc à $\mathcal{S}^2_c$.

Démonstration. La preuve consiste à utiliser la méthode itérative de Picard comme dans le cas déterministe. Une démonstration s'appuyant sur un argument de point fixe est également possible.

Pour $X \in \mathcal{S}^2_c$, posons, pour tout $t \in [0,T]$,

$$\Phi(X)_t = Z + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r.$$

Le processus $\Phi(X)$ est bien défini et est continu si $X \in \mathcal{S}^2_c$.

Si X et Y sont deux éléments de $\mathcal{S}^2_c$, comme $(a+b)^2 \leq 2a^2 + 2b^2$, on a, pour tout $0 \leq t \leq u \leq T$,

$$\left|\Phi(X)_t - \Phi(Y)_t\right|^2 \leq 2 \sup_{0\leq t\leq u} \left| \int_0^t \big(b(r,X_r) - b(r,Y_r)\big)dr\right|^2 + 2 \sup_{0\leq t\leq u} \left| \int_0^t \big(\sigma(r,X_r) - \sigma(r,Y_r)\big)dW_r\right|^2.$$

Utilisant les propriétés de l'intégrale stochastique, il vient :

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\leq t\leq u}\left|\Phi(X)_t - \Phi(Y)_t\right|^2\Big] \quad &\leq \quad 2\mathbb{E}\left[\Big(\int_0^u \left|b(r,X_r) - b(r,Y_r)\right|dr\Big)^2\right] \\ &\quad +8\mathbb{E}\left[\int_0^u \left\|\sigma(r,X_r) - \sigma(r,Y_r)\right\|^2 dr\right].\end{aligned}$$

L'inégalité de Hölder donne alors la majoration

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\leq t\leq u}\left|\Phi(X)_t - \Phi(Y)_t\right|^2\Big] \quad &\leq \quad 2T\mathbb{E}\left[\int_0^u \left|b(r,X_r) - b(r,Y_r)\right|^2 dr\right] \\ &\quad +8\mathbb{E}\left[\int_0^u \left\|\sigma(r,X_r) - \sigma(r,Y_r)\right\|^2 dr\right].\end{aligned}$$

Comme les fonctions b et σ sont Lipschitz en espace, on obtient, pour tout $u \in [0,T]$,

$$\mathbb{E}\Big[\sup_{0\leq t\leq u}\left|\Phi(X)_t - \Phi(Y)_t\right|^2\Big] \leq 2K^2(T+4)\mathbb{E}\Big[\int_0^u \sup_{0\leq t\leq r}|X_t - Y_t|^2 dr\Big]. \qquad (2)$$

De plus, notant 0 le processus nul, on a, comme $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$,

$$\left|\Phi(0)_t\right|^2 \leq 3Z^2 + 3\sup_{0\leq t\leq T}\left|\int_0^t b(r,0)dr\right|^2 + 3\sup_{0\leq t\leq T}\left|\int_0^t \sigma(r,0)dW_r\right|^2,$$

d'où l'on tire en utilisant l'inégalité de Doob et la croissance linéaire de b et σ,

$$\mathbb{E}\Big[\sup_{0\leq t\leq T}\big|\Phi(0)_t\big|^2\Big] \leq 3\Big(\mathbb{E}\big[Z^2\big]+K^2T^2+4K^2T\Big). \tag{3}$$

Les estimations (2) et (3) montrent alors que le processus $\Phi(X)$ appartient à $\mathcal{S}^2_{\mathrm{c}}$ dès que X appartient à $\mathcal{S}^2_{\mathrm{c}}$.

On définit alors par récurrence une suite de processus de $\mathcal{S}^2_{\mathrm{c}}$ en posant

$$X_0 = 0, \quad \text{ et, } \quad X^{n+1} = \Phi(X^n), \text{ pour } n \geq 0.$$

On obtient très facilement à l'aide de la formule (2), pour tout $n \geq 0$, notant C à la place de $2K^2(T+4)$,

$$\mathbb{E}\Big[\sup_{0\leq t\leq T}\big|X^{n+1}_t - X^n_t\big|^2\Big] \leq \frac{C^nT^n}{n!}\mathbb{E}\Big[\sup_{0\leq t\leq T}\big|X^1_t\big|^2\Big],$$

soit encore, notant D le majorant de l'inégalité (3),

$$\mathbb{E}\Big[\sup_{0\leq t\leq T}\big|X^{n+1}_t - X^n_t\big|^2\Big] \leq D\frac{C^nT^n}{n!}.$$

Il résulte de cette dernière inégalité que

$$\sum_{n\geq 0}\Big\|\sup_{0\leq t\leq T}\big|X^{n+1}_t - X^n_t\big|\Big\|_{\mathrm{L}^1} \leq \sum_{n\geq 0}\Big\|\sup_{0\leq t\leq T}\big|X^{n+1}_t - X^n_t\big|\Big\|_{\mathrm{L}^2} \leq \sqrt{D}\sum_{n\geq 0}\frac{(CT)^{n/2}}{\sqrt{n!}} < \infty.$$

Ainsi, la série $\sum_n \sup_t \big|X^{n+1}_t - X^n_t\big|$ converge $\mathbb{P}$–p.s. et donc, $\mathbb{P}$–p.s., X^n converge uniformément sur $[0,T]$ vers un processus X continu. De plus $X \in \mathcal{S}^2_{\mathrm{c}}$ puisque la convergence a lieu dans $\mathcal{S}^2$ – voir l'inégalité précédente. On vérifie très facilement que X est solution de l'EDS (1) en passant à la limite dans la définition $X^{n+1} = \Phi(X^n)$.

Si X et Y sont deux solutions de l'EDS (1) dans $\mathcal{S}^2_{\mathrm{c}}$ alors $X = \Phi(X)$ et $Y = \Phi(Y)$. L'inégalité (2) donne alors, pour tout $u \in [0,T]$,

$$\mathbb{E}\Big[\sup_{0\leq t\leq u}\big|X_t - Y_t\big|^2\Big] \leq 2K^2(T+4)\int_0^u \mathbb{E}\Big[\sup_{0\leq t\leq r}|X_t - Y_t|^2\Big]dr,$$

et le lemme de Gronwall montre que

$$\mathbb{E}\Big[\sup_{0\leq t\leq T}\big|X_t - Y_t\big|^2\Big] = 0,$$

ce qui prouve que X et Y sont indistinguables.

Pour montrer l'unicité des solutions de (1) au sens de la définition 2, nous devons montrer que toute solution appartient à $\mathcal{S}^2_{\mathrm{c}}$ c'est à dire, comme toute solution est continue par définition, appartient à $\mathcal{S}^2$.

Pour cela, considérons le temps d'arrêt $\tau_n = \inf\{t \in [0,T],\ |X_t| > n\}$ avec la convention $\inf\emptyset = +\infty$. Si $u \in [0,t]$, on a

$$\big|X_{u\wedge\tau_n}\big|^2 \leq 3\bigg(|Z|^2 + \sup_{0\leq u\leq t}\Big|\int_0^{u\wedge\tau_n} b(r,X_r)dr\Big|^2 + \sup_{0\leq u\leq t}\Big|\int_0^{u\wedge\tau_n}\sigma(r,X_r)dW_r\Big|^2\bigg).$$

Il vient alors,

$$\mathbb{E}\Big[\sup_{0\leq u\leq t\wedge\tau_n}\big|X_u\big|^2\Big] \leq 3\bigg(\mathbb{E}\big[|Z|^2\big] + \mathbb{E}\Big[\Big(\int_0^{t\wedge\tau_n}\big|b(r,X_r)\big|dr\Big)^2\Big] + 4\mathbb{E}\Big[\int_0^{t\wedge\tau_n}\big\|\sigma(r,X_r)\big\|^2dr\Big]\bigg),$$

et utilisant la croissance linéaire de b et σ, on obtient :

$$\mathbb{E}\Big[\sup_{0\le u\le t\wedge\tau_n}\big|X_u\big|^2\Big] \le 3\Big(\mathbb{E}\big[|Z|^2\big] + 2K^2T^2 + 8K^2T + (2K^2T+8K^2)\int_0^t \mathbb{E}\Big[\sup_{0\le u\le r\wedge\tau_n}\big|X_u\big|^2\Big]\Big).$$

On obtient, en appliquant le lemme de Gronwall, à la fonction $t \longmapsto \mathbb{E}\Big[\sup_{0\le u\le t\wedge\tau_n}\big|X_u\big|^2\Big]$

$$\mathbb{E}\Big[\sup_{0\le u\le T\wedge\tau_n}\big|X_u\big|^2\Big] \le 3\Big(\mathbb{E}\big[|Z|^2\big] + 2K^2T^2 + 8K^2T\Big)\exp\{3(2K^2T+8K^2)T\},$$

et le lemme de Fatou donne

$$\mathbb{E}\Big[\sup_{0\le u\le T}\big|X_u\big|^2\Big] < 3\Big(\mathbb{E}\big[|Z|^2\big] + 2K^2T^2 + 8K^2T\Big)\exp\{3(2K^2T+8K^2)T\}.$$

Ceci implique l'unicité des solutions de l'EDS (1).

Cette remarque termine la preuve. □

1.3. Exemples. Donnons deux exemples classiques d'EDS. Le premier est emprunté au monde de la finance. Le prix d'une action est généralement modélisé par l'EDS :

$$dS_t = S_t\big(\mu dt + \sigma dW_t\big), \qquad S_0 \text{ donné};$$

le paramètre σ s'appelle la volatilité et est très important. On montre facilement à l'aide de la formule d'Itô que

$$S_t = S_0 \exp\big\{(\mu - \sigma^2/2)t + \sigma W_t\big\}.$$

Considérons à présent l'équation de Langevin :

$$dX_t = -cX_t dt + \sigma dW_t, \qquad X_0 = x.$$

L'unique solution X de cette EDS s'appelle le processus d'Ornstein–Ulhenbeck. Posons $Y_t = X_t e^{ct}$. La formule d'intégration par parties donne

$$dY_t = e^{ct}dX_t + ce^{ct}X_t dt = e^{ct}\sigma dW_t.$$

On a alors $X_t = e^{-ct}x + \int_0^t e^{c(s-t)}\sigma dW_s$. On peut montrer que le processus X est un processus gaussien et donc en particulier que X_t est une variable aléatoire gaussienne. Calculons sa moyenne et sa variance. On a

$$\mathbb{E}[X_t] = e^{-ct}x, \quad \text{et,} \quad \mathrm{Var}(X_t) = \mathbb{E}\Big[\Big(\int_0^t e^{c(s-t)}\sigma dW_s\Big)^2\Big] = \sigma^2 e^{-2ct}\int_0^t e^{2cs}ds = \sigma^2\frac{1-e^{-2ct}}{2c}.$$

2. Propriétés élémentaires du flot

On va travailler ici avec des conditions initiales déterministes ce qui permet de prendre comme filtration la filtration naturelle du mouvement brownien $\{\mathcal{F}_t^W\}_{t\ge0}$. On suppose que les fonctions b et σ vérifie les hypothèses du théorème 3 – Lipschitz en espace et croissance linéaire. D'après le résultat précédent, on peut construire pour tout $(s,x)\in[0,T]\times\mathbb{R}^n$, la solution de l'EDS

$$X_t^{s,x} = x + \int_s^t b(r, X_r^{s,x})dr + \int_s^t \sigma(r, X_r^{s,x})dW_r, \quad s\le t\le T, \tag{4}$$

et l'on conviendra que $X_t^{s,x} = x$ si $0\le t\le s$.

Dans le cas déterministe, si $\sigma = 0$, le flot de l'équation différentielle, noté $\varphi_t^{s,x}$ dans ce cas, possède de nombreuses propriétés ; en particulier :

1. $\varphi_t^{s,x}$ est Lipschitz en (s,x,t) ;
2. pour $r \le s \le t$, $\varphi_t^{r,x} = \varphi_t^{s,\varphi_s^{r,x}}$;
3. si $s \le t$, $x \longmapsto \varphi_t^{s,x}$ est un homéomorphisme de $\mathbb{R}^n$.

Dans le cas stochastique, $X_t^{s,x}$ possède aussi des propriétés du même type. Dans le paragraphe suivant, nous étudierons la continuité de $(s,x,t) \longmapsto X_t^{s,x}$, puis le paragraphe suivant sera consacré à la propriété de composition. Nous n'aborderons pas le point 3 ; je vous renvoie aux livres de H. Kunita [Kun84, Kun90].

2.1. Continuité. Pour démontrer les propriétés de continuité du flot, nous allons utiliser le critère de Kolmogorov (théorème 36). Pour cela, nous devons établir des estimations sur les moments de $X_t^{s,x}$. Les démonstrations sont un peu techniques mais ne sont pas difficiles : il s'agit souvent d'utiliser les inégalités de Hölder, de Burkholder-Davis-Gundy – BDG dans la suite – et le lemme de Gronwall.

Proposition 4. *Soit $p \ge 1$. Il existe une constante C, dépendant de T et p, telle que :*

$$\forall s \in [0,T], \quad \forall x \in \mathbb{R}^n, \qquad \mathbb{E}\Big[\sup_{0\le t\le T} \big|X_t^{s,x}\big|^p\Big] \le C\Big(1+|x|^p\Big) \tag{5}$$

Démonstration. On fait la preuve dans le cas $n = d = 1$.

Nous commençons par le cas $p \ge 2$. Fixons s et x. Nous notons X_t à la place de $X_t^{s,x}$ pour alléger l'écriture. Dans ce qui suit C est une constante dépendant de p et T dont la valeur peut changer d'une ligne sur l'autre mais qui ne dépend pas de (s,x).

On a, tout d'abord,

$$\sup_{t\in[0,T]} |X_t|^p \le \sup_{t\in[0,s]} |X_t|^p + \sup_{t\in[s,T]} |X_t|^p \le |x|^p + \sup_{t\in[s,T]} |X_t|^p;$$

il suffit donc d'établir l'inégalité pour $\mathbb{E}\big[\sup_{t\in[s,T]} |X_t|^p\big]$.

Comme nous ne savons pas à priori si cette quantité est finie ou non, on introduit le temps d'arrêt $\tau_n = \inf\{t \in [0,T],\ |X_t| > n\}$, et on prend $n > |x|$ de sorte que $\tau_n > s$.

L'inégalité $(a+b+c)^p \le 3^{p-1}(a^p+b^p+c^p)$ fournit l'estimation, pour tout $u \in [s,T]$,

$$\begin{aligned}\big|X_{u\wedge\tau_n}\big|^p &\le 3^{p-1}\bigg(|x|^p + \sup_{s\le u\le t}\Big|\int_s^{u\wedge\tau_n} b(r,X_r)dr\Big|^p + \sup_{s\le u\le t}\Big|\int_s^{u\wedge\tau_n} \sigma(r,X_r)dW_r\Big|^p\bigg)\\ &\le 3^{p-1}\bigg(|x|^p + \Big(\int_s^{t\wedge\tau_n} \big|b(r,X_r)\big|dr\Big)^p + \sup_{s\le u\le t}\Big|\int_s^{u\wedge\tau_n} \sigma(r,X_r)dW_r\Big|^p\bigg).\end{aligned}$$

L'inégalité de BDG conduit à :

$$\mathbb{E}\Big[\sup_{s\le u\le t\wedge\tau_n} \big|X_u\big|^p\Big] \le C\left(|x|^p + \mathbb{E}\bigg[\Big(\int_s^{t\wedge\tau_n} \big|b(r,X_r)\big|dr\Big)^p\bigg] + \mathbb{E}\bigg[\Big(\int_s^{t\wedge\tau_n} \big|\sigma(r,X_r)\big|^2dr\Big)^{p/2}\bigg]\right),$$

et utilisant l'inégalité de Hölder ($p/2 \ge 1$), on a, notant p^* le conjugué de p et q celui de $p/2$,

$$\mathbb{E}\Big[\sup_{s\le u\le t\wedge\tau_n} \big|X_u\big|^p\Big] \le C\left(|x|^p + T^{p/p^*}\mathbb{E}\bigg[\int_s^{t\wedge\tau_n} \big|b(r,X_r)\big|^pdr\bigg] + T^{p/2q}\mathbb{E}\bigg[\int_s^{t\wedge\tau_n} \big|\sigma(r,X_r)\big|^pdr\bigg]\right).$$

De plus, comme b et σ sont à croissance linéaire, on a :

$$\mathbb{E}\bigg[\int_s^{t\wedge\tau_n} \big|b(r,X_r)\big|^pdr\bigg] \le K^p\mathbb{E}\bigg[\int_s^{t\wedge\tau_n} \big(1+|X_r|\big)^pdr\bigg] \le C\left(1+\mathbb{E}\bigg[\int_s^{t\wedge\tau_n} |X_r|^pdr\bigg]\right),$$

et donc
$$\mathbb{E}\left[\int_s^{t\wedge\tau_n} \big|b(r,X_r)\big|^p dr\right] \le C\left(1+\mathbb{E}\left[\int_s^t \sup_{s\le u\le r\wedge\tau_n} |X_u|^p dr\right]\right) \;;$$
et la même inégalité est valable pour le terme en σ.

Par suite, on obtient :
$$\mathbb{E}\left[\sup_{s\le u\le t\wedge\tau_n} \big|X_u\big|^p\right] \le C\left(1+|x|^p+\int_s^t \mathbb{E}\left[\sup_{s\le u\le r\wedge\tau_n} |X_u|^p\right] dr\right),$$
où C ne dépend pas de n. Le lemme de Gronwall donne alors, pour tout n,
$$\mathbb{E}\Big[\sup_{s\le u\le T\wedge\tau_n} \big|X_u\big|^p\Big] \le C\big(1+|x|^p\big).$$

On fait tendre n vers l'infini et on applique le lemme de Fatou pour avoir :
$$\mathbb{E}\Big[\sup_{s\le u\le T} \big|X_u\big|^p\Big] \le C\big(1+|x|^p\big),$$
ce qui termine la preuve dans le cas $p \ge 2$.

Si maintenant $1 \le p < 2$ alors $2p \ge 2$ et l'inégalité de Hölder donne
$$\mathbb{E}\Big[\sup_{s\le u\le T} \big|X_u\big|^p\Big] \le \left(\mathbb{E}\Big[\sup_{s\le u\le T} \big|X_u\big|^{2p}\Big]\right)^{1/2} \le C^{1/2}\big(1+|x|^{2p}\big)^{1/2},$$
ce qui conduit à, notant que $\sqrt{a+b} \le \sqrt{a}+\sqrt{b}$,
$$\mathbb{E}\Big[\sup_{s\le u\le T} \big|X_u\big|^p\Big] \le C^{1/2}\big(1+|x|^p\big).$$

Cette dernière inégalité termine la preuve de cette proposition. □

Nous savons à présent que la solution de l'EDS a des moments de tout ordre ; nous montrons une estimation du même type pour les moments des accroissements de X.

Proposition 5. *Soit $2 \le p < \infty$. Il existe une constante C telle que, pour tout $(s,x),(s',x')$ appartenant à $[0,T]\times\mathbb{R}^n$,*
$$\mathbb{E}\Big[\sup_{0\le t\le T} \big|X_t^{s,x}-X_t^{s',x'}\big|^p\Big] \le C\Big(|x-x'|^p+|s-s'|^{p/2}\big(1+|x'|^p\big)\Big). \tag{6}$$

Démonstration. Fixons (s,x) et (s',x'). Trivialement,
$$|X_t^{s,x}-X_t^{s',x'}|^p \le 2^{p-1}\big(|X_t^{s,x}-X_t^{s,x'}|^p+|X_t^{s,x'}-X_t^{s',x'}|^p\big)$$
de sorte que nous montrons l'inégalité pour chacun des deux termes précédents.

Commençons par le premier, $|X_t^{s,x}-X_t^{s,x'}|^p$. Il est inutile de prendre un temps d'arrêt car la proposition précédente nous dit que l'espérance du sup en t est fini. On a
$$\sup_{t\in[0,T]} |X_t^{s,x}-X_t^{s,x'}|^p \le \sup_{t\in[0,s]} |X_t^{s,x}-X_t^{s,x'}|^p + \sup_{t\in[s,T]} |X_t^{s,x}-X_t^{s,x'}|^p,$$
de sorte que
$$\sup_{t\in[0,T]} |X_t^{s,x}-X_t^{s,x'}|^p \le |x-x'|^p + \sup_{t\in[s,T]} |X_t^{s,x}-X_t^{s,x'}|^p;$$
par suite, on s'intéresse seulement au second membre de cette inégalité.

Pour tout $u \in [s,T]$, on a :

$$\begin{aligned}|X_u^{s,x} - X_u^{s,x'}|^p &\leq 3^{p-1}\Big(|x-x'|^p + \Big(\int_s^t \big|b(r,X_r^{s,x}) - b(r,X_r^{s,x'})\big|dr\Big)^p \\ &\qquad + \sup_{u\in[s,t]}\Big|\int_s^u \big(\sigma(r,X_r^{s,x}) - \sigma(r,X_r^{s,x'})\big)dW_r\Big|^p\Big).\end{aligned}$$

Les inégalités de BDG et de Hölder conduisent à l'inégalité, notant p^* le conjugué de p :

$$\begin{aligned}\mathbb{E}\Big[\sup_{u\in[s,t]}|X_u^{s,x} - X_u^{s,x'}|^p\Big] &\leq C\Big(|x-x'|^p + T^{p/p*}\mathbb{E}\Big[\int_s^t \big|b(r,X_r^{s,x}) - b(r,X_r^{s,x'})\big|^p dr\Big] \\ &\qquad + \mathbb{E}\Big[\Big(\int_s^t \big|\sigma(r,X_r^{s,x}) - \sigma(r,X_r^{s,x'})\big|^2 dr\Big)^{p/2}\Big]\Big).\end{aligned}$$

Utilisant une fois encore l'inégalité de Hölder, on obtient, notant q le conjugué de $p/2$,

$$\mathbb{E}\Big[\Big(\int_s^t \big|\sigma(r,X_r^{s,x}) - \sigma(r,X_r^{s,x'})\big|^2 dr\Big)^{p/2}\Big] \leq T^{p/2q}\mathbb{E}\Big[\int_s^t \big|\sigma(r,X_r^{s,x}) - \sigma(r,X_r^{s,x'})\big|^p dr\Big];$$

b et σ étant Lipschitz, l'inégalité précédente donne

$$\mathbb{E}\Big[\sup_{u\in[s,t]}|X_u^{s,x} - X_u^{s,x'}|^p\Big] \leq C\Big(|x-x'|^p + \int_s^t \mathbb{E}\Big[\sup_{u\in[s,r]}|X_u^{s,x} - X_u^{s,x'}|^p\Big]dr\Big).$$

Le lemme de Gronwall donne alors – en changeant C –

$$\mathbb{E}\Big[\sup_{u\in[s,T]}|X_u^{s,x} - X_u^{s,x'}|^p\Big] \leq C|x-x'|^p.$$

Il reste à étudier le terme $\mathbb{E}\big[\sup_t |X_t^{s,x'} - X_t^{s',x'}|^p\big]$. On suppose, sans perte de généralité, que $s \leq s'$ et on coupe en trois morceaux i.e.

$$\sup_{t\in[0,T]}|X_t^{s,x'} - X_t^{s',x'}|^p \leq \sup_{t\in[0,s]}|X_t^{s,x'} - X_t^{s',x'}|^p + \sup_{t\in[s,s']}|X_t^{s,x'} - X_t^{s',x'}|^p + \sup_{t\in[s',T]}|X_t^{s,x'} - X_t^{s',x'}|^p,$$

d'où l'on déduit que

$$\sup_{t\in[0,T]}|X_t^{s,x'} - X_t^{s',x'}|^p \leq \sup_{t\in[s,s']}|X_t^{s,x'} - x'|^p + \sup_{t\in[s',T]}|X_t^{s,x'} - X_t^{s',x'}|^p.$$

Pour le premier terme du membre de droite de l'inégalité précédente, on a,

$$\mathbb{E}\Big[\sup_{t\in[s,s']}|X_t^{s,x'} - x'|^p\Big] \leq 2^{p-1}\Big(\mathbb{E}\Big[\Big(\int_s^{s'} \big|b(r,X_r^{s,x'})\big|dr\Big)^p\Big] + \mathbb{E}\Big[\sup_{t\in[s,s']}\Big|\int_s^t \sigma(r,X_r^{s,x'})dW_r\Big|^p\Big]\Big).$$

L'inégalité de Hölder et la majoration (5) (proposition 4) donnent, utilisant la croissance linéaire de b,

$$\mathbb{E}\Big[\Big(\int_s^{s'} \big|b(r,X_r^{s,x'})\big|dr\Big)^p\Big] \leq (s'-s)^p\mathbb{E}\Big[\sup_{u\in[s,s']}\big|b(u,X_u^{s,x'})\big|^p\Big] \leq CT^{p/2}|s-s'|^{p/2}\big(1+|x'|^p\big).$$

D'autre part, l'inégalité de BDG, donne

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\in[s,s']}\Big|\int_s^t \sigma(r,X_r^{s,x'})dW_r\Big|^p\Big] &\leq \mathbb{E}\Big[\Big(\int_s^{s'} \big|\sigma(r,X_r^{s,x'})\big|^2 dr\Big)^{p/2}\Big] \\ &\leq (s'-s)^{p/2}\mathbb{E}\Big[\sup_{u\in[s,s']}\big|\sigma(u,X_u^{s,x'})\big|^p\Big],\end{aligned}$$

et, via la croissance de σ et l'estimation (5), on obtient

$$\mathbb{E}\Big[\sup_{t\in[s,s']}\Big|\int_s^t \sigma(r,X_r^{s,x'})dW_r\Big|^p\Big] \leq C|s-s'|^{p/2}\big(1+|x'|^p\big).$$

Finalement,

$$\mathbb{E}\Big[\sup_{t\in[s,s']}|X_t^{s,x'}-X_t^{s',x'}|^p\Big] \leq C|s-s'|^{p/2}\big(1+|x'|^p\big). \tag{7}$$

Étudions pour finir le terme $\mathbb{E}\big[\sup_{t\in[s',T]}|X_t^{s,x'}-X_t^{s',x'}|^p\big]$. Remarquons que, pour $t\in[s',T]$,

$$\begin{aligned} X_t^{s,x'} &= X_{s'}^{s,x'} + \int_{s'}^t b(r,X_r^{s,x'})dr + \int_{s'}^t \sigma(r,X_r^{s,x'})dW_r, \\ X_t^{s',x'} &= x' + \int_{s'}^t b(r,X_r^{s',x'})dr + \int_{s'}^t \sigma(r,X_r^{s',x'})dW_r. \end{aligned}$$

On a donc, pour tout $u\in[s',t]$,

$$\begin{aligned} \big|X_u^{s,x'}-X_u^{s',x'}\big|^p \leq\ & 3^{p-1}\Bigg(\big|X_{s'}^{s,x'}-x'\big|^p + \Big(\int_{s'}^t \big|b(r,X_r^{s,x'})-b(r,X_r^{s',x'})\big|dr\Big)^p \\ & + \sup_{u\in[s',t]}\Big|\int_{s'}^t \big(\sigma(r,X_r^{s,x'})-\sigma(r,X_r^{s',x'})\big)dW_r\Big|^p\Bigg), \end{aligned}$$

et un calcul désormais familier, utilisant les inégalités de Hölder, et de BDG, conduit à, via la majoration (7) et le fait que b et σ sont Lipschitz,

$$\begin{aligned} \mathbb{E}\Big[\sup_{u\in[s',t]}\big|X_u^{s,x'}-X_u^{s',x'}\big|^p\Big] &\leq C\Bigg(|s-s'|^{p/2}\big(1+|x'|^p\big) + \mathbb{E}\Big[\int_{s'}^t \big|X_r^{s,x'}-X_r^{s',x'}\big|^p dr\Big]\Bigg) \\ &\leq C\Bigg(|s-s'|^{p/2}\big(1+|x'|^p\big) + \mathbb{E}\Big[\int_{s'}^t \sup_{u\in[s',r]}\big|X_u^{s,x'}-X_u^{s',x'}\big|^p dr\Big]\Bigg). \end{aligned}$$

Le lemme de Gronwall appliqué à $r\longmapsto \sup_{u\in[s',r]}\big|X_u^{s,x'}-X_u^{s',x'}\big|^p$ donne alors

$$\mathbb{E}\Big[\sup_{u\in[s',t]}\big|X_u^{s,x'}-X_u^{s',x'}\big|^p\Big] \leq C|s-s'|^{p/2}\big(1+|x'|^p\big),$$

ce qui termine la preuve. □

Corollaire 6. *Il existe une modification du processus X telle que l'application de $[0,T]\times\mathbb{R}^n$ à valeurs dans $\mathcal{S}^2_{\mathrm{c}}$, $(s,x)\longmapsto(t\mapsto X_t^{s,x})$ soit continue.*

En particulier, $(s,x,t)\longmapsto X_t^{s,x}$ est continue.

Démonstration. C'est une application directe de l'estimation précédente – valable pour tout $p\geq 2$ et du critère de Kolmogorov (théorème 36) vu au chapitre 1. □

Remarque. Il est important de remarquer que le corollaire précédent implique en particulier que, $\mathbb{P}$–p.s., pour tout (s,x,t), l'équation (4) est valable.

Nous terminons ce paragraphe en précisant la régularité du flot généré par l'EDS.

Proposition 7. *Soit $2\leq p<\infty$. Il existe une constante C telle que, pour tout (s,x,t), (s',x',t'),*

$$\mathbb{E}\Big[\big|X_t^{s,x}-X_{t'}^{s',x'}\big|^p\Big] \leq C\Big(|x-x'|^p + \big(1+|x'|^p\big)\big(|s-s'|^{p/2}+|t-t'|^{p/2}\big)\Big). \tag{8}$$

En particulier, les trajectoires $(s,x,t)\longmapsto X_t^{s,x}$ sont hölderiennes (localement en x) de paramètre β en s, α en x et β en t pour tout $\beta<1/2$ et $\alpha<1$.

Démonstration. La régularité des trajectoires résulte du critère de Kolmogorov et de l'estimation que nous allons montrer rapidement.

On a :

$$\left|X_t^{s,x} - X_{t'}^{s',x'}\right| \leq \left|X_t^{s,x} - X_t^{s',x'}\right| + \left|X_t^{s',x'} - X_{t'}^{s',x'}\right|.$$

Comme $(a+b)^p \leq 2^{p-1}(a^p+b^p)$, il suffit d'établir l'estimation pour le second terme du majorant – le premier relevant de l'estimation (6) de la proposition 5. On suppose alors que $t \leq t'$; trois cas se présentent : $t \leq t' \leq s'$ (trivial), $t \leq s' \leq t'$ et $s' \leq t \leq t'$. Notons que le second cas se ramène au troisième puisque, si $t \leq s' \leq t'$, $X_t^{s',x'} = X_{s'}^{s',x'}$ et $|t-t'| \geq |t'-s'|$. On suppose donc que $s' \leq t \leq t'$. On a alors

$$X_{t'}^{s',x'} - X_t^{s',x'} = \int_t^{t'} b(r, X_r^{s',x'})dr + \int_t^{t'} b(r, X_r^{s',x'})dW_r,$$

et on doit estimer

$$\mathbb{E}\Big[\Big(\int_t^{t'} \left|b(r, X_r^{s',x'})\right|dr\Big)^p\Big], \quad \text{et}, \quad \mathbb{E}\Big[\Big|\int_t^{t'} \sigma(r, X_r^{s',x'})dr\Big|^p\Big].$$

Or, d'après un calcul déjà effectué lors de la démonstration de la proposition 5 – voir page 7 – chacun des deux termes précédents est majoré par $C|t-t'|^{p/2}\big(1+|x'|^p\big)$ ce qui termine la démonstration. □

2.2. Propriété de Markov. Nous allons établir la propriété de Markov pour les solutions de l'EDS (4) comme conséquence de la propriété de flot.

Proposition 8. *Soit $x \in \mathbb{R}^n$; soient $0 \leq r \leq s \leq t$. On a*

$$X_t^{r,x} = X_t^{s,X_s^{r,x}}, \qquad \mathbb{P}\text{–}p.s.$$

Démonstration. C'est la remarque 2.1 qui fournit le résultat. En effet, $\mathbb{P}$–p.s.

$$\forall s, y, t, \qquad X_t^{s,y} = y + \int_s^t b(u, X_u^{s,y})du + \int_s^t \sigma(u, X_u^{s,y})dW_u.$$

Il vient alors, $\mathbb{P}$–p.s.,

$$\forall t, \qquad X_t^{s,X_s^{r,x}} = X_s^{r,x} + \int_s^t b\big(u, X_u^{s,X_s^{r,x}}\big)du + \int_s^t \sigma\big(u, X_u^{s,X_s^{r,x}}\big)dW_u.$$

On remarque que $X_\cdot^{r,x}$ est aussi solution de cette dernière EDS sur $[s,T]$ puisque

$$\begin{aligned} X_t^{r,x} &= x + \int_r^t b(u, X_u^{r,x})du + \int_r^t \sigma(u, X_u^{r,x})dW_u \\ &= X_s^{r,x} + \int_s^t b(u, X_u^{r,x})du + \int_s^t \sigma(u, X_u^{r,x})dW_u. \end{aligned}$$

L'unicité des solutions d'une EDS à coefficients Lipschitz donne alors – en fait à l'indistinguabilité près sur $[s,T]$ – $X_t^{r,x} = X_t^{s,X_s^{r,x}}$. □

Remarque. En fait, par continuité, l'égalité $X_t^{r,x} = X_t^{s,X_s^{r,x}}$ a lieu pour tout x et pour tout $0 \leq r \leq s \leq t \leq T$ en dehors d'un ensemble $\mathbb{P}$–négligeable.

Nous allons à présent établir la propriété de Markov. Rappelons qu'un processus X est markovien s'il ne dépend du passé que par l'intermédiaire du présent. Mathématiquement cela conduit à la définition suivante :

Définition 9. Propriété de Markov. Soit X un processus progressivement mesurable par rapport à la filtration $\{\mathcal{F}_t\}_{t\geq 0}$. On dit que X possède la propriété de Markov par rapport à $\{\mathcal{F}_t\}_{t\geq 0}$ si, pour toute fonction borélienne bornée f, et pour tout $s \leq t$, on a

$$\mathbb{E}\big(f(X_t) \mid \mathcal{F}_s\big) = \mathbb{E}\big(f(X_t) \mid X_s\big), \quad \mathbb{P}\text{–}p.s.$$

Montrons la propriété de Markov pour les solutions de l'EDS (4) par rapport à la tribu du mouvement brownien W.

Théorème 10. *Soient $x \in \mathbb{R}^n$ et $s \in [0,T]$. $\big(X_t^{s,x}\big)_{0\leq t\leq T}$ est un processus de Markov par rapport à la filtration du MB et si f est mesurable et bornée alors, pour $s \leq r \leq t$,*

$$\mathbb{E}\big(f(X_t^{s,x}) \mid \mathcal{F}_r\big) = \Lambda(X_r^{s,x}) \quad \mathbb{P}\text{–}p.s.$$

avec $\Lambda(y) = \mathbb{E}\big[f(X_t^{r,y})\big]$.

Démonstration. Soient $0 \leq s \leq r \leq t \leq T$. D'après la propriété du flot, on a $X_t^{s,x} = X_t^{r,X_r^{s,x}}$. De plus, nous démontrerons au chapitre 5 – ou alors voir par exemple [Fri75] – que $X_t^{r,y}$ est mesurable par rapport à la tribu des accroissements du mouvement brownien $\sigma\{W_{r+u} - W_r,\ u \in [0, t-r]\}$. Donc,

$$X_t^{r,y} = \Phi\big(y, W_{r+u} - W_r; 0 \leq u \leq t-r\big),$$

où Φ est mesurable. Par suite,

$$X_t^{s,x} = X_t^{r,X_r^{s,x}} = \Phi\big(X_r^{s,x}, W_{r+u} - W_r; 0 \leq u \leq t-r\big).$$

Pour conclure, il faut encore remarquer que $X_r^{s,x}$ est $\mathcal{F}_r$–mesurable et noter que $\mathcal{F}_r$ est indépendante de la tribu $\sigma\{W_{r+u} - W_r,\ u \in [0, t-r]\}$. On a alors

$$\begin{aligned}
\mathbb{E}\big(f(X_t^{s,x}) \mid \mathcal{F}_r\big) &= \mathbb{E}\Big(f\big\{\Phi\big(X_r^{s,x}, W_{r+u} - W_r; 0 \leq u \leq t-r\big)\big\} \mid \mathcal{F}_r\Big) \\
&= \mathbb{E}\Big[f\big\{\Phi\big(y, W_{r+u} - W_r; 0 \leq u \leq t-r\big)\big\}\Big]_{|y = X_r^{s,x}} \\
&= \mathbb{E}\Big[f\big(X_t^{r,y}\big)\Big]_{|y = X_r^{s,x}} \\
&= \Lambda\big(X_r^{s,x}\big).
\end{aligned}$$

Ce qui montre la propriété annoncée. □

On peut également montrer que, si g est mesurable et, par exemple bornée,

$$\mathbb{E}\Big(\int_r^T g(u, X_u^{s,x})du \mid \mathcal{F}_r\Big) = \Gamma(r, X_r^{s,x}),$$

où $\Gamma(r,y) = \mathbb{E}\big[\int_r^T g(u, X_u^{r,y})du\big]$.

Remarque. Si b et σ ne dépendent pas du temps – on dit que l'EDS est autonome – on peut montrer que la loi de $X_t^{s,x}$ est la même que la loi de $X_{t-s}^{0,x}$. On a alors :

$$\mathbb{E}\big(f(X_t^{s,x}) \mid \mathcal{F}_r\big) = \mathbb{E}\big[f(X_{t-r}^{0,y})\big]_{|y = X_r^{s,x}}.$$

Références

[Fri75] A. Friedman, *Stochastic differential equations and applications I*, Probab. Math. Stat., vol. 28, Academic Press, New York, 1975.

[Kun84] H. Kunita, *Stochastic differential equations and stochastic flows of diffeomorphisms*, Ecole d'été de probabilités de Saint-Flour XII—1982 (P.-L. Hennequin, ed.), Lecture Notes in Math., vol. 1097, Springer-Verlag, Berlin Heidelberg New York, 1984, pp. 143–303.

[Kun90] ———, *Stochastic flows and stochastic differential equations*, Cambridge Stud. Adv. Math., vol. 24, Cambridge University Press, Cambridge, 1990.

Chapitre 5

Le cadre markovien

Nous considérons ici des EDSR « markoviennes » : il s'agit d'un cadre très spécifique dans lequel la dépendance de la condition terminale et du générateur de l'EDSR dans l'aléat ω se fait au travers de la solution d'une EDS. Nous verrons que la propriété de Markov pour les EDS se transfère aux EDSR ; d'où le nom du chapitre.

1. Le modèle

1.1. Hypothèses et notations. La donnée de base est toujours un espace probabilisé complet $(\Omega, \mathcal{F}, \mathbb{P})$ sur lequel est défini un mouvement brownien W d-dimensionnel. La filtration $\{\mathcal{F}_t\}_{t\geq 0}$ est la filtration naturelle de W augmentée de sorte que les conditions habituelles sont satisfaites.

On considère deux fonctions $b : [0,T] \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ et $\sigma : [0,T] \times \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times d}$ continues. On suppose qu'il existe une constante $K \geq 0$ telle que, pour tout t, pour tout x, x' de $\mathbb{R}^n$,

1. $\big|b(t,x) - b(t,x')\big| + \big\|\sigma(t,x) - \sigma(t,x')\big\| \leq K|x - x'|$;
2. $\big|b(t,x)\big| + \big\|\sigma(t,x)\big\| \leq K(1 + |x|)$.

Sous ces hypothèses, on peut construire, étant donnés un réel $t \in [0,T]$ et une variable aléatoire $\theta \in \mathrm{L}^2(\mathcal{F}_t)$, $\{X_u^{t,\theta}\}_{t\leq u\leq T}$ la solution de l'EDS

$$X_u^{t,\theta} = \theta + \int_t^u b\big(r, X_r^{t,\theta}\big)dr + \int_t^u \sigma\big(r, X_r^{t,\theta}\big)dW_r, \qquad t \leq u \leq T \ ; \tag{1}$$

on convient, que si $0 \leq u \leq t$, $X_u^{t,\theta} = \mathbb{E}(\theta \,|\, \mathcal{F}_u)$. Les propriétés de $\{X_u^{t,\theta}\}_{t\leq u\leq T}$ ont été étudiées au Chapitre 4 surtout dans le cas $\theta = x \in \mathbb{R}^n$.

Considérons aussi deux fonctions continues $g : \mathbb{R}^k \longrightarrow \mathbb{R}^k$ et $f : [0,T]\times\mathbb{R}^n\times\mathbb{R}^k\times\mathbb{R}^{k\times d} \longrightarrow \mathbb{R}^k$. Nous supposons que f et g vérifient les hypothèses suivantes : il existe deux réels μ et $p \geq 1$ tels que, pour tout (t,x,y,y',z,z'),

1. $(y - y')\cdot\big(f(t,x,y,z) - f(t,x,y',z)\big) \leq \mu|y - y'|^2$;
2. $\big|f(t,x,y,z) - f(t,x,y,z')\big| \leq K\|z - z'\|$;
3. $\big|g(x)\big| + \big|f(t,x,y,z)\big| \leq K\big(1 + |x|^p + |y| + \|z\|\big)$.

Sous ces hypothèses, si θ appartient à $\mathrm{L}^{2p}(\mathcal{F}_t)$, on peut résoudre l'EDSR

$$Y_u^{t,\theta} = g\big(X_T^{t,\theta}\big) + \int_u^T f\big(r, X_r^{t,\theta}, Y_r^{t,\theta}, Z_r^{t,\theta}\big)dr - \int_u^T Z_r^{t,\theta}dW_r, \qquad 0 \leq u \leq T. \tag{2}$$

Dans tout ce chapitre, nous supposerons que les hypothèses précédentes sur les coefficients b, σ, f et g sont satisfaites. Parfois, nous ferons une hypothèse plus forte sur f et g en supposant que g est K–Lipschitz et que f est K–Lipschitz en x uniformément en (t,y,z). Dans ce cas, $p = 1$. Nous désignerons cette hypothèse par (Lip).

1.2. Premières propriétés.

Rappels sur les EDS. Rappelons tout d'abord quelques propriétés des EDS que nous avons établies lors du Chapitre 4. Nous commençons par des estimations dans L^q avec $q \geq 2$.

La première chose à signaler est que le théorème d'Itô sur les EDS (Théorème 3) garantit l'existence et l'unicité de $\{X_u^{t,\theta}\}_{0\leq u\leq T}$. De plus, il existe une constante C telle que, pour tout $t \in [0,T]$ et θ, $\theta' \in \mathrm{L}^q(\mathcal{F}_t)$,

$$\mathbb{E}\left[\sup_{0\leq u\leq T}\left|X_u^{t,\theta}\right|^q\right] \leq C\left(1+\mathbb{E}\left[|\theta|^q\right]\right), \qquad \mathbb{E}\left[\sup_{0\leq u\leq T}\left|X_u^{t,\theta}-X_u^{t,\theta'}\right|^q\right] \leq C\,\mathbb{E}\left[|\theta-\theta'|^q\right].$$

Dans le même esprit, on a

$$\mathbb{E}\left[\sup_{0\leq u\leq T}\left|X_u^{t,x}-X_u^{t',x'}\right|^q\right] \leq C\left\{|x-x'|^q+|t-t'|^{q/2}\left(1+|x|^q\right)\right\}.$$

Notons si $x\in\mathbb{R}^n$ et $u \geq t$, $\Lambda(\omega,t,x,u) = X_u^{t,x}(\omega)$. On est naturellement amené à comparer deux objets : $\Lambda(\omega,t,\theta(\omega),u)$ d'une part et $X_u^{t,\theta}(\omega)$ d'autre part. Nous avons vu au Chapitre 4 que $\mathbb{P}$–p.s.,

$$\Lambda\big(\omega,t,\theta(\omega),u\big) = X_u^{t,\theta}(\omega), \qquad t\leq u\leq T.$$

L'EDSR (2). Nous donnons deux propriétés élémentaires de l'équation (2).

Proposition 1. *Si $\theta \in \mathrm{L}^{2p}(\mathcal{F}_t)$, l'EDSR (2) possède une unique solution (vérifiant $Z \in \mathrm{M}^2$), $\left\{(Y_u^{t,\theta}, Z_u^{t,\theta})\right\}_{0\leq u\leq T}$. Il existe une constante C telle que, pour tout t, pour tout $\theta\in \mathrm{L}^{2p}(\mathcal{F}_t)$,*

$$\mathbb{E}\left[\sup_{0\leq u\leq T}\left|Y_u^{t,\theta}\right|^2+\int_0^T\left\|Z_r^{t,\theta}\right\|^2dr\right] \leq C\left(1+\mathbb{E}\left[|\theta|^{2p}\right]\right).$$

Démonstration. Commençons par écrire l'équation (2) sous la forme d'une EDSR telle que nous les avons étudiées dans les deux chapitres précédents. Pour cela, on résout d'abord l'EDS (1), puis on définit

$$\bar{\xi} = g\big(X_T^{t,\theta}\big), \qquad \bar{f}(\omega,r,y,z) = f\big(r,X_r^{t,\theta}(\omega),y,z\big).$$

Avec ces notations, l'équation (2) n'est rien d'autre que l'EDSR

$$Y_u = \bar{\xi}+\int_u^T \bar{f}(r,Y_r,Z_r)dr - \int_t^T Z_r dW_r, \qquad 0\leq u\leq T.$$

Pour la première partie de la proposition, il suffit de montrer, vues les hypothèses du chapitre, que $\bar{\xi}$ est de carré intégrable et qu'il en va de même pour le processus $\{\bar{f}(r,0,0)\}_r$. Mais, d'après la croissance de g et f, on a

$$\forall r\in[0,T], \qquad \left|\bar{\xi}\right|+\left|\bar{f}(r,0,0)\right| \leq K\left(1+\sup_{0\leq u\leq T}\left|X_u^{t,\theta}\right|^p\right) \;;$$

le résultat s'en suit immédiatement puisque, comme $\theta\in\mathrm{L}^{2p}$, il en est de même de $\sup_{0\leq u\leq T}\left|X_u^{t,\theta}\right|$ d'après le rappel sur les EDS.

Pour établir la majoration, on utilise les estimations à priori du Chapitre 3 cf. Proposition 1, si $\alpha = 2\mu+2K^2$,

$$\mathbb{E}\left[\sup_{0\leq u\leq T}\left|Y_u^{t,\theta}\right|^2+\int_0^T\left\|Z_r^{t,\theta}\right\|^2dr\right] \leq C_u(1+T)e^{|\alpha|T}\,\mathbb{E}\left[\left|g\big(X_T^{t,\theta}\big)\right|^2+\int_0^T\left|f\big(r,X_r^{t,\theta},0,0\big)\right|^2dr\right],$$

ce qui compte tenu de la croissance de f et g donne

$$\mathbb{E}\left[\sup_{0\leq u\leq T}\left|Y_u^{t,\theta}\right|^2+\int_0^T\left\|Z_r^{t,\theta}\right\|^2dr\right] \leq C\left(1+\mathbb{E}\left[\sup_{0\leq u\leq T}\left|X_u^{t,\theta}\right|^{2p}\right]\right),$$

et par suite, pour une autre constante C,

$$\mathbb{E}\left[\sup_{0\le u\le T}\left|Y_u^{t,\theta}\right|^2+\int_0^T\left\|Z_r^{t,\theta}\right\|^2dr\right]\le C\left(1+\mathbb{E}\left[|\theta|^{2p}\right]\right),$$

ce que l'on voulait établir. □

Remarque. En particulier, lorsque θ est constante égale à x, on a

$$\mathbb{E}\left[\sup_{0\le u\le T}\left|Y_u^{t,x}\right|^2+\int_0^T\left\|Z_r^{t,x}\right\|^2dr\right]\le C\left(1+|x|^{2p}\right).$$

Proposition 2. *Si θ_n converge vers θ dans $\mathrm{L}^{2p}(\mathcal{F}_t)$, alors*

$$\mathbb{E}\left[\sup_{0\le u\le T}\left|Y_u^{t,\theta}-Y_u^{t,\theta_n}\right|^2+\int_0^T\left\|Z_r^{t,\theta}-Z_r^{t,\theta_n}\right\|^2dr\right]\longrightarrow 0,\qquad si\quad n\to\infty.$$

De plus, sous l'hypothèse (Lip), on a, si θ et θ' sont $\mathcal{F}_t$–mesurables et de carré intégrable,

$$\mathbb{E}\left[\sup_{0\le u\le T}\left|Y_u^{t,\theta}-Y_u^{t,\theta'}\right|^2+\int_0^T\left\|Z_r^{t,\theta}-Z_r^{t,\theta'}\right\|^2dr\right]\le C\,\mathbb{E}\left[|\theta-\theta'|^2\right].$$

Démonstration. Le point de départ est encore une fois les estimations à priori cf. Corollaire 2. On a donc, si θ et θ' sont deux éléments de $\mathrm{L}^{2p}(\mathcal{F}_t)$, notant $\alpha=2\mu+2K^2$, un contrôle de

$$\mathbb{E}\left[\sup_{0\le u\le T}\left|Y_u^{t,\theta}-Y_u^{t,\theta'}\right|^2+\int_0^T\left\|Z_r^{t,\theta}-Z_r^{t,\theta'}\right\|^2dr\right]$$

par la quantité, à la constante multiplicative $C_u(1+T)e^{|\alpha|T}$ près,

$$\mathbb{E}\left[\left|g\big(X_T^{t,\theta}\big)-g\big(X_T^{t,\theta'}\big)\right|^2+\int_0^T\left|f\big(r,X_r^{t,\theta},Y_r^{t,\theta},Z_r^{t,\theta}\big)-f\big(r,X_r^{t,\theta'},Y_r^{t,\theta},Z_r^{t,\theta}\big)\right|^2dr\right].$$

Sous l'hypothèse (Lip), cette dernière s'estime facilement ; on obtient comme majorant, C, C' désignant deux constantes dépendant de T, μ, K,

$$C'\,\mathbb{E}\left[\sup_{0\le r\le T}\left|X_r^{t,\theta}-X_r^{t,\theta'}\right|^2\right]\le C\,\mathbb{E}\left[|\theta-\theta'|^2\right],$$

d'après les résultats sur les EDS.

En l'absence de l'hypothèse (Lip), on doit montrer que l'on conserve la convergence vers 0 si on prend $\theta'=\theta_n$ avec $\theta_n\longrightarrow\theta$ dans L^{2p}. Les estimations sur les EDS donnent,

$$\mathbb{E}\left[\sup_{0\le r\le T}\left|X_r^{t,\theta}-X_r^{t,\theta_n}\right|^{2p}\right]\le C\,\mathbb{E}\left[|\theta-\theta_n|^{2p}\right]\longrightarrow 0\;;\tag{3}$$

ceci fournit la convergence en probabilité de $g\big(X_T^{t,\theta_n}\big)$ vers $g\big(X_T^{t,\theta}\big)$ par continuité de g ainsi que la convergence en $\mathbb{P}\otimes m$–mesure de

$$f\big(r,X_r^{t,\theta_n},Y_r^{t,\theta},Z_r^{t,\theta}\big)\quad\text{vers}\quad f\big(r,X_r^{t,\theta},Y_r^{t,\theta},Z_r^{t,\theta}\big)$$

vue la continuité de f. Il reste donc à établir l'équi-intégrabilité. Pour cela il suffit de noter que la croissance de f et g conduit à l'inégalité

$$\left|g\big(X_T^{t,\theta_n}\big)\right|+\left|f\big(r,X_r^{t,\theta_n},Y_r^{t,\theta},Z_r^{t,\theta}\big)\right|\le K\left(1+\sup_{0\le r\le T}\left|X_r^{t,\theta_n}\right|^p+\left|Y_r^{t,\theta}\right|+\left\|Z_r^{t,\theta}\right\|\right),$$

et le dernier majorant converge dans L^2 essentiellement en vertu de (3). Le résultat s'en suit directement. □

2. La propriété de Markov

Dans ce paragraphe, nous allons établir que la propriété de Markov des solutions d'EDS se transfère aux solutions d'EDSR. Cette propriété est importante pour les applications aux EDP. Nous commençons par montrer que $Y_t^{t,x}$ est une quantité déterministe. Introduisons une nouvelle notation : si $t \le u$, $\mathcal{F}_u^t = \sigma\big(\mathcal{N},\ W_r - W_t,\ t \le r \le u\big)$.

Proposition 3. *Soit* $(t,x) \in [0,T]\times\mathbb{R}^n$. $\{(X_u^{t,x}, Y_u^{t,x},)\}_{t\le u\le T}$ *est adapté par rapport à la filtration* $\{\mathcal{F}_u^t\}_{t\le u\le T}$. *En particulier,* $Y_t^{t,x}$ *est déterministe.*

On peut choisir une version de $\{Z_u^{t,x}\}_{t\le u\le T}$ *adapté par rapport à la filtration* $\{\mathcal{F}_u^t\}_{t\le u\le T}$.

Démonstration. Considérons le processus $\{B_u\}_{u\ge 0}$ définie par $B_u = W_{t+u} - W_t$ et notons $\{\mathcal{G}_u\}_u$ sa filtration naturelle augmentée. B est un mouvement brownien et on a de plus $\mathcal{G}_u = \mathcal{F}_{t+u}^t$.

Soit $\{X_u\}_{0\le u\le T-t}$ la solution $\{\mathcal{G}_u\}_u$–adaptée de l'EDS

$$X_u = x + \int_0^u b(t+r, X_r)dr + \int_0^u \sigma(t+r, X_r)dB_r, \qquad 0 \le u \le T-t.$$

En particulier, pour tout $v \in [t,T]$, on a

$$X_{v-t} = x + \int_0^{v-t} b(t+r, X_r)dr + \int_0^{v-t} \sigma(t+r, X_r)dB_r.$$

Dans les deux intégrales, on fait le changement de variables $s = r+t$ pour obtenir – le changement dans l'intégrale stochastique sera justifiée à la fin de la preuve –

$$\int_0^{v-t} b(t+r, X_r)dr = \int_t^v b(s, X_{s-t})ds, \qquad \int_0^{v-t} \sigma(t+r, X_r)dB_r = \int_t^v \sigma(s, X_{s-t})dW_s,$$

et par suite,

$$X_{v-t} = x + \int_t^v b(s, X_{s-t})ds + \int_t^v \sigma(s, X_{s-t})dW_s, \qquad t \le v \le T\ ;$$

on a, d'autre part, par définition du processus $X^{t,x}$,

$$X_v^{t,x} = x + \int_t^v b(s, X_s^{t,x})ds + \int_t^v \sigma(s, X_s^{t,x})dW_s, \qquad t \le v \le T.$$

L'EDS de coefficients b, σ et de condition initiale x en t possède deux solutions : X_{v-t} et $X_v^{t,x}$. Par unicité des solutions d'EDS à coefficients Lipschitz, ces deux processus sont indistinguables : $\mathbb{P}$–p.s., $\forall v \in [t,T]$, $X_{v-t} = X_v^{t,x}$. En particulier, $X_v^{t,x}$ est mesurable par rapport à $\mathcal{G}_{v-t}$ puisqu'il en est ainsi pour X_{v-t}. Or $\mathcal{G}_{v-t} = \mathcal{F}_v^t$.

Le résultat pour les EDSR se déduit de celui que nous venons d'établir pour les EDS. En effet, on considère la solution $\{(Y_u, Z_u)\}_{0\le u\le T-t}$, $\mathcal{G}_u$–adaptée de l'EDSR

$$Y_u = g(X_{T-t}) + \int_u^{T-t} f(t+r, X_r, Y_r, Z_r)dr - \int_u^{T-t} Z_r dB_r, \qquad 0 \le u \le T-t,$$

soit encore

$$Y_{v-t} = g(X_{T-t}) + \int_{v-t}^{T-t} f(t+r, X_r, Y_r, Z_r)dr - \int_{v-t}^{T-t} Z_r dB_r, \qquad t \le v \le T.$$

On effectue le changement de variables $s = t + r$ dans les deux intégrales, pour obtenir

$$Y_{v-t} = g(X_{T-t}) + \int_v^T f(s, X_{s-t}, Y_{s-t}, Z_{s-t})ds - \int_v^T Z_{s-t}dW_s, \qquad t \le v \le T.$$

Il s'en suit que $\{Y_{v-t}, Z_{v-t}\}_{v\in[t,T]}$ est solution sur $[t,T]$ de l'EDSR (2) pour $\theta = x$ puisque nous savons déjà que $X_{v-t} = X_v^{t,x}$. L'unicité des solutions des EDSR donne, dans $\mathcal{S}_c^2 \times \mathrm{M}^2$, $\{Y_{v-t}, Z_{v-t}\}_{v\in[t,T]} = \{Y_v^{t,x}, Z_v^{t,x}\}_{v\in[t,T]}$ et la mesurabilité recherchée puisque $\{Y_{v-t}, Z_{v-t}\}_{v\in[t,T]}$ est adapté par rapport à $\mathcal{G}_{v-t} = \mathcal{F}_v^t$.

Pour être tout à fait complet, justifions le changement de variables dans les intégrales stochastiques i.e. si $\{h(r)\}_{0\le r\le T-t}$ est un processus de carré intégrable et $\mathcal{G}_r$–adapté,

$$\int_0^u h(r)dB_r = \int_t^{t+u} h(s-t)dW_s, \qquad 0 \le u \le T-t,$$

où rappelons-le $B_r = W_{t+r} - W_t$. Le résultat est immédiat si $h(r) = h_a \mathbf{1}_{]a,b]}(r)$ avec h_a une v.a. bornée, $\mathcal{G}_a$–mesurable. En effet, on a

$$\int_0^u h(r)dB_r = h_a\left(B_{u\wedge b} - B_{u\wedge a}\right) = h_a\left(W_{t+u\wedge b} - W_{t+u\wedge a}\right),$$

et aussi, comme $\mathcal{G}_a \subset \mathcal{F}_{t+a}$,

$$\int_t^{u+t} h(s-t)dW_s = \int_t^{u+t} h_a \mathbf{1}_{[t+a,t+b[}(s)dW_s = h_a\left(W_{(t+u)\wedge(t+b)} - W_{(t+u)\wedge(t+a)}\right).$$

Par linéarité, le résultat est valable pour tous les processus simples. Si maintenant, h est un processus de carré intégrable et $\mathcal{G}_u$–adapté, il existe une suite processus simples h_n, $\mathcal{G}_u$–adaptés, tels que

$$\mathbb{E}\left[\int_0^{T-t} |h_n(r) - h(r)|^2\, dr\right] \longrightarrow 0.$$

On en déduit immédiatement que

$$\mathbb{E}\left[\int_t^{T} |h_n(s-t) - h(s-t)|^2\, ds\right] \longrightarrow 0.$$

Comme $\mathcal{G}_s \subset \mathcal{F}_{t+s}$, $\{h_n(s-t)\}_s$ et par suite $\{h(s-t)\}_s$ sont $\mathcal{F}_s$–progressivement mesurables. Pour finir, il suffit de noter que, d'une part

$$\mathbb{E}\left[\sup_{0\le u\le T-t}\left|\int_0^u \left(h_n(r) - h(r)\right) dB_r\right|^2\right] \le 4\,\mathbb{E}\left[\int_0^{T-t} |h_n(r) - h(r)|^2\, dr\right]$$

et que d'autre part

$$\mathbb{E}\left[\sup_{t\le u\le T}\left|\int_t^u \left(h_n(s-t) - h(s-t)\right) dW_s\right|^2\right] \le 4\,\mathbb{E}\left[\int_t^T |h_n(s-t) - h(s-t)|^2\, ds\right],$$

ce qui achève la démonstration de cette proposition. □

Puisque nous savons que $Y_t^{t,x}$ est une quantité déterministe, on peut définir une fonction en posant,

$$\forall (t,x) \in [0,T]\times\mathbb{R}^n, \qquad u(t,x) := Y_t^{t,x}.$$

Commençons par étudier la croissance et la continuité de cette fonction.

Proposition 4. *La fonction u est continue et à croissance polynomiale. On a*

$$\forall (t,x) \in [0,T]\times\mathbb{R}^n, \qquad |u(t,x)| \le C\left(1 + |x|^p\right).$$

Sous l'hypothèse (Lip), la fonction u vérifie de plus, pour tout (t,x), (t',x'),

$$|u(t,x) - u(t',x')| \le C\left(|x-x'| + |t-t'|^{1/2}\left(1+|x|\right)\right).$$

Démonstration. La croissance de la fonction u a déjà été notée lors de la Remarque 1.2. Montrons la continuité. C désigne une constante dépendant de K et T qui peut changer au cours du texte. Soient (t,x) et (t',x') deux points de $[0,T]\times\mathbb{R}^n$. On a

$$u(t',x')-u(t,x)=Y_{t'}^{t',x'}-Y_t^{t,x}=\mathbb{E}\left[Y_{t'}^{t',x'}-Y_t^{t,x}\right]=\mathbb{E}\left[Y_{t'}^{t',x'}-Y_{t'}^{t,x}\right]+\mathbb{E}\left[Y_{t'}^{t,x}-Y_t^{t,x}\right]\ ;$$

si $t\geq t'$ – on procède de même si $t<t'$ –,

$$Y_{t'}^{t,x}=Y_t^{t,x}+\int_{t'}^t f\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)\,dr-\int_{t'}^t Z_r^{t,x}\,dW_r,$$

et par suite, on obtient, à l'aide de l'inégalité de Hölder,

$$\begin{aligned}\left|\mathbb{E}\left[Y_{t'}^{t,x}-Y_t^{t,x}\right]\right|^2 &= \left|\mathbb{E}\left[\int_{t'}^t f\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)\,dr\right]\right|^2\\ &\leq |t-t'|\;\mathbb{E}\left[\int_0^T \left|f\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)\right|^2\,dr\right].\end{aligned}$$

Or, utilisant la croissance de f, on obtient,

$$\begin{aligned}\mathbb{E}\left[\int_0^T \left|f\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)\right|^2\,dr\right] &\leq C\,\mathbb{E}\left[1+\sup_{0\leq r\leq T}\left\{\left|X_r^{t,x}\right|^{2p}+\left|Y_r^{t,x}\right|^2\right\}+\int_0^T\left\|Z_r^{t,x}\right\|^2\,dr\right]\\ &\leq C\left(1+|x|^{2p}\right).\end{aligned}$$

D'autre part, on a, $\left|\mathbb{E}\left[Y_{t'}^{t',x'}-Y_{t'}^{t,x}\right]\right|^2\leq\mathbb{E}\left[\sup_{r\in[0,T]}\left|Y_r^{t',x'}-Y_r^{t,x}\right|^2\right]$, et les estimations à priori fournissent

$$C\,\mathbb{E}\left[\left|g\big(X_T^{t',x'}\big)-g\big(X_T^{t,x}\big)\right|^2+\int_0^T\left|f\big(r,X_r^{t',x'},Y_r^{t,x},Z_r^{t,x}\big)-f\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)\right|^2\,dr\right]$$

comme majorant. Notons $\Delta(t,x,t',x')$ cette quantité (sans la constante multiplicative). Nous venons de montrer que

$$|u(t',x')-u(t,x)|^2\leq C\left(\Delta(t,x,t',x')+|t-t'|\left(1+|x|^{2p}\right)\right).\tag{4}$$

Sous l'hypothèse (Lip), nous avons $p=1$ et, de plus,

$$\Delta(t,x,t',x')\leq C\,\mathbb{E}\left[\sup_{0\leq r\leq T}\left|X_r^{t',x'}-X_r^{t,x}\right|^2\right]\leq C\left(|x-x'|^2+|t-t'|\left(1+|x|^2\right)\right),$$

d'après la Proposition 5.

Dans le cas général, soit $\{(t_n,x_n)\}_n$ une suite de points de $[0,T]\times\mathbb{R}^n$ convergeant vers (t,x). On prend $(t',x')=(t_n,x_n)$ dans l'inégalité (4) ; il suffit de montrer que $\Delta_n=\Delta(t,x,t_n,x_n)$ tend vers 0 quand $n\to\infty$. Or, d'après la Proposition 5, on a, pour tout $q\geq 1$,

$$\sup_{0\leq r\leq T}\left|X_r^{t,x}-X_r^{t_n,x_n}\right|\longrightarrow 0,\qquad \text{dans } \mathrm{L}^q.$$

La continuité de g et f donnent alors $g\big(X_T^{t_n,x_n}\big)\longrightarrow g\big(X_T^{t,x}\big)$ en probabilité et

$$f\big(r,X_r^{t_n,x_n},Y_r^{t,x},Z_r^{t,x}\big)\longrightarrow f\big(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x}\big)\qquad \text{en } \mathbb{P}\otimes m\text{–mesure.}$$

Pour conclure, nous devons établir l'équi-intégrabilité des carrés. Mais cela résulte simplement de l'inégalité

$$\left|g\big(X_T^{t_n,x_n}\big)\right|^2+\left|f\big(r,X_r^{t_n,x_n},Y_r^{t,x},Z_r^{t,x}\big)\right|^2\leq C\left(1+\sup_{0\leq r\leq T}\left|X_r^{t_n,x_n}\right|^{2p}+\sup_{0\leq r\leq T}\left|Y_r^{t,x}\right|^2+\left\|Z_r^{t,x}\right\|^2\right)$$

puisque comme déjà dit $\sup_{0\leq r\leq T}\left|X_r^{t_n,x_n}\right|$ converge dans tous les L^q. □

À l'aide de cette fonction, nous pouvons établir la propriété de Markov pour les EDSR.

Théorème 5. *Soient* $t \in [0,T]$ *et* $\theta \in \mathrm{L}^{2p}(\mathcal{F}_t)$*. On a :*

$$\mathbb{P}\text{–}p.s. \qquad Y_t^{t,\theta} = u(t,\theta) = Y_t^{t,\cdot} \circ \theta.$$

Démonstration. Supposons tout d'abord que θ soit une variable aléatoire étagée : $\theta = \sum_{i=1}^{l} x_i \, \mathbf{1}_{A_i}$ où $(A_i)_{1\leq i\leq l}$ est une partition $\mathcal{F}_t$–mesurable de Ω et $x_i \in \mathbb{R}^n$ pour tout $i = 1,\ldots,l$. Notons $\big(X_r^i, Y_r^i, Z_r^i\big)_{0\leq r\leq T}$ à la place de $(X_r^{t,x_i}, Y_r^{t,x_i}, Z_r^{t,x_i})_{0\leq r\leq T}$. On a alors, pour $t \leq r \leq T$,

$$X_r^{t,\theta} = \sum\nolimits_i \mathbf{1}_{A_i} X_r^i, \quad Y_r^{t,\theta} = \sum\nolimits_i \mathbf{1}_{A_i} Y_r^i, \quad Z_r^{t,\theta} = \sum\nolimits_i \mathbf{1}_{A_i} Z_r^i.$$

En effet, par définition, pour tout i, si $r \geq t$,

$$X_r^i = x_i + \int_t^r b(u, X_u^i) du + \int_t^r \sigma(u, X_u^i) dW_u,$$

et, en multipliant par $\mathbf{1}_{A_i}$ et faisant la somme sur i, on obtient – les indicatrices rentrent à l'intérieur de l'intégrale stochastique puisque $A_i \in \mathcal{F}_t$ –

$$\sum\nolimits_i \mathbf{1}_{A_i} X_r^i = \theta + \int_t^r \sum\nolimits_i \mathbf{1}_{A_i} b(u, X_u^i) du + \int_t^r \sum\nolimits_i \mathbf{1}_{A_i} \sigma(u, X_u^i) dW_u \;;$$

or on a $\sum_i \mathbf{1}_{A_i} H(T_i) = H\left(\sum_i \mathbf{1}_{A_i} T_i\right)$. Il s'en suit que

$$\sum\nolimits_i \mathbf{1}_{A_i} X_r^i = \theta + \int_t^r b\Big(u, \sum\nolimits_i \mathbf{1}_{A_i} X_u^i\Big)\, du + \int_t^r \sigma\Big(u, \sum\nolimits_i \mathbf{1}_{A_i} X_u^i\Big)\, dW_u.$$

Par unicité des solutions d'EDS à coefficients Lipschitz, on a

$$\mathbb{P}\text{–}p.s., \quad \forall t \leq r \leq T, \qquad X_r^{t,\theta} = \sum\nolimits_i \mathbf{1}_{A_i} X_r^i.$$

De la même façon, on a, pour tout i, si $t \leq r \leq T$,

$$Y_r^i = g\big(X_T^i\big) + \int_r^T f\left(u, X_u^i, Y_u^i, Z_u^i\right) du - \int_r^T Z_u^i \, dW_u,$$

et multipliant par $\mathbf{1}_{A_i}$ et faisant la somme sur i, on montre, utilisant la relation déjà établie pour la diffusion que $\left(\sum_i \mathbf{1}_{A_i} Y_r^i, \sum_i \mathbf{1}_{A_i} Z_r^i\right)$ est solution sur $[t,T]$ de l'EDSR

$$Y_r' = g\big(X_T^{t,\theta}\big) + \int_r^T f\left(u, X_u^{t,\theta}, Y_u', Z_u'\right) du - \int_r^T Z_u' \, dW_u$$

dont $\left(Y_r^{t,\theta}, Z_r^{t,\theta}\right)$ est également solution par définition. L'unicité des solutions des EDSR donnent

$$\mathbb{P}\text{–}p.s., \quad \forall t \leq r \leq T, \quad Y_r^{t,\theta} = \sum\nolimits_i \mathbf{1}_{A_i} Y_r^i, \qquad Z_r^{t,\theta} = \sum\nolimits_i \mathbf{1}_{A_i} Z_r^i, \quad \mathbb{P}\otimes m\text{–p.p. sur } \Omega \times [t,T].$$

En particulier, pour $r = t$, on a, revenant aux notations de départ,

$$\mathbb{P}\text{–p.s.} \qquad Y_t^{t,\theta} = \sum\nolimits_i \mathbf{1}_{A_i} Y_t^i = \sum\nolimits_i \mathbf{1}_{A_i} Y_t^{t,x_i} = \sum\nolimits_i \mathbf{1}_{A_i} u(t,x_i) = u\left(t, \sum\nolimits_i \mathbf{1}_{A_i} x_i\right) = u(t,\theta).$$

Avant de traiter le cas général, examinons ce qui se passe sous l'hypothèse (Lip). Considérons une suite de v.a. θ_n, $\mathcal{F}_t$–étagées, qui converge vers θ dans $\mathrm{L}^2(\mathcal{F}_t)$. La Proposition 2 fournit l'estimation

$$\mathbb{E}\left[\left|Y_t^{t,\theta_n} - Y_t^{t,\theta}\right|^2\right] \leq C\,\mathbb{E}\left[|\theta_n - \theta|^2\right]$$

et, d'après la Proposition 4, $u(t,\cdot)$ est Lipschitzienne ce qui implique que

$$\mathbb{E}\left[|u(t,\theta_n)-u(t,\theta)|^2\right] \le C\,\mathbb{E}\left[|\theta_n-\theta|^2\right].$$

Or d'après l'étape précédente, pour tout entier n, $u(t,\theta_n)=Y_t^{t,\theta_n}$; on en déduit directement que $u(t,\theta)=Y_t^{t,\theta}$, $\mathbb{P}$–p.s.

Dans le cas général, considérons une suite de v.a. θ_n, $\mathcal{F}_t$–étagées, qui converge vers θ dans $\mathrm{L}^{2p}(\mathcal{F}_t)$. La Proposition 2 entraîne toujours que

$$Y_t^{t,\theta_n} \longrightarrow Y_t^{t,\theta}, \quad \text{dans } \mathrm{L}^2, \quad \text{si } n\to\infty,$$

et, de plus, la continuité de l'application $u(t,\cdot)$ implique, en particulier, la convergence de $u(t,\theta_n)$ vers $u(t,\theta)$ en probabilité. La Remarque 1.2 montre que $|u(t,x)|$ est majoré par $C(1+|x|^p)$. Par suite,

$$|u(t,\theta_n)|^2 \le C'\left(1+|\theta_n|^{2p}\right),$$

ce qui montre que $\left(|u(t,\theta_n)|^2\right)_n$ est une suite équi-intégrable ; donc $u(t,\theta_n)$ vers $u(t,\theta)$ en probabilité et dans L^2. Comme $u(t,\theta_n)=Y_t^{t,\theta_n}$, $\mathbb{P}$–p.s., on obtient $u(t,\theta)=Y_t^{t,\theta}$, $\mathbb{P}$–p.s. □

Corollaire 6. *Soient $t\in[0,T]$ et $\theta\in\mathrm{L}^{2p}(\mathcal{F}_t)$. Alors, on a, $\mathbb{P}$–p.s.,*

$$\forall s\in[t,T], \qquad Y_s^{t,\theta}=u\left(s,X_s^{t,\theta}\right).$$

Démonstration. Fixons $(t,\theta)\in[0,T]\times\mathrm{L}^{2p}(\mathcal{F}_t)$ et prenons $s\in[t,T]$. On a d'après le théorème précédent,

$$\mathbb{P}\text{–}p.s., \qquad Y_s^{s,X_s^{t,\theta}}=u(s,X_s^{t,\theta}).$$

Or $\left\{\left(Y_r^{s,X_s^{t,\theta}},Z_r^{s,X_s^{t,\theta}}\right)\right\}_r$ est solution sur $[s,T]$ de l'EDSR

$$Y_u=g\Big(X_T^{s,X_s^{t,\theta}}\Big)+\int_u^T f\Big(r,X_r^{s,X_s^{t,\theta}},Y_r,Z_r\Big)\,dr-\int_u^T Z_r\,dW_r, \qquad s\le u\le T.$$

Or, par unicité des solutions sur $[s,T]$ de l'EDS,

$$X_r=X_s^{t,\theta}+\int_s^r b(u,X_u)\,du+\int_s^r \sigma(u,X_u)\,dW_u$$

on obtient facilement

$$\mathbb{P}\text{–}p.s., \qquad \forall r\in[s,T], \qquad X_r^{s,X_s^{t,\theta}}=X_r^{t,\theta}.$$

Par conséquent, $\left\{\left(Y_r^{s,X_s^{t,\theta}},Z_r^{s,X_s^{t,\theta}}\right)\right\}_r$ et $\left\{\left(Y_r^{t,\theta},Z_r^{t,\theta}\right)\right\}_r$ sont deux solutions sur $[s,T]$ de l'EDSR

$$Y_u=g\big(X_T^{t,\theta}\big)+\int_u^T f\big(r,X_r^{t,\theta},Y_r,Z_r\big)\,dr-\int_u^T Z_r\,dW_r, \qquad s\le u\le T.$$

L'unicité des solutions de cette EDSR donne en particulier

$$\mathbb{P}\text{–}p.s., \qquad Y_s^{t,\theta}=Y_s^{s,X_s^{t,\theta}}=u(s,X_s^{t,\theta}).$$

Ceci est valable pour tout $s\in[t,T]$; comme les deux processus $\left\{Y_s^{t,\theta}\right\}_s$ et $\left\{u(s,X_s^{t,\theta})\right\}_s$ sont continus via la continuité de u, on obtient le résultat. □

Remarque. On utilise souvent ce résultat avec θ constante égale à x. Cette *propriété de Markov*, $Y_s^{t,x}=u(s,X_s^{t,x})$, joue un rôle important lorsque l'on tente de construire la solution d'une EDP à l'aide d'une EDSR : nous le verrons plus loin.

3. Formule de Feynman-Kac

Nous finissons ce chapitre en faisant un lien entre les EDSR et les EDP. Au paragraphe précédent, nous avons vu que $Y_r^{t,x} = u(r, X_r^{t,x})$ où u est une fonction déterministe ; nous allons voir que u est solution d'une équation aux dérivées partielles – EDP. Nous notons $\mathcal{L}$ l'opérateur différentiel du second ordre suivant : pour h régulière

$$\begin{aligned}\mathcal{L}h(t,x) &= \frac{1}{2}\mathrm{trace}\big(\sigma\sigma^*(t,x)\mathrm{D}^2h(t,x)\big) + b(t,x)\cdot\nabla h(t,x)\\ &= \frac{1}{2}\sum_{i,j=1}^{n}\big(\sigma\sigma^*\big)_{i,j}(t,x)\,\partial_{x_i,x_j}h(t,x) + \sum_{i=1}^{n} b_i(t,x)\,\partial_{x_i}h(t,x).\end{aligned}$$

En ce qui concerne les notations, si h est une fonction de t et x nous notons $\partial_t h$ ou h' la dérivée partielle en temps, ∇h le gradient en espace (vecteur colonne) et $\mathrm{D}h = (\nabla h)^*$, $\mathrm{D}^2 h$ la matrice des dérivées secondes.

Nous supposerons également que le processus Y est réel c'est à dire que $k = 1$. Z est donc une ligne de dimension d (celle du brownien). En résumé, $f : [0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^{1\times d}\longrightarrow\mathbb{R}$.

L'objectif de ce paragraphe est d'établir des relations entre $\{Y_r^{t,x}, Z_r^{t,x}\}_{0\le r\le T}$ solution de l'EDSR

$$Y_r^{t,x} = g\big(X_T^{t,x}\big) + \int_r^T f\big(u, X_u^{t,x}, Y_u^{t,x}, Z_u^{t,x}\big)\,du - \int_r^T Z_u^{t,x}\,dW_u, \qquad 0\le r\le T, \tag{5}$$

où $\{X_r^{t,x}\}_{0\le r\le T}$ est la solution de l'EDS – avec la convention $X_r^{t,x} = x$ si $0\le r\le t$ –

$$X_r^{t,x} = x + \int_t^r b\big(u, X_u^{t,x}\big)\,du + \int_t^r \sigma\big(u, X_u^{t,x}\big)\,dW_u, \qquad t\le r\le T, \tag{6}$$

d'une part et « la solution » v de l'EDP parabolique suivante :

$$\partial_t v(t,x) + \mathcal{L}v(t,x) + f\,(t,x,v(t,x),\mathrm{D}v(t,x)\sigma(t,x)) = 0, \quad (t,x)\in]0,T[\times\mathbb{R}^n, \qquad v(T,\cdot) = g. \tag{7}$$

Proposition 7. *Supposons que l'EDP (7) possède une solution v, de classe $\mathcal{C}^{1,2}\big([0,T]\times\mathbb{R}^n\big)$, telle que pour tout $(t,x)\in[0,T]\times\mathbb{R}^n$, $|\nabla v(t,x)|\le C\,(1+|x|^q)$.*

Alors, pour tout $(t,x)\in[0,T]\times\mathbb{R}^n$, la solution de l'EDSR (5), $\{Y_r^{t,x}, Z_r^{t,x}\}_{t\le r\le T}$, est donnée par le couple de processus $\big\{v\big(r,X_r^{t,x}\big), \mathrm{D}v\sigma\big(r,X_r^{t,x}\big)\big\}_{t\le r\le T}$. En particulier, on obtient la formule, $u(t,x) = Y_t^{t,x} = v(t,x)$.

Démonstration. La condition de croissance sur le gradient de v assure que $\big(\mathrm{D}v\sigma\big(r,X_r^{t,x}\big)\big)_r$ est un processus de carré intégrable. De plus, comme v est régulière, on obtient en appliquant la formule d'Itô,

$$dv\big(s,X_s^{t,x}\big) = v'\big(s,X_s^{t,x}\big)ds + \nabla v\big(s,X_s^{t,x}\big)\cdot dX_s^{t,x} + \frac{1}{2}\mathrm{trace}\big\{\sigma\sigma^*\big(s,X_s^{t,x}\big)\mathrm{D}^2 v\big(s,X_s^{t,x}\big)\big\}ds\ ;$$

soit encore, utilisant $dX_s^{t,x}$ pour $s\ge t$, et la définition de $\mathcal{L}$,

$$dv\big(s,X_s^{t,x}\big) = \big\{v'\big(s,X_s^{t,x}\big) + \mathcal{L}v\big(s,X_s^{t,x}\big)\big\}\,ds + \mathrm{D}v\sigma\big(s,X_s^{t,x}\big)dW_s.$$

Or v est solution de l'EDP (7), donc $v' + \mathcal{L}v = -f$; il vient alors

$$dv\big(s,X_s^{t,x}\big) = -f\big(s,X_s^{t,x}, v\big(s,X_s^{t,x}\big), \mathrm{D}v\sigma\big(s,X_s^{t,x}\big)\big)\,ds + \mathrm{D}v\sigma\big(s,X_s^{t,x}\big)dW_s.$$

On conclut en intégrant de r à T notant que $v\big(T,X_T^{t,x}\big) = g\big(X_T^{t,x}\big)$. □

Le résultat précédent donne une formule de représentation probabiliste pour la solution d'une EDP parabolique non-linéaire – on dit semi-linéaire car la non-linéarité n'est pas très forte –. Ce type de formules, connues sous le nom de formule de Feynman-Kac – suite aux travaux de Richard Feynman et Mark Kac – s'appliquent à l'origine à des problèmes linéaires. Nous l'obtenons comme corollaire :

Corollaire 8. *Prenons $f(t,x,y,z) = c(t,x)y + h(t,x)$, où c et h sont deux fonctions continues bornées. Sous les hypothèses précédentes, pour tout $(t,x) \in [0,T] \times \mathbb{R}^n$,*

$$v(t,x) = \mathbb{E}\left[g\big(X_T^{t,x}\big) \exp\left(\int_t^T c(r, X_r^{t,x})\, dr \right) + \int_t^T h(r, X_r^{t,x}) \exp\left(\int_t^r c(s, X_s^{t,x})\, ds \right) dr \right].$$

Démonstration. D'après la proposition précédente, nous savons que $v(t,x) = Y_t^{t,x}$. Mais lorsque la fonction f prend la forme $f(t,x,y,z) = c(t,x)y + h(t,x)$, l'EDSR que l'on doit résoudre est une EDSR linéaire. On utilise alors la formule vue au cours du Chapitre 2 pour conclure. □

Remarque. Dans les deux derniers énoncés, nous avons supposé l'existence d'une solution classique v et nous en avons déduit la solution de l'EDSR et la formule $v(t,x) = Y_t^{t,x}$. Si tous les coefficients sont réguliers, on peut montrer que u qui est définie par $u(t,x) = Y_t^{t,x}$ est une fonction régulière qui est solution de l'EDP (7).

La démarche que nous avons eue précédemment consiste à étudier l'EDP, puis en déduire les solutions de l'EDSR. Mais on peut également étudier l'EDSR et en déduire la construction de la solution de l'EDP sans supposer les coefficients réguliers. Pour cela nous utiliserons la notion de solutions de viscosité des EDP dont nous rappelons rapidement la définition. Je vous renvoie au livre de Guy Barles [Bar94] pour plus de détails.

Définition 9. Soit u une fonction continue sur $[0,T] \times \mathbb{R}^n$ vérifiant la condition $u(T,x) = g(x)$. u est une sous-solution (resp. sursolution) de viscosité de l'EDP (7) si, pour toute fonction φ, de classe $\mathcal{C}^{1,2}\big([0,T] \times \mathbb{R}^n\big)$, on a, en tout point $(t_0,x_0) \in]0,T[\times \mathbb{R}^n$ de maximum (resp. minimum) local de $u - \varphi$:

$$\partial_t\varphi(t_0,x_0) + \mathcal{L}\varphi(t_0,x_0) + f\left(t_0, x_0, u(t_0,x_0), \mathrm{D}\varphi\sigma(t_0,x_0)\right) \geq 0, \qquad (\text{resp. } \leq 0).$$

u est solution de viscosité si elle est à la fois sous-solution et sursolution de viscosité.

Théorème 10. *La fonction $u(t,x) := Y_t^{t,x}$ est solution de viscosité de l'EDP (7).*

Démonstration. Remarquons que la fonction u est continue et vérifie de plus $u(T,\cdot) = g$. Nous montrons seulement que u est sous-solution (la démonstration de u sursolution est identique).

Soit donc φ une fonction de classe $\mathcal{C}^{1,2}$ telle que $u - \varphi$ possède en (t_0,x_0) un maximum local ($0 < t_0 < T$). Quitte à remplacer φ par $\varphi - \varphi(t_0,x_0) + u(t_0,x_0)$ – opération qui n'affecte pas les dérivées de φ, on peut supposer que $\varphi(t_0,x_0) = u(t_0,x_0)$. Nous devons montrer que

$$\varphi'(t_0,x_0) + \mathcal{L}\varphi(t_0,x_0) + f\big(t_0, x_0, u(t_0,x_0), \mathrm{D}\varphi\sigma(t_0,x_0)\big) \geq 0.$$

Supposons le contraire i.e. il existe $\delta > 0$ tel que

$$\varphi'(t_0,x_0) + \mathcal{L}\varphi(t_0,x_0) + f\big(t_0, x_0, u(t_0,x_0), \mathrm{D}\varphi\sigma(t_0,x_0)\big) = -\delta < 0.$$

$u - \varphi$ possède un maximum local en (t_0,x_0) qui est nul donc par continuité, il existe $0 < \alpha \leq T - t_0$ tel que, si $t_0 \leq t \leq t_0 + \alpha$ et $|x - x_0| \leq \alpha$,

$$u(t,x) \leq \varphi(t,x), \qquad \text{et}, \qquad \varphi'(t,x) + \mathcal{L}\varphi(t,x) + f\big(t, x, u(t,x), \mathrm{D}\varphi\sigma(t,x)\big) \leq -\delta/2.$$

Considérons le temps d'arrêt $\tau = \inf\{u \geq t_0 \ ; \ |X_u^{t_0,x_0} - x_0| > \alpha\} \wedge t_0 + \alpha$. Comme X^{t_0,x_0} est un processus continu on a $|X_\tau^{t_0,x_0} - x_0| \leq \alpha$.

La formule d'Itô appliquée à $\varphi(r, X_r^{t_0,x_0})$ donne

$$d\varphi(r, X_r^{t_0,x_0}) = \left\{\varphi'(r, X_r^{t_0,x_0}) + \mathcal{L}\varphi(r, X_r^{t_0,x_0})\right\} dr + \mathrm{D}\varphi\sigma(r, X_r^{t_0,x_0}) dW_r,$$

et, en intégrant de $u \wedge \tau$ à $(t_0 + \alpha) \wedge \tau = \tau$ pour $t_0 \leq u \leq t_0 + \alpha$, on obtient

$$\varphi\big(u \wedge \tau, X_{u\wedge\tau}^{t_0,x_0}\big) = \varphi\big(\tau, X_\tau^{t_0,x_0}\big) - \int_{u\wedge\tau}^{\tau} \{\varphi' + \mathcal{L}\varphi\} \left(r, X_r^{t_0,x_0}\right) dr - \int_{u\wedge\tau}^{\tau} \mathrm{D}\varphi\sigma\big(r, X_r^{t_0,x_0}\big)\, dW_r \ ;$$

notant, pour $t_0 \leq u \leq t_0 + \alpha$, $Y'_u = \varphi\big(u \wedge \tau, X_{u\wedge\tau}^{t_0,x_0}\big)$ et $Z'_u = \mathbf{1}_{u\leq\tau} \mathrm{D}\varphi\sigma\big(r, X_r^{t_0,x_0}\big)$, l'égalité précédente prend la forme

$$Y'_u = \varphi\big(\tau, X_\tau^{t_0,x_0}\big) + \int_u^{t_0+\alpha} -\mathbf{1}_{r\leq\tau} \{\varphi' + \mathcal{L}\varphi\} \left(r, X_r^{t_0,x_0}\right) dr - \int_u^{t_0+\alpha} Z'_r\, dW_r, \qquad t_0 \leq u \leq t_0 + \alpha.$$

De la même façon, si on pose, pour $t_0 \leq u \leq t_0 + \alpha$, $Y_u = Y_{u\wedge\tau}^{t_0,x_0}$ et $Z_u = \mathbf{1}_{u\leq\tau} Z_u^{t_0,x_0}$, on a

$$Y_u = Y_{t_0+\alpha} + \int_u^{t_0+\alpha} \mathbf{1}_{r\leq\tau}\, f\big(r, X_r^{t_0,x_0}, Y_r, Z_r\big)\, dr - \int_u^{t_0+\alpha} Z_r\, dW_r, \qquad t_0 \leq u \leq t_0 + \alpha.$$

Or la propriété de Markov – Corollaire 6 et la remarque qui le suit – implique que, $\mathbb{P}$–p.s. pour tout $t_0 \leq r \leq t_0 + \alpha$, $Y_r^{t_0,x_0} = u(r, X_r^{t_0,x_0})$ d'où l'on déduit que $Y_{t_0+\alpha} = Y_\tau^{t_0,x_0} = u(\tau, X_\tau^{t_0,x_0})$. L'égalité précédente devient alors

$$Y_u = u\big(\tau, X_\tau^{t_0,x_0}\big) + \int_u^{t_0+\alpha} \mathbf{1}_{r\leq\tau}\, f\big(r, X_r^{t_0,x_0}, u\big(r, X_r^{t_0,x_0}\big), Z_r\big)\, dr - \int_u^{t_0+\alpha} Z_r\, dW_r, \quad t_0 \leq u \leq t_0 + \alpha.$$

Nous allons appliquer le théorème de comparaison à $(Y'_u, Z'_u)_u$ et $(Y_u, Z_u)_u$ qui sont solutions d'EDSR. Par définition de τ, on a $u\big(\tau, X_\tau^{t_0,x_0}\big) \leq \varphi\big(\tau, X_\tau^{t_0,x_0}\big)$ et

$$\begin{aligned} \mathbf{1}_{r\leq\tau}\, f\big(r, X_r^{t_0,x_0}, u(r, X_r^{t_0,x_0}), Z'_r\big) &= \mathbf{1}_{r\leq\tau}\, f\big(r, X_r^{t_0,x_0}, u(r, X_r^{t_0,x_0}), \mathrm{D}\varphi\sigma(r, X_r^{t_0,x_0})\big) \\ &\leq -\mathbf{1}_{r\leq\tau} \{\varphi' + \mathcal{L}\varphi\} \big(r, X_r^{t_0,x_0}\big). \end{aligned}$$

De plus, on a, toujours par définition de τ,

$$\mathbb{E}\left[\int_{t_0}^{t_0+\alpha} -\mathbf{1}_{r\leq\tau} \left(\varphi' + \mathcal{L}\varphi + f\right) \big(r, X_r^{t_0,x_0}, u(r, X_r^{t_0,x_0}), \mathrm{D}\varphi\sigma(r, X_r^{t_0,x_0})\big)\, dr\right] \geq \mathbb{E}\left[\tau - t_0\right] \delta/2.$$

Cette quantité est strictement positive : en effet, $\delta > 0$ et $\tau > t_0$ car $|X_{t_0}^{t_0,x_0} - x_0| = 0 < \alpha$. On peut donc appliquer la version stricte du théorème de comparaison, voir la remarque à la suite du Théorème 9, pour obtenir $u(t_0, x_0) = Y_{t_0} < Y'_{t_0} = \varphi(t_0, x_0)$. Ceci est impossible puisque $u(t_0, x_0) = \varphi(t_0, x_0)$. u est donc bien une sous-solution. $\square$

Références

[Bar94] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, Math. Appl., vol. 17, Springer-Verlag, Paris, 1994.