Stochastics, 1982, Vol. 6, pp. 193–231
0090-9491/82/0604-0193 $06.50/0

Printed in Great Britain

Équations du Filtrage non Linéaire de la Prédiction et du Lissage

E. PARDOUX
U.E.R. de Mathématiques, Université de Provence, 3, Place Victor Hugo, 13331 Marseille Cedex 3, France

(*Accepted for publication April 26, 1981*)

We establish equations of non linear filtering, prediction (extrapolation) and smoothing (interpolation) in the case where the signal is a non degenerate diffusion process, and the observation is a noisy functional of the signal. We consider both the case of observation noise correlated with the signal, and the opposite case where we establish "robust" form of the equations. We study finally the case of unbounded coefficients, and the case where there is a feedback from the observation to the signal.

1. INTRODUCTION

Soit X_t un processus markovien de diffusion. Supposons que l'on observe le processus:

$$Y_t = \int_0^t h(X_s)\,ds + W_t$$

où W_t est un processus de wiener-éventuellement corrélé avec X_t. Soit $\mathscr{F}_t$ la tribu des observations jusqu'à l'instant t. Le problème de *filtrage* consiste à calculer la loi conditionnelle de X_t, sachant $\mathscr{F}_t$. Si $s<t$, le problème de *prédiction* consiste à calculer la loi conditionnelle de X_t, sachant $\mathscr{F}_s$, et le problème de *lissage* à calculer la loi conditionnelle de X_s, sachant $\mathscr{F}_t$. Le problème de filtrage a été étudié par de nombreux auteurs—voir en particulier Liptzer–Shiryaev [12]. La solution du problème de prédiction est un corollaire facile de celle du problème de filtrage. Mais le problème de lissage—dans le cas non-linéaire—a été

beaucoup moins étudié. On trouvera cependant des résultats dans ce sens dans [12] et [13].

Le but de cet article est principalement de montrer comment la méthode que nous avons utilisée en [14] pour établir l'équation de Zakai du filtrage non-linéaire, permet de résoudre également le problème du lissage.

De nouvelles démonstrations—dûes à Krylov–Rosovskii [8]—des résultats essentiels de [12, II° Partie] sont indiquées; elles sont beaucoup plus simple que dans [14].

Dans une seconde partie, nous établissons par une méthode directe la forme dite "robuste" (voir Clark [2], Davis [3]), des équations du filtrage, de la prédiction et du lissage, dans le cas où le bruit d'observation est indépendant du signal X_t.

La fin de cet article étudie deux généralisations. L'une concerne le cas où certains coefficients sont non bornés, de façon à inclure comme cas particulier le cas linéaire. L'autre concerne le cas où le processus observé Y_t apparaît dans les coefficients de l'équation du signal X_t.

La plupart des résultats des §1 à 5 sont également décrits dans [16].

2. FORMULATION DU PROBLEME—HYPOTHESES

On considère le système différentiel stochastique:

$$
\begin{aligned}
dX_t &= b(t, X_t)\,dt + \sigma(t, X_t)\,dW_t \\
dY_t &= h(t, X_t)\,dt + g(t)\,dW_t + \tilde{g}(t)\,d\tilde{W}_t
\end{aligned}
\tag{2.1}
$$

où X_t prend ses valeurs dans R^N, Y_t dans R^D, et $(W_t/\tilde{W}_t)$ est un wiener standard à valeurs dans R^{N+D}.

On fait les hypothèses suivantes sur les coefficients:

$\sigma = \sigma^*$ est une application mesurable et bornée de $R_+ \times R^N$, à valeurs matrices $N \times N$ (2.2)

$x \to \sigma(t, x)$ est continue, uniformément sur tout compact de $R_+ \times R^N$ (2.3)

$$\frac{\partial \sigma_{ij}}{\partial x_i} \in L^\infty(R_+ \times R^N),\ i, j = 1 \ldots N \tag{2.4}$$

$$\exists \alpha > 0 \text{ t.q } a(t, x) = \sigma 0 \sigma(t, x) \geqq \alpha I, \forall (t, x). \tag{2.5}$$

b et h sont mesurables et bornées, de $R_+ \times R^N$, à valeurs dans R^N et R^D respectivement (2.6)

g et $\tilde{g}$ sont mesurables, de R_+ à valeurs matrices $D \times N$ et $D \times D$ respectivement. (2.7)

On suppose que l'on a normalisé le bruit d'observation, i.e.:

$$g(t)g^*(t) + \tilde{g}(t)\tilde{g}^*(t) \equiv I \tag{2.8}$$

On suppose enfin:

$$\exists \beta > 0 \text{ t.q. } \tilde{g}(t)\tilde{g}^*(t) \geqq \beta I,\ \forall t \geqq 0. \tag{2.9}$$

On pose $\Omega = C(R+; R^{N+D})$,

$$\begin{pmatrix} X_t(\omega) \\ Y_t(\omega) \end{pmatrix} = \omega(t), \mathscr{G}_t^s = \sigma(\omega(\theta), s \leqq \theta \leqq t),$$

$$\mathscr{G}^s = \bigvee_{t \geqq s} \mathscr{G}_t^s, \mathscr{G}_t = \mathscr{G}_t^0, \mathscr{G} = \mathscr{G}^0.$$

Il résulte des hypothèses cidessus—cf. Stroock–Varadhan [18]—que $\forall (s,x)$, $\exists$ une unique loi de probabilité P_{sx} sur $(\Omega, \mathscr{G}^s)$ telle que:

i) $P_{sx}(X_s = x, Y_s = 0) = 1$

ii) P_{sx} est solution du problème de martingales associé à (2.1)

Soit π_0 une loi de probabilité sur R^N, de densité $p_0(x)$. On suppose:

$$p_0 \in L^2(R^N) \tag{2.10}$$

On notera P la probabilité sur $(\Omega, \mathscr{G})$, solution du problème de martingales associté à (2.1), tel que sous P, la loi de X_0 soit π_0, et $Y_0 = 0$ p.s. On supposera que toutes les tribus introduites ci-dessus— et plus loin— contiennent tous les ensembles de P-mesure nulle de la tribu $\mathscr{G}$.

On pose

$$Z_t^s = \exp\left\{\int_s^t h(X_\theta) . dY_\theta - \tfrac{1}{2}\int_s^t |h(X_\theta)|^2 d\theta\right\}, \ Z_t = Z_t^0$$

et on définit les lois $\mathring{P}$ sur $(\Omega, \mathscr{G})$ et $\mathring{P}_{sx}$ sur $(\Omega, \mathscr{G}^s)$ par: $\forall t > s$,

$$\left.\frac{d\mathring{P}}{dP}\right|_{\mathscr{G}_t} = Z_t^{-1}, \left.\frac{d\mathring{P}_{sx}}{dP_{sx}}\right|_{\mathscr{G}_t^s} = (Z_t^s)^{-1}$$

Alors il existe un $\mathring{P}$-wiener standard $(W_t'/\tilde{W}_t')$ tel que $\mathring{P}$ p.s.:

$$\begin{cases} dX_t = [b(t, X_t) - c^*(t, X_t) h(t, X_t)]\, dt + \sigma(t, X_t)\, dW_t' \\ dY_t = g(t)\, dW_t' + \tilde{g}(t)\, d\tilde{W}_t' \end{cases} \tag{2.11}$$

où $c = g\sigma^*$

Dans toute la suite, on écrira $\mathring{\Omega}$ pour $(\Omega, \mathscr{G}, \mathring{P})$

On pose $\mathscr{F}_t^s = \sigma\{Y_\theta - Y_s, s \leqq \theta \leqq t\}, \mathscr{F}_t = \mathscr{F}_t^0$.

On notera E[resp. E, E_{sx}, $\mathring{E}_{sx}$] l'espérance suivant la loi P [resp. $\mathring{P}$, P_{sx}, $\mathring{P}_{sx}$].

Le résultat suivant est facile à vérifier, et permet de ramener l'étude des lois conditionnelles sous la loi P, à celle des lois conditionnelles sous la loi $\mathring{P}$:

LEMME 2.1 *Soient λ et $\mu > 0$. On pose $\theta = \lambda v \mu$. Alors $\forall V$ v.a.r. $\mathscr{G}_\lambda$ mesurable et bornée,*

$$E(V/\mathscr{F}_\mu) = \frac{\mathring{E}(VZ_\theta \mid \mathscr{F}_\mu)}{\mathring{E}(Z_\mu \mid \mathscr{F}_\mu)} \text{ p.s.}$$

Introduisons enfin deux opérateurs aux dérivées partielles. On désigne par L_t le générateur infinitésimal du processus de Markov X_t:

$$L_t = \tfrac{1}{2} \sum_{i,j=1}^N a_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^N b_i(t,x) \frac{\partial}{\partial x_i}$$

Soient $H^1 \triangleq H^1(R^N) = \{u \in L^2(R^N); (\partial u/\partial x_i) \in L^2(R^N), i = 1 \dots N\}$ et $H^{-1} = (H^1)'$. Grâce à (2.4), L_t peut-être redéfini comme une famille d'éléments de $\mathscr{L}(H^1; H^{-1})$ de la façon suivante: $\forall u, v \in H^1$,

$$\langle L_t u, v \rangle = -\tfrac{1}{2} \sum_{i,j=1}^N \int_{R^N} a_{ij}(t,x) \frac{\partial u}{\partial x_i}(x) \frac{\partial v}{\partial x_j}(x)\, dx$$

$$+ \sum_{i=1}^N \int_{R^N} \bar{b}_i(t,x) \frac{\partial u}{\partial x_i}(x) v(x)\, dx$$

où

$$\bar{b}_i = b_i - \tfrac{1}{2} \sum_{j=1}^{N} \frac{\partial a_{ij}}{\partial x_j}$$

On définit de plus $B_t \in \mathscr{L}(H^1; (L^2(R^N))^D)$ par:

$$(B_{kt}u)(x) = h_k(t,x)u(x) + \sum_{i=1}^{N} c_{ki}(t,x)\frac{\partial u}{\partial x_i}(x), \; k=1,\ldots,D$$

Remarque 2.2 Il nous arrivera dans cet article d'utiliser des processus du type $\alpha(X_t)$, où $\alpha(x)$ n'est défini que pour presque tout $x \in R^N$. Les hypothèses faites ci-dessus entraînent que $\forall t$, la loi de X_t admet une densité par rapport à dx. Donc $\alpha(X_t)$ est alors bien défini comme une classe de variables aléatoires p.s. égales. ■

3. LE CAS GENERAL

3.1 Le problème de filtrage

Nous allons rappeler les résultats de [14] en donnant les démonstrations proposées depuis par Krylov–Rozovskii [voir [8], où ces résultats sont d'ailleurs généralisés au cas où α et β de (2.5)-(2.9) peuvent s'annuler], lesquelles sont beaucoup plus simples que celles de [14].

Introduisons tout d'abord la notion d'intégrale de Ito rétrograde. (Ω, $\mathscr{F}$, $\mathscr{F}_s$, $\mathring{P}$, Y_s) étant un processus de Wiener standard à valeurs dans $\mathbb{R}^D$, fixons $t>0$—qui jouera le rôle d'instant final.

Posons $\mathscr{F}_t^s = \sigma\{Y_\theta - Y_t, s \leqq \theta \leqq t\}$. $\tilde{Y}_s \triangleq Y_s - Y_t$ est un "$\mathscr{F}_t^s$ processus de Wiener rétrograde"; i.e. $\forall 0 \leqq \theta \leqq s$, $\tilde{Y}_\theta - \tilde{Y}_s$ est une v.a. gaussienne centrée, d'opérateurs de covariance $(s-\theta)I$, indépendante de $\mathscr{F}_t^s$. Si $\{\xi_s, s \in [0,t]\}$ est un processus à valeurs dans $\mathbb{R}^D$, $\mathscr{F}_t^s$-adapté, à trajectoires continues et borné, on peut définir l'intégrale de Ito rétrograde.

$$\int_s^t \xi_\theta \oplus dY_\theta \triangleq P\text{-}\lim_{\delta_n \downarrow 0} \sum_{i=1}^{n} \xi_{t_{i+1}} \cdot (Y_{t_{i+1}} - Y_{t_i})$$

où $s = t_0 < t_1 < \ldots < t_n = t$, et $\delta_n = \sup_{k \leqq n} t_k - t_{k-1}$.

Cette définition s'étend sans peine comme la définition usuelle.

Si ξ_s est mesurable et $\mathscr{F}_t^s$ adapté, avec

$$E\left[\left(\int_0^t |\xi_s|^2 \, ds\right)^{1/2}\right] < \infty, \text{ alors } \left\{\int_s^t \xi_\theta \oplus dY_\theta, 0 \leqq s \leqq t\right\}$$

est une $\mathscr{F}_t^s$ martingale rétrograde. On a de plus une formule de Ito rétrograde. Si ξ_s et φ_s sont mesurables et $\mathscr{F}_t^s$ adaptés à valeurs dans $\mathbb{R}^D$ et $\mathbb{R}$, avec

$$\int_0^t |\xi_s|^2\,ds<\infty \text{ et } \int_0^t |\varphi_s|^2\,ds<\infty \text{ p.s., et si } \Phi\in C^2(\mathbb{R})$$

alors avec:

$$X_s=x+\int_s^t \xi_\theta \oplus dY_s+\int_s^t \varphi_\theta\,d\theta,$$

on a:

$$d\Phi(X_s)=-\Phi'(X_s)\varphi_s\,ds-\Phi'(X_s)\xi_s\oplus dY_s-\tfrac{1}{2}\Phi''(X_s)\xi_s^2\,ds$$

Considrons l'EDP stochastique rétrograde:

$$\begin{cases} dv(s)+L_s v(s)\,ds+B_s v(s)\oplus dY_s=0, \ s\leqq t \\ v(t)=f \end{cases} \tag{3.1}$$

où $f\in L^2(R^N)\cap L^\infty(R^N)$. On cherche une solution $v(s)$ adaptée à $\mathscr{F}_t^s$. Alors (3.1) admet une solution unique:

$$v\in L^2(\mathring{\Omega}\times]0,t[;H^1)\cap L^2(\mathring{\Omega};\subset([0,t];L^2(R^N)))$$

On a de plus le résultat suivant—qui est une sorte de généralisation de la formule de Feynman–Kac:

THÉORÈME 3.1 $\forall s\in[0,t]$, on a l'égalité suivant $d\mathring{P}\times dx$ p.p.:

$$v(s,x)=\mathring{E}_{sx}(f(X_t)Z_t^s\,|\,\mathscr{F}_t^s) \tag{3.2}$$

Preuve Etant donné $\varphi\in L^\infty(0,t;\mathbb{R}^D)$, on définit le processus:

$$\rho^s=\exp\left\{\int_s^t \varphi_\theta\,dY_\theta-\tfrac{1}{2}\int_s^t|\varphi_\theta|^2\,d\theta\right\}$$

φ étant déterministe, l'intégrale stochastique dans l'exponentielle peut être prise en n'importe quel sens. ρ^s est $\mathscr{F}_t^s$ adapté, et il résulte de la formule de Ito rétrograde:

$$d\rho^s=-\rho^s\varphi_s\oplus dY_s$$

Soient $u \in V$, $V(s)=\rho^s v(s)$. Appliquant la formule de Ito rétrograde, on obtient, en désignant—ici et dans toute la suite—par $(.,.)$ le produit scalaire dans $L^2(\mathbb{R}^N)$:

$$d(V(s),u)+\langle LV(s),u\rangle ds+(BV(s),u)\oplus dY_s$$
$$+\varphi_s(V(s),u)\oplus dY_s+\varphi_s(BV(s),u)\,ds=0$$

Cette égalité est vraie $\forall u \in V$. Posons $\bar{V}(s)=\mathring{E}[V(s)]$. On obtient

$$\begin{cases} \dfrac{d}{ds}\bar{V}(s)+L\bar{V}(s)+\varphi_s B\bar{V}(s)=0, & s \leqq t \\ \bar{V}(t)=f \end{cases}$$

Définissons une nouvelle loi de probabilité P^{φ}_{sx} sur $(\Omega, \mathscr{G}^s_t)$ par:

$$\left.\frac{dP^{\varphi}_{sx}}{dP_{sx}}\right|_{\mathscr{G}^s_t}=\rho^s\exp\left[-\int_s^t \varphi_\theta h(\theta, X_\theta)\,d\theta\right]$$

Il résulte du théorème de Girsanov qu'il existe un P^{φ}_{sx} Wiener $(W''_t/\tilde{W}''_t)t$ $-q$:

$$\begin{cases} dX_s=[b(s,X_s)+c^*(s,X_s)\varphi_s]\,ds+\sigma(s,X_s)\,dW''_s \\ dY_s=[h(s,X_s)+\varphi_s]\,ds+g(s)\,dW''_s+\tilde{g}(s)\,d\tilde{W}''_s \end{cases}$$

La solution $\bar{V}$ de l'équation ci-dessus peut s'écrire à l'aide de la formule de Feynmann-Kac:

$$\bar{V}(s,x)=E^{\varphi}_{sx}\left[f(X_t)\exp\left[\int_s^t \varphi_\theta h(\theta, X_\theta)\,d\theta\right]\right\}$$

Ce résultat est classique (cf. par example Bensoussan–Lions [1]) avec un peu plus de régularité sur les coefficients. Notre résultat s'en déduit par passage à la limite (cf. [5]). La formule ci-dessus peut se réécrire:

$$\bar{V}(s,x)=\mathring{E}_{sx}\{\rho^s\mathring{E}_{sx}[f(X_t)Z^s_t/\mathscr{F}^s_t]\}.$$

Or

$$\bar{V}(s,x)=\mathring{E}_{sx}[\rho^s v(s,x)]$$

Mais lorsqu'on fait varier φ dans $L^\infty(0,t;\mathbb{R}^D)$, les combinaisons linéaires des ρ^s correspondants forment un ensemble dense dans $L^2(\Omega,\mathscr{F}_t^s,\mathring{P}_{sx})$. ■

On considère alors l'EDP stochastique progressive:†

$$\begin{aligned} dp(s) &= L_s^* p(s)\,ds + B_s^* p(s)\cdot dY_s \\ p(0) &= po \end{aligned} \tag{3.3}$$

Cette fois, un cherche une solution $\mathscr{F}_s$-adaptée, et de nouveau (3.3) a une solution unique:

$$p \in L^2(\mathring{\Omega}x]0,t[\,;H^1) \cap L^2(\mathring{\Omega};\, C([0,T];L^2(R^N)))$$

Le résultat intéressant est que (3.3) est l'équation adjointe de (3.1), au sens suivant:

THÉORÈME 3.2 *Le processus* $\{(v(s),p(s)),\ 0 \leqq s \leqq t\}$ *est à trajectoires p.s. constantes.*

Preuve Il suffit de montrer que $\forall s_1, s_2 \in [0,t]$,

$$(v(s_1),p(s_1)) = (v(s_2),p(s_2)) \quad \text{p.s.}$$

Soient φ et ρ^s définis comme dans la preuve du théorème 3.1. On pose:

$$\rho_s = \exp\left[\int_0^s \varphi_\theta\,.\,dY_\theta - \tfrac{1}{2}\int_0^s |\varphi_\theta|^2\,d\theta\right],\ \rho = \rho^s\rho_s$$

D'après le dernier argument utilisé au Théorème 3.1, il suffit de montrer que $\forall s_1, s_2 \in [0,t]$,

$$\mathring{E}[\rho(v(s_1),p(s_1))] = \mathring{E}[\rho(v(s_2),p(s_2))]$$

i.e. il suffit de montrer que l'application:

$$s \to \mathring{E}[\rho(v(s),p(s))]$$

†En Anglais, forward.

est constante, de $[0,t]$ dans $\mathbb{R}$. Or:

$$\begin{aligned}\mathring{E}[\rho(v(s),p(s))] &= \mathring{E}[(\rho^s v(s),\rho_s p(s))]\\ &= (\mathring{E}[\rho^s v(s)],\mathring{E}[\rho_s p(s)])\\ &= (V(s),\bar{P}(s))\end{aligned}$$

où $\bar{V}$ et $\bar{P}$ sont les solutions de:

$$\begin{cases}\dfrac{d}{ds}\bar{V}(s)+L\bar{V}(s)+\varphi_s B\bar{V}(s)=0\\ \bar{V}(t)=f\end{cases}$$

$$\begin{cases}\dfrac{d}{ds}\bar{P}(s)=L^*\bar{P}(s)+\varphi_s B^*\bar{P}(s)\\ \bar{P}(0)=p_0\end{cases}$$

et il est immédiat que

$$\frac{d}{ds}(\bar{V}(s),\bar{P}(s))=0 \qquad \blacksquare$$

Il résulte alors des théorèmes 3.1 et 3.2:

$$\mathring{E}(f(X_t)Z_t\,|\,\mathscr{F}_t)=\int_{R^N} p_0(x)\mathring{E}_{0x}(f(X_t)Z_t\,|\,\mathscr{F}_t)\,dx$$

$$\mathring{E}(f(X_t)Z_t\,|\,\mathscr{F}_t)=(p(t),f) \tag{3.4}$$

Ceci est vrai $\forall f\in L^2(R^N)\cap L^\infty(R^N)$, et $\forall t>0$, donc d'après le lemme 2.1, $p(t,x)$ est la "densité conditionnelle non normalisée" de X_t, sachant $\mathscr{F}_t$; i.e.:

COROLLAIRE 3.3 *$p(t,x)(\int_{R^N}p(t,x)\,dx)^{-1}$ est la densité de la loi conditionnelle de X_t, sachant $\mathscr{F}_t$.*

Remarque 3.4 L'équation (3.3) est souvent appelée "équation de Zakai". Cet auteur l'a en effet établie—voir [19]—dans un cadre un peu moins général, en particulier en supposant $g=0$. Le densité conditionnelle satisfait une EDP stochastique non linéaire, appelée équation de Kushner–Stratonovtich, qui est moins facile à traiter, tant théoriquement que numeriquement.

3.2 Le problème de prédiction

Supposons que l'on observe Y jusqu'à l'instant $s<t$; et que l'on cherche la loi conditionnelle de X_t, sachant $\mathscr{F}_s$. Ce problème—dit de prédiction se ramène artificiellement à un problème de filtrage de la façon suivante. Posons:

$$\hat{Y}_t = Y_{s\wedge t} + \tilde{W}_t - \tilde{W}_{s\wedge t}, \hat{\mathscr{F}}_t^s = \sigma\{\hat{Y}_\theta - \hat{Y}_s, s \leqq \theta \leqq t\},$$

$$\hat{\mathscr{F}}_t = \hat{\mathscr{F}}_t^0, \hat{Z}_t = Z_{t\wedge s}, \left.\frac{d\hat{P}}{dP}\right|_{\mathscr{G}_t} = (\hat{Z}_t)^{-1}$$

Alors, d'après les résultats du §3.1, si $\hat{p}$ désigne la solution de:

$$\begin{cases} d\hat{p}(\theta) = L_\theta^* \hat{p}(\theta)\, d\theta + 1_{\{\theta<s\}} B_\theta^* \hat{p}(\theta). dY_\theta \\ \hat{p}(0) = p_0 \end{cases} \tag{3.5}$$

$$E(f(X_t) \mid \hat{\mathscr{F}}_t) = \frac{\hat{E}(f(X_t)\hat{Z}_t \mid \hat{\mathscr{F}}_t)}{\hat{E}(\hat{Z}_t \mid \hat{\mathscr{F}}_t)}$$

$$= (\hat{p}(t), f)(\hat{p}(t), 1)^{-1}$$

De plus, puisque $\hat{\mathscr{F}}_t = \mathscr{F}_s \vee \hat{\mathscr{F}}_t^s$, et que $\hat{\mathscr{F}}_t^s$ et $\mathscr{F}_s \vee \sigma(X_t)$ sont indépendantes sous la loi P,

$$E(f(X_t) \mid \hat{\mathscr{F}}_t) = E(f(X_t) \mid \mathscr{F}_s)$$

On a donc le:

THÉORÈME 3.5

$$\forall 0 \leqq s < t,\ \hat{p}(t,x)\left(\int_{R^N} \hat{p}(t,x)\, dx\right)^{-1}$$

est la densité de la loi de X_t, conditionnée par $\mathscr{F}_s$. ∎

Remarquons que la technique que nous venons d'employer permet aussi de résoudre des problèmes de filtrage (ou prédiction, ou lissage) lorsque l'observation est intermittente.

3.3 Le problème de lissage

On cherche maintenant la loi conditionnelle de X_s, sachant $\mathscr{F}_t$ ($s<t, s$ et t fixés dans tout ce paragraphe). On va voir que la solution de ce problème s'exprime à l'aide de deux équations de type (3.1)—rétrograde, et (3.3)—progressive.

L'équation (3.1) n'est plus seulement un outil intermédiaire pour les démonstrations.

On considère l'EDP stochastique rétrograde:

$$\begin{cases} dv(\theta)+L_\theta v(\theta)\,d\theta+B_\theta v(\theta)\oplus dY_\theta=0, \ \theta\leqq t \\ v(t,x)\equiv 1 \end{cases}$$

On choisit une "fonction poids" ρ par exemple $\rho(x)=(1+|x|^2)^{-N}$ telle que $\rho(x)\geqq 0$, $|\partial\rho/\partial x_i(x)|\leqq c\rho(x)$ et $L^\infty(R^N)\subset L^2(R^N;\rho(x)\,dx)$. On note L^2_ρ et H^1_ρ les espaces $L^2(R^N)$ et H^1, où l'on a remplacé la mesure dx par $\rho(x)\,dx$. Alors la méthode développée dans [14, I° partie] permet de montrer l'existence et l'unicité d'une solution de (3.6):

$$v\in L^2(\mathring{\Omega}x]0,t[;H^1_\rho)\cap L^2(\mathring{\Omega};C([0,t];L^2_\rho))$$

$v(\theta)$ étant $\mathscr{F}^\theta_t$ adapté. De plus, il résulte du théorème 3.1, par un passage à la limite monotone:

PROPOSITION 3.6 *L'égalité suivante a lieu $d\mathring{P}\times dx$ p.p.:*

$$v(s,x)=\mathring{E}_{sx}(Z^s_t\,|\,\mathscr{F}^s_t) \qquad \blacksquare$$

On désigne maintenant à nouveau par p la solution de l'équation (3.3). On a le:

THÉORÈME 3.7 $\forall f\in C_0(R^N)$†,

$$\mathring{E}[f(X_s)Z_t\,|\,\mathscr{F}_t]=\int_{R^N} p(s,x)v(s,x)f(x)\,dx$$

Admettons un instant ce résultat. Il résulte alors:

COROLLAIRE 3.8 *$p(s,x)v(s,x)(p(s),v(s))^{-1}$ est la densité de la loi conditionnelle de X_s, sachant $\mathscr{F}_t$.*

†$C_0(R^N)$ désigne l'espace des fonctions de R^N dans R, continues à support compact.

Preuve D'après (3.4) et la proposition 3.6, $p(s,x)v(s,x) \geqq 0\, d\mathring{P} \times dx$ p.p. Il résulte alors du théorème 3.7, par un passage à la limite monotone:

$$\mathring{E}[Z_t/\mathscr{F}_t] = (p(s), v(s))$$

et cette quantité appartient p.s. à $]0, +\infty[$. Il résulte alors du Lemme 2.1:

$$E[f(X_s)/\mathscr{F}_t] = \frac{(p(s), v(s)f)}{(p(s), v(s))}$$

et ceci $\forall f \in C_0(R^N)$. ■

Le démonstration du théorème utilise les deux Lemmes:

LEMME 3.9 *Si* $f \in C_0(R^N)$,

$$\mathring{E}[f(X_s)Z_t \,|\, \mathscr{F}_t] = \mathring{E}^{\mathscr{F}_t}[f(X_s)Z_s \mathring{E}_{sX_s}^{\mathscr{F}_t^s}(Z_t^s)]$$

Preuve Il suffit d'établir que $\forall \varphi$ v.a.r. $\mathscr{F}_s$ mesurable et bornée, $\forall \psi$ v.a.r. $\mathscr{F}_t^s$ mesurable et bornée, les produits scalaires dans $L^2(\mathring{\Omega})$ des deux membres de l'égalité ci-dessus avec $\varphi\psi$ coïncident. On va utiliser ci-dessous la propriété de Markov du processus $\{(X_\theta, Y_{\theta \vee s} - Y_s), \theta \geqq 0\}$, dont l'état à l'instant $\theta = s$ ne dépend que de X_s.

$$\begin{aligned}\mathring{E}[\mathring{E}(f(X_s)Z_t \,|\, \mathscr{F}_t)\varphi\psi] &= \mathring{E}[f(X_s)Z_s\varphi \mathring{E}^{\mathscr{G}_s}(Z_t^s\psi)] \\ &= \mathring{E}[f(X_s)Z_s\varphi \mathring{E}_{sX_s}(\psi \mathring{E}_{sX_s}^{\mathscr{F}_t^s}(Z_t^s))] \\ &= \mathring{E}[f(X_s)Z_s \mathring{E}_{sX_s}^{\mathscr{F}_t^s}(Z_t^s)\varphi\psi] \quad ■\end{aligned}$$

On note $\mathscr{B}_N$ la tribu borélienne de R^N

LEMME 3.10 *Soit* $(x,\omega) \to G(x,\omega)$ *une application* $\mathscr{B}_N \otimes \mathscr{F}_t^s$ *mesurable à valeurs dans* R_+. *Alors:*

$$\mathring{E}[G(X_s)Z_s \,|\, \mathscr{F}_t] = \int_{R^N} p(s,x)G(x)\,dx \quad \text{p.s.}$$

Preuve D'après le théorème des classes monotones, il suffit d'établir le résultat pour:

$$G(x,\omega) = g(x)1_A(\omega)$$

$$\forall A \in \mathscr{F}_t^s, \forall g = 1_B, B \in \mathscr{B}_N.$$

Il suffit donc de montrer que $\forall \varphi$, ψ comme au lemme précédent,

$$\mathring{E}[g(X_s)1_A Z_s \varphi\psi] = \mathring{E}[\varphi\psi 1_A (p(s), g)]$$

On va utiliser (3.4):

$$\begin{aligned}\mathring{E}[g(X_s)Z_s 1_A \varphi\psi] &= \mathring{E}[g(X_s)Z_s\varphi]\mathring{E}[1_A\psi]\\ &= \mathring{E}[\varphi(p(s), g)]\mathring{E}[1_A\psi]\\ &= \mathring{E}[\varphi\psi 1_A(p(s), g)] \quad \blacksquare\end{aligned}$$

Preuve de théoreme 3.7 Il suffit de démontrer le théorème avec $f \geqq 0$. On utilise successivement le Lemme 3.9, la Proposition 3.6, et le Lemme 3.10:

$$\begin{aligned}\mathring{E}(f(X_s)Z_t \mid \mathscr{F}_t) &= \mathring{E}^{\mathscr{F}_t}(f(X_s)Z_s \mathring{E}^{\mathscr{F}_t^s}_{sX_s}(Z_t^s))\\ &= \mathring{E}^{\mathscr{F}_t}(f(X_s)Z_s v(s, X_s))\\ &= \int_{R^N} p(s,x) f(x) v(s,x)\, dx\end{aligned}$$

4. LE CAS OU LE BRUIT D'OBSERVATION EST INDEPENDANT DU SIGNAL

On rajoute aux hypothèses du §2:

$$g \equiv 0 \tag{4.1}$$

$$\frac{\partial h}{\partial t}, \frac{\partial h}{\partial x_i} \in L^\infty(R^+ \times R^N),\ i = 1\ldots, N \tag{4.2}$$

Il résulte de (4.1) et (2.8) que la seconde équation de (2.1) se réécrit:

$$Y_t = \int_0^t h(s, X_s)\, ds + \tilde{W}_t$$

De plus, $c = 0$ dans (2.11), et dans l'expression de B.

4.1 Le problème de filtrage

v et p étant les quantités définies au §3.1, on considère maintenant:

$$\begin{cases} u(t,x) = v(t;x)\exp[Y_t, h(t,x)] \\ q(t,x) = p(t,x)\exp[-Y_t. h(t,x)] \end{cases} \tag{4.3}$$

On pourrait déduire de (3.1) et (3.3) les équations pour u et q-comme l'on fait Rozovskii [17] Liptzer–Shiryaev [12] et Clark [2] pour q. Le changement de variable (4.3) est un cas particulier des changements de variable proposés par Doss [4] et Sussmann [19], pour ramener la résolution d'une équation de Ito à celle d'une équation différentielle ordinaire. Cependant nous allons plutôt indiquer une dérivation directe des équations du u et q suivant l'exposé [15]. L'intérêt est que cette dérivation est tout à fait élémentaire (sauf en ce qui concerne les notations !), et se prête à des généralisations commodes, par exemple en théorie du contrôle—voir Fleming–Pardoux [5]. Davis [3] a également établi l'équation de q, dans le cas où X_t appartient à une classe plus générale de processus de Markov que celle considérée ici.

Remarquons tout d'abord que sous les lois $\mathring{P}$ et $\mathring{P}_{sx}$, $X.$ et $Y.$ sont indépendants, i.e.:

$$\mathring{P}_{sx}(dX, dY) = \bar{P}_{sx}(dX) \times W_{s0}(dY)$$

$$\mathring{P}(dX, dY) = \bar{P}(dX) \times W(dY)$$

où W[resp. W_{s0}] désigne la mesure de Wiener sur $C(R+; R^D)$ [resp. $C([s, +\infty[; R^D)]$, et $\bar{P}$, $\bar{P}_{sx}$ des solutions du problème de martingales associées à:

$$dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$$

Alors

$$\mathring{E}_{sx}(f(X_t)Z_t^s/\mathscr{F}_t^s) = \bar{E}_{sx}(f(X_t)Z_t^s), W_{s0} \quad \text{p.s.} \tag{4.4}$$

Supposons pour l'instant que la fonction h est "régulière". Alors, en intégrant par parties l'intégrale stochastique qui se trouve dans l'expression de Z_t^s, on obtient:

$$\exp(Y_s.h(X_s))Z_t^s=\exp(Y_t.h(X_t))$$

$$\times\exp\left\{-\int_s^t Y_\theta.(\sigma\nabla h)(X_\theta)\,dW_\theta-\tfrac{1}{2}\int_s^t (aY_\theta.\nabla h,\,Y_\theta.\nabla h)(X_\theta)\,d\theta\right\}$$

$$\times\exp\left\{\int_s^t e(\theta,X_\theta,Y_\theta)\,d\theta\right\}$$

où

$$e(\theta,X,X)=\tfrac{1}{2}(aY.\nabla h(X),\,Y.\nabla h(X))-Y.(h'_\theta+L_\theta h)(X)-\tfrac{1}{2}|h(X)|^2$$

Fixons maintenant une trajectoire de Y. On définit alors une nouvelle mesure $\check{P}_{sx}^Y$ sur $(\Omega,\sigma(X_\theta,s\leqq\theta\leqq t))$ par:

$$\frac{d\check{P}_{sx}^Y}{d\bar{P}_{sx}}=\exp\left\{-\int_s^t Y_\theta.(\sigma\nabla h)(X_\theta)\,dW_\theta-\tfrac{1}{2}\int_s^t (aY_\theta.\nabla h,\,Y_\theta.\nabla h)(X_\theta)\,d\theta\right\}. \tag{4.5}$$

Alors compte tenu de (3.2), (4.3), (4.4), (4.5) et du calcul fait ci-dessus:

$$u(s,x)=\check{E}_{sx}^Y\left\{f(X_t)\exp\left[Y_t.h(X_t)+\int_s^t e(\theta,X_\theta,Y_\theta)\,d\theta\right]\right\} \tag{4.6}$$

D'après la formule de Feyman-Kac, on est amené à considérer l'EDP rétrograde suivante:

$$\frac{d\bar{u}}{ds}+\check{L}_s\bar{u}+e(s,Y_s)\bar{u}=0,\quad s\leqq t$$

$$\bar{u}(t)=f\exp(Y_t h(t))$$

où

$$\check{L}_s u=L_s u-\sum_{i=1}^N (aY_s.\nabla h)_i\frac{\partial u}{\partial x_i} \tag{4.7}$$

Après intégration par parties, pour faire disparaître les dérivées secondes de h, (4.7) devient:

$$\frac{d\bar{u}}{ds}+\hat{L}_s\bar{u}+\hat{e}(s,Y_s)\bar{u}=0,\quad s\leqq t$$

$$\bar{u}(t)=f\exp(Y_th(t)) \tag{4.8}$$

où $\hat{L}$ et $\hat{e}$ sont définis par: $\forall u,\ v\in H^1$,

$$\langle\hat{L}_su,v\rangle=\langle L_su,v\rangle-\tfrac{1}{2}\sum_i\left((aY_s.\nabla h)_i\frac{\partial u}{\partial x_i},v\right)$$

$$+\tfrac{1}{2}\sum_i\left((aY_s.\nabla h)_iu,\frac{\partial v}{\partial x_i}\right)$$

$\hat{e}(s,Y_s,x)$ s'obtient à partir de $e(s,Y_s,x)$ en remplaçant L_sh par $\bar{b}.\nabla h$

$$\left(\bar{b}_i=b_i-\tfrac{1}{2}\sum_j\frac{\partial a_{ij}}{\partial x_j}\right).$$

Supposons que $f\in C_0(R^N)$. Alors d'après par exemple Bensoussan–Lions [1], pour chaque trajectoire de Y. fixée, l'équation (4.8) a une solution unique:

$$u\in L^2(0,t;H^1)\cap C([0,t];L^2(R^N))$$

et u dépend continûment de Y.—voir Fleming–Pardoux [5]. De plus:

THÉORÈME 4.1 *$\forall s\in[0,t]$, l'égalité suivante a lieu $dx\times dW_{s0}$ p.p.*

$$u(s,x)=\exp[Y_sh(s,x)]\mathring{E}_{sx}(f(X_t)Z_t^s/\mathscr{F}_t^s) \tag{4.9}$$

i.e. la solution de (4.8) *est bien le quantité définie par* (4.3).

Preuve Il suffit de démontrer le résultat avec des coefficients b, σ et h "réguliers", puisque à la fois u—solution de (4.8)—et $\mathring{P}_{sx}$ dépendent continûment de ces coefficients; le résultat général s'obtient alors par passage à la limite—voir les détails de cet argument dans [5].

Or si $h\in C_b^{1,2}(R_+\times R^N;R^D)$, et si σ et b sont un peu plus réguliers que supposé ci-dessus (par exemple chaque b_i et σ_{ij} dans $C_b^{0,1}(R+\times R^N)$), alors d'après Bensoussan–Lions [1], la solution de (4.7)- donc aussi de

(4.8)— est donnée par (4.6), et donc aussi par (4.9), compte tenu du calcul qui nous a mené à (4.6). ■

Considérons maintenant l'équation adjointe de (4.8):

$$\begin{cases} \dfrac{dq}{ds} = \hat{L}_s^* q + \hat{e}(s, Y_s)u \quad s \geqq 0 \\ \\ q(0) = p_0 \end{cases} \tag{4.10}$$

A nouveau pour chaque trajectoire de Y, cette équation a une solution unique:

$$q \in L^2(0, t; \mathrm{H}^1) \cap \mathrm{C}([0, \mathrm{t}]; L^2(R^N))$$

et de plus, pour presque tout $s \in [0, t]$,

$$\frac{d}{ds}(u(s), q(s)) = 0$$

Donc $(q(t), f \exp(Y_t h)) = \int p_0(x) \mathring{E}_{0x}(f(X_t) Z_t \,|\, \mathscr{F}_t)\, dx$

$$(q(t), f \exp(Y_t h)) = \mathring{E}(f(X_t) Z_t \,|\, \mathscr{F}_t). \tag{4.11}$$

Nous avons démontré le:

THÉORÈME 4.2 $\forall t \geqq 0$

$$q(t, x) \exp(Y_t h(t, x)) q(t), \exp(Y_t h(t))^{-1}$$

est la densité de la loi conditionnelle de X_t, *sachant* $\mathscr{F}_t$. ■

Remarque 4.3 L'équation (4.10) permet de construire une application:

$$Y. \to E[f(X_t)/\mathscr{F}_t]$$

définie sur tout l'espace C $([0, t]; R^D)$, et continue. Ce résultat très important est dû à Clark [2]. ■

4.2 Le problème de prédiction

D'après (4.2), on ne peut plus ramener le problème de prédiction à un problème de filtrage, comme nous l'avons fait au §3.2.

Mais si $s<t$, $f\in C_0(R^N)$,

$$E(f(X_t)\,|\,\mathscr{F}_s)=E^{\mathscr{F}_s}E^{\mathscr{G}_s}(f(X_t))$$

Or il résulte d'un argument similaire à celui du Théorème 4.1 que:

$$E^{\mathscr{G}_s}[f(X_t)]=\tilde{u}(s,X_s)$$

où $\tilde{u}$ est la solution de:

$$\begin{aligned}&\frac{d}{d\theta}\tilde{u}+L_\theta\tilde{u}=0,\ \theta\leqq t\\&\tilde{u}(t)=f\end{aligned}\tag{4.12}$$

Et d'après le théorème 4.2, si q désigne à nouveau la solution de (4.10),

$$E^{\mathscr{F}_s}[\tilde{u}(s,X_s)]=\frac{(q(s)\exp(Y_sh(s),\tilde{u}(s))}{(q(s),\exp(Y_sh(s))}$$

De plus, si p désigne la solution de:

$$\begin{aligned}&\frac{d}{d\theta}p=L_\theta^*p,\quad\theta\geqq s\\&p(s,x)=q(s,x)\exp(Y_sh(s,x))\end{aligned}\tag{4.13}$$

Il résulte que la dualité entre (4.12) et (4.13):

$$(q(s)\exp(Y_sh(s)),\tilde{u}(s))=(p(t),f)$$

De plus, puisque (4.13) est une équation de Fokker–Planck:

$$(q(s),\exp(Y_sh(s)))=(p(t),1)$$

on a montré le:

THÉORÈME 4.4 *$\forall t>s$, la densité de la loi conditionnelle de X_t, sachant $\mathscr{F}_s$, est donnée par $p(t,x)(p(t),1)^{-1}$.*

4.3 Le problème de lissage

On cherche maintenant la loi de X_s, conditionée par $\mathscr{F}_t(s<t)$. On

considère l'équation:

$$
\begin{aligned}
&\frac{du}{d\theta}+\hat{L}_\theta u+\hat{e}(\theta, Y_\theta)u=0 \quad \theta\leqq t\\
&u(t)=\exp(Y_t h(t))
\end{aligned}
\tag{4.14}
$$

Si ρ, L^2_ρ et H^1_ρ sont définis comme au §3.3, alors pour toute trajectoire $Y_.$, l'équation (4.14) a une solution unique:

$$u\in L^2(0,t;\mathrm{H}^1_\rho)\cap C([0,t];L^2_\rho)$$

Il résulte de Théorème 4.1, par un passage à la limite monotone:

$$u(s,x)=\exp(Y_s h(s,x))\bar{E}_{sx}(Z^s_t)$$

Alors, d'après la propriété de Markov du processus X_t, et l'égalité $\bar{E}\varphi=\mathring{E}(\varphi|\mathscr{F}_t)$ pour φ $\mathscr{G}_t$ mesurable,

$$
\begin{aligned}
\bar{E}[f(X_s)Z_t]&=\bar{E}[f(X_s)Z_s\bar{E}_{sX_s}(Z^s_t)]\\
&=\bar{E}[f(X_s)Z_s\exp(-Y_s h(X_s))u(s,X_s)]\\
&=(q(s),fu(s))
\end{aligned}
$$

d'après (4.11), si q désigne toujours la solution de (4.10). Il résulte alors du Lemme 2.1:

THÉORÈME 4.5 $\forall s\in[0,t]$, *la densité de la loi conditionnelle de* X_s, *sachant* $\mathscr{F}_t$, *est donnée par*

$$q(s,x)u(s,x)(q(s),u(s))^{-1}$$

ou q *désigne la solution de* (4.10), u *celle de* (4.14). ■

5. LE CAS DES COEFFICIENTS NON BORNES

5.1 Hypothèses

Parmi les hypothèses faites ci-dessus, l'une des plus gênantes est que les coefficients sont tous supposés bornés. Cela interdit en particulier de considérer le cas linéaire comme un cas particulier de la théorie faite ci-dessus. Dans [10], Kunita a considéré une situation assez générale, avec des coefficients non bornés. Cependant, il n'obtient qu'un résultat d'existence locale (en t) pour l'équation du filtrage. Nous allons nous limiter à un problème plus particulier, mais qui inclut le cas linéaire, et donner un résultat d'existence globale pour les équations du filtrage, du lissage et de la prédiction.

Nous nous limitons à la situation étudiée au §4. On fait sur les coefficients de diffusion les mêmes hypothèses qu'au §3; i.e. on suppose vérifiées (2.2), (2.3), (2.4), (2.5), (2.8) et (4.1). On suppose en outre:

$$\begin{cases} b, h, h'_t \text{ sont mesurables de } R_+ \times R^N \text{ à valeurs dans } R^N, R^D \text{ et } R^D; \text{ et} \\ \exists c \text{ t.q.} \\ |b(t,x)| + |h(t,x)| + |h'_t(t,x)| \leqq c(1+|x|), \forall t, x \end{cases} \tag{5.1}$$

$$\frac{\partial h}{\partial x_i} (i=1\ldots,N), \sum_{i,j=1}^{N} a_{ij} \frac{\partial^2 h}{\partial x_i \partial x_j} \in L^\infty(R_+ \times R^N; R^D) \tag{5.2}$$

$$\tfrac{1}{2} \sum_{i,j=1}^{N} \frac{\partial^2 a_{ij}}{\partial x_i \partial x_j} - \sum_{i=1}^{N} \frac{\partial b_i}{\partial x_i} \in L^\infty(R_+ \times R^N) \tag{5.3}$$

$$p_0 \in L^2(R^N) \tag{5.4}$$

5.2 Existence et unicité de solutions aux EDP associées

Considérons tout d'abord l'équation rétrograde (4.7):

$$\begin{cases} \dfrac{d\bar{u}}{ds} + \check{L}_s \bar{u} + e(s, Y_s)\bar{u} = 0 \\ \qquad \bar{u}(t) = f \exp(Y_t h(t)) \end{cases} \tag{5.5}$$

On suppose $f \in C_b(R^N)$.

On cherche une solution u de la forme (4.6) i.e;

$$u(s,x) = \check{E}^Y_{sx} \{ f(X_t) \exp [Y_t h(X_t) + \int_s^t e(\theta, X_\theta, Y_\theta)\, d\theta] \} \tag{5.6}$$

Remarquons que d'après les hypothèses ci-dessus, et l'expression de e, $\exists$ une constante C—qui dépend de $\operatorname{Sup}_{s \leqq \theta \leqq t} |Y_\theta|$—telle que:

$$e(\theta, X_\theta, Y_\theta) \leqq C(1+|X_\theta|), \quad s \leqq \theta \leqq t \tag{5.7}$$

$$\check{b}(\theta, X_\theta, Y_\theta) \leqq C(1+(X_\theta)), \quad s \leqq \theta \leqq t \tag{5.8}$$

où $\check{b} = b - a[Y.\nabla h]$

On peut déduire de (5.6), (5.7) et (5.8), en utilisant une inégalité exponentielle sur les martingales, que $|u(s,x)|$ est majoré par $\exp(d+|x|)e^{d(t-s)+1}$ où d est une constante qui dépend de $\sup_{s\leq\theta\leq t}|Y_\theta|$†.

On pose alors:

$$\varphi(s,x)=\exp[(d+|x|)e^{d(t-s+1)}] \tag{5.9}$$

LEMME 5.1 *Pour toute trajectoire de Y, $\exists d$ tel que l'équation* (5.5) *ait au plus une solution, parmi les fonctions $u(s,x)$ qui vérifient:*

$$\varphi^{-1}u\in L^2(0,t;H^1) \tag{5.10}$$

Preuve Posons

$$\tilde{b}_i=\breve{b}_i-\tfrac{1}{2}\sum_{j=1}^{N}\frac{\partial a_{ij}}{\partial x_j}.$$

Soient u_1 et u_2 deux solutions de (5.5) qui vérifient l'hypothèse du Lemme, $\bar{u}=u_1-u_2$, $\lambda=\varphi^{-1}\bar{u}$. Alors:

$$\varphi\frac{\partial\lambda}{\partial s}+\lambda\frac{-\partial\varphi}{\partial s}+\tfrac{1}{2}\sum_{i,j}\frac{\partial}{\partial x_j}\left(a_{ij}\left[\frac{\partial\lambda}{\partial x_i}\varphi+\frac{\partial\varphi}{\partial x_i}\lambda\right]\right)$$

$$+\sum_i\tilde{b}_i\left[\frac{\partial\lambda}{\partial x_i}\varphi+\frac{\partial\varphi}{\partial x_i}\lambda\right]+e\lambda\ \varphi=0$$

On peut diviser cette égalité par φ, d'où il résulte que $\lambda'_s\in L^2(0,t;H^{-1})$.

On peut donc multiplier scalairement l'égalité résultante par λ Or, d'après un résultat rappelé dans Bensoussan–Lions [1, p. 110],

$$\frac{d}{ds}|\lambda|^2{}_{L^2(R^N)}=\langle\lambda,\lambda'_s\rangle\text{p.p.}$$

On obtient donc après calculs:

$$\tfrac{1}{2}\frac{d}{ds}|\lambda|^2-\tfrac{1}{2}\sum_{i,j}(a_{ij}\lambda'_{x_i},\lambda'_{x_j})$$

$$+(\theta+e-d(d+|x|)e^{d(t-s)},\lambda^2)=0$$

†La démonstration de ce résultat, qui m'a été suggérée par W. Fleming, a été le point de point de départ du raisonnement qui va suivre.

où θ est un coefficient qui vérifie une inégalité de type (5.7). Donc, en choisissant d assez grand, on obtient, compte tenu de (2.5),

$$-\frac{d}{ds}|\lambda(s)|^2+\frac{\alpha}{2}\|\lambda(s)\|_{H^1}^2\leqq 0 \quad \text{p.p.} \tag{5.11}$$

Mais $\lambda(t)=0$, donc $\lambda(s)=0$, $\forall s\in[0,t]$.

THÉORÈME 5.2 *Pour toute trajectoire de Y, l'équation* (5.5) *a une solution unique u vérifiant* (5.9)-(5.10), *avec d suffisamment grand De plus,* $\forall s\in[0,t], dx\times d\mathring{P}$ p.p.

$$u(s,x)=\exp(Y_s h(x))\mathring{E}_{sx}(f(X_t)Z_t^s/\mathscr{F}_t^s) \tag{5.12}$$

Preuve Soit b_n, h_n une suite de coefficients qui vérifient les mêmes hypothèses que b et h, sont bornés, et convergent—pour chaque s fixévers b et h, uniformément sur tout compact. Soit u_n la solution de l'équation $(5.5)_n$ correspondante, et $\lambda_n=u_n\varphi^{-1}$, où d dans φ a été choisi tel que le Lemme 5.1 s'applique, et que $\lambda_n(t)$ reste dans un borné de $L^2(R^N)$. Alors λ_n vérifie (5.11).

Donc λ_n reste dans un borné de $L^2(0,t;H^1)$, et on peut extraire une sous-suite (encore notée λ_n) telle que $\lambda_n\to\lambda$ dans $L^2(0,t;H^1)$ faible. Soit $v\in C_0^\infty(R_+\times R^N)$. Alors, si $(5.5)_n$ s'écrit:

$$\begin{cases}\dfrac{du_n}{ds}+\check{L}_n u_n+e_n u_n=0\\ \qquad u_n(t)=f\exp[Y_t h_n(t)]\end{cases}$$

$$(u_n(t),v)+\int_0^t\langle\lambda_n,\varphi(\check{L}_n)^*v\rangle ds$$
$$+\int_0^t(\lambda_n,\varphi e_n v)\,ds=\int_0^t\left(\lambda_n,\varphi\frac{dv}{ds}\right)ds$$

On vérifie aisément que:

$$\varphi(\check{L}_n)^*v\to\varphi\check{L}^*v \quad \text{dans} \quad L^2 0,T;H^{-1}) \quad \text{fort}$$

$$\varphi e^n v\to\varphi e v \quad \text{dans} \quad L^2(0,T;L^2(R^N)) \quad \text{fort}$$

On peut donc passer à la limite dans l'égalité ci-dessus, d'où il résulte que $u=\varphi\lambda$ une solution unique de (5.5)

Il reste à montrer (5.12).

Fixons $s \in [0, t]$. On peut supposer, quitte à extraire une sous-suite, que

$$u_n(s) \to u(s) \quad \text{dans} \quad L^2(R^N) \quad \text{faible}. \tag{5.13}$$

Il découle du théorème 4.1:

$$u_n(s, x) = \exp[Y_s h^n(s, x)] \mathring{E}^n_{sx}(f(X_t)^n Z^s_t \mid \mathscr{F}^s_t)$$

Il suffit d'établir (5.12) dans le cas $f \geqq 0$. Soit θ une application continue de Ω dans $\mathbb{R}, \geqq 0$, bornée et $\mathscr{F}^s_t$ mesurable. On pose $\theta_R = \theta 1_{\{|Y_s| \leqq R\}}$ Soit $\rho \in C_0(R^N)$, $\rho \geqq 0$, $\int \rho(x)dx = 1$. On note $P^n_{s\rho}$ la loi $\int \rho(x) P^n_{sx}(\cdot)dx$ et on définit de même $P_{s\rho}$. On note W la mesure de Wiener sur $C(R_+; R^D)$. Il résulte de l'indépendance entre Y_s et $\mathscr{F}^s_t$, sous la loi $\mathring{P}$:

$$W\left[\theta_R\left(\frac{\rho}{\exp(Y_s \cdot h^n(s))}, u^n(s)\right)\right] = E^n_{s\rho}[f(X_t)\theta] W(|Y_s| \leqq R) \tag{5.14}$$

Mais $P^n_{s\rho}$ converge étroitement vers $P_{s\rho}$ d'après Stroock–Varadhan [18]; et avec (5.13) et le théorème de convegence dominée de Lebesque, on peut passer à la limite dans le terme de gauche de (5.14) quand $n \to \infty$. On fait ensuite tendre $R \to \infty$, par convergence monotone, d'où:

$$W\left[\theta\left(\frac{\rho}{\exp(Y_s h(s))}, u(s)\right)\right] = E_{s\rho}[f(X_t)\theta]$$

(5.12) découle alors de la latitude de choix de θ et ρ, et du fait que

$$\left(\frac{\rho}{\exp(Y_s h(s))}, u(s)\right) \text{ est } \mathscr{F}^s_t \text{ mesurable, comme limite de v.a.r.}$$

$\mathscr{F}^s_t$ mesurables.

5.3 Equations du filtrage, du lissage et de la prédiction

On considère l'équation adjointe de (5.5):

$$\begin{aligned} \frac{dq}{ds} &= \check{L}^*_s q + e(s, Y_s)q \\ q(0) &= p_0 \end{aligned} \tag{5.15}$$

THÉORÈME 5.3

$$\forall t \geqq 0, q(t,x)\exp[Y_t h(t,x)](q(t),\exp[Y_t h(t)])^{-1}$$

est la densité de la loi conditionnelle de X_t, sachant $\mathscr{F}_t$.

Preuve Soit q^n la solution de (5.15), où b et h sont remplacés par b_n et h_n. D'après (4.11), $\forall f \in C_0(R^N)$:

$$\mathring{E}^n(f(X_t)^n Z_t \mid \mathscr{F}_t) = (q^n(t), f\exp[Y_t h^n(t)]) \tag{5.16}$$

En reprenant la démonstration du théorème 5.2, on s'aperçoit que, la suite q^n se comportant comme la suite u_n, $\varphi^{-1}q^n(t)$ reste dans un borné de $L^2(R^N)$. Donc, quitte à extraire une sous-suite, on peut supposer que $q^n(t) \to q(t)$ dans $L^2_{loc}(R^N)$ faible. Il en résulte que l'on peut passer à la limite dans le membre de droite de (5.16). Le passage à la limite dans le membre de gauche résulte de ce que $\mathring{P}^n$ converge étroitement vers $\mathring{P}$. Donc:

$$E(f(X_t)Z_t \mid \mathscr{F}_t) = (q(t), f\exp[Y_t h(t)]) \tag{5.17}$$

Par un passage à la limite monotone, on obtient la relation (5.17) avec $f = 1$. Le Théorème découle alors du Lemme 2.1.

On peut de la même façon passer à la limite sur le résultat du Théorème 4.4, pour établir l'équation de la prédiction. Les équations du lissage s'obtiennent aisément à partir de (5.12) et de (5.17), par le raisonnement fait au §4.3.

6. LE CAS OU L'OBSERVATION APPARAIT DANS L'EQUATION DU SIGNAL

6.1 Hypothèses et notations

Le modèle considéré jusqu'à présent peut se généraliser en:

$$\begin{cases} dX_t = b(t,X_t,Y_t)\,dt + \sigma(t,X_t,Y_t)\,dW_t \\ dY_t = h(t,X_t,Y_t)\,dt + g(t,Y_t)\,dW_t + \tilde{g}(t,Y_t)\,d\tilde{W}_t \end{cases} \tag{6.1}$$

où (X_t, Y_t) est donc un processus de Markov dont seule la deuxième "composante" Y_t est observée. A nouveau, $(W_t/\tilde{W}_t)$ est un wiener standard à valeurs dans R^{N+D}.

Remarque 6.1 La seule dissymétrie que l'on est obligé de conserver entre X et Y, est la non dépendance en X du coefficient de diffusion de Y. Cette restriction va nous permettre de supposer que l'on a fait la normalisation $gg^*+\tilde{g}\tilde{g}^*=1$, puisque $\bar{Y}_t=\int_0^t[gg^*+\tilde{g}\tilde{g}^*]^{-1/2}dY$ est alors observable au même titre que Y. Cette normalisation fait que, sous la loi $\mathring{P}$, Y est un wiener, ce qui est crucial dans nos dérivations.

Cette restriction semble être cruciale pour pouvoir utiliser une méthode de "probabilité de référence"—voir Szpirglas–Mazzioto [20]. A notre connaissance, le seul travail qui ne fait pas cette hypothèse est celui de Kunita [9], qui utilise une méthode de type "innovation".

Signalons que les équations du filtrage et de la prédiction, pour le problème (6.1), ont déjà été établies par Krylov–Rozovskii [7], par une méthode différente. ■

On fait les hypothèses suivantes sur les coefficients:

b, $\sigma=\sigma^*$, h sont des applications bornées de $R_+\times R^N\times R^D$ à valeurs dans R^N, $R^{N\times N}$ et R^D, continues en (x,y), uniformément sur tout compact de $R_+\times R^N\times R^D$. $\quad (6.2_1)$

$g,\tilde{g}$ sont des applications bornées de $R_+\times R^D$, à valeurs dans $R^{D\times N}$ et $R^{D\times D}$, continues en y, uniformément sur tout compact de $R_+\times R^D$. $\quad (6.2_2)$

$\dfrac{\partial\sigma_{ij}}{\partial x_i}\in L^\infty(R_+\times R^N\times R^D)$ et est continue en (x,y), uniformément sur tout compact de $R_+\times R^N\times R^D$; $i,j=1\ldots,N$ $\quad (6.3)$

$$\exists\alpha>0 \quad \text{t.q.} \quad a(t,x,y)=\sigma\,0\,\sigma(t,x,y)\geqq\alpha I, \qquad \forall(t,x,y). \tag{6.4}$$

$$gg^*+\tilde{g}\tilde{g}^*\equiv I \tag{6.5}$$

$$\exists\beta>0 \quad \text{t.q.} \quad \tilde{g}\tilde{g}^*\geqq\beta I, \forall(t,y) \tag{6.6}$$

$\sum_{k=1}^{D}\dfrac{\partial h_k}{\partial y_k}, \sum_{k=1}^{D}\dfrac{\partial c_{ik}}{\partial y_k}\in L^\infty(R_+\times R^N\times R^D)$ et sont continues en (x,y), uniformément sur tout compact de $R_+\times R^N\times R^D$; $i=1,\ldots,N$. $\quad (6.7)$

On définit Ω, $\mathscr{G}^s$, $\mathscr{G}_t$, P, $\mathring{P}$, Z_t^s et $\mathscr{F}_t^s$ comme au §2, avec à nouveau l'hypothèse (2.10). On définit enfin $\mathring{P}_{sxy}$ et $\mathring{P}_{sx}$, correspondant aux conditions initiales:

sous $\mathring{P}_{sxy}$, $X_s = x$ et $Y_s = y$ p.s.; sous $\mathring{P}_{sx}$, $X_s = x$ p.s. et Y_s est un vecteur aléatoire gaussien centré d'opérateur de covariance sI.

Comme dans (2.11),

$$Y_t = \int_0^t g(s, Y_s)\, dW'_s + \int_0^t \tilde{g}(s, Y_s)\, d\tilde{W}'_s$$

où $(W'_t/\tilde{W}'_t)$ est un $\mathring{P}$-wiener standard à valeurs dans R^{N+D}. Il résulte de (6.5), (6.6) et de ce que gg^* et g^*g ont mêmes valeurs propres:

$$I - g^*g \geqq \beta I$$

On définit

$$\tilde{Y}_t = \int_0^t [I - g^*(s, Y_s) g(s, Y_s)]^{-1/2} (dW'_s - g^*(s, Y_s)\, dY_s)$$

Il résulte alors du théorème de P. Lévy que $(Y_t/\tilde{Y}_t)$ est un $\mathring{P}$-processus de Wiener standard à valeurs dans R^{N+D}. De plus,

$$dX_t = [b(t, X_t, Y_t) - c^*h(t, X_t, Y_t)]\, dt + c^*(t, X_t, Y_t)\, dY_t + \tilde{c}(t, X_t, Y_t)\, d\tilde{Y}_t \tag{6.8}$$

où $c = g\sigma, \tilde{c} = \sigma[I - g^*g]^{1/2}$

On définit enfin les opérateurs à coefficients stochastiques:

$$L_t = \frac{1}{2} \sum_{i,j=1}^{N} a_{ij}(t, x, Y_t) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{N} b_i(t, x, Y_t) \frac{\partial}{\partial x_i}$$

$$B_{kt} = h_k(t, x, Y_t) + \sum_{i=1}^{N} c_{ki}(t, x, Y_t) \frac{\partial}{\partial x_i}$$

$$C_{kt} = \sum_{k=1}^{D} \frac{\partial h_k}{\partial y_k}(t, x, Y_t) + \sum_{k=1}^{D} \sum_{i=1}^{N} \frac{\partial c_{ki}}{\partial y_k}(t, x, Y_t) \frac{\partial}{\partial x_i}$$

6.2 Etude d'une EDP stochastique rétrograde

L'EDP stochastique progressive (3.3) se généralise de façon naturelle et sans ambiguïté, lorsqu'on remplace les opérateurs L_s^* et B_s^* par les adjoints des deux premiers opérateurs à coefficients stochastiques définis ci-dessus. Il n'en va pas de même de l'E.D.P.S. rétrograde (3.1). Comme les opérateurs dépendent de Y_s, la solution $v(s)$ ne peut pas être $\mathscr{F}_t^s=\sigma\{Y_\theta - Y_s, s\leqq\theta\leqq t\}$—adaptée, mais elle sera $\bar{\mathscr{F}}_t^s=\sigma\{Y_\theta; s\leqq\theta\leqq t\}$—adaptée.

Or au §3, on a utilisé le fait que $\tilde{Y}_s=Y_s-Y_t$ est un "$\mathring{P}$—$\mathscr{F}_t^s$—Wiener rétrograde", i.e. $\forall s_1<s_2$, $\tilde{Y}_{s_1}-\tilde{Y}_{s_2}$ est un vecteur aléatoire gaussien centré d'opérateur de covariance $(s_2-s_1)I$, indépendent de $\mathscr{F}_t^s$. Il est clair que $\tilde{Y}_s$ n'est pas un $\bar{\mathscr{F}}_t^s$—Wiener rétrograde.

Mais

$$\bar{\mathscr{F}}_t^s=\mathscr{F}_t^s \vee \sigma(Y_t)$$

$$=\mathscr{F}_t^s \vee \sigma(\tilde{Y}_0)$$

Alors d'après un résultat de Ito [6], traduit en termes "rétrogrades", le processus β_s, défini par:

$$\beta_s=Y_s-Y_t+\int_s^t \frac{Y_\theta}{\theta}\,d\theta$$

est un $\mathring{P}-\bar{\mathscr{F}}_t^s$ Wiener rétrograde.

Compte-tenu d'un terme correctif dû à la dépendance de B_s par rapport à Y_s (qui s'introduit pour des raisons qui apparaîtront plus loin), on est amené à considérer l'E.D.P.S. rétrograde:

$$\left.\begin{aligned} &dv(s)+L_s v(s)\,ds+B_s v(s)\oplus d\beta_s \\ &\quad+\frac{Y_s}{s}B_s v(s)\,ds=C_s v(s)\,ds, \quad 0<s\leqq t \\ &v(t)=f \end{aligned}\right\} \tag{6.9}$$

où l'on suppose:

$$f\in C_0^1(R^N) \tag{6.10}$$

La dépendance en Y_s des opérateurs dans (6.9) ne pose pas de problèmes particulier pour l'existence et l'unicité d'une solution. Par contre, le terme multiplicatif Y_s/s affaiblit le résultat: $1/s$ interdit toute estimation au voisinage de 0, et Y_s détruit l'estimation que l'on avait au §3 pour les moments d'ordre 2 de v. On démontre à partir des résultats de [14, I° partie], par une technique de localisation par des temps d'arrêt.

THÉORÈME 6.1 *L'équation* (6.9) *a une solution unique* $\bar{\mathscr{F}}_t^s$*—adaptée:* $v \in L^2(\varepsilon, t; H^1) \cap C([\varepsilon, t]; L^2(R^N))$; $\forall \varepsilon > 0$, $\mathring{P}$ a.s.

Nous allons avoir besoin d'un résultat de convergence d'un schéma de discrétisation en t de (6.9). Soient $s \in]0, t[$, et $s = t_1 < t_2 < \ldots < t_n = t$, avec

$$t_{i+1} - t_i = \Delta t = \frac{t-s}{n}, i = 1, \ldots, n-1.$$

On note $\Delta\beta_i = \beta_{t_{i+1}} - \beta_{t_i}$, et on considère le schéma:

$$\left.\begin{aligned} & v_{i+1} - v_i + \left(\int_{t_i}^{t_{i+1}} L(r, Y_{t_i})\,dr\right)v_i + \left(\int_{t_i}^{t_{i+1}} B(r, Y_{t_{i+1}})\frac{dr}{\Delta t}\right)v_{i+1}\Delta\beta_i \\ & + \int_{t_i}^{t_{i+1}} \frac{Y_r}{r}\,dr\left(\int_{t_i}^{t_{i+1}} B(r, Y_{t_{i+1}})\frac{dr}{\Delta t}\right)v_{i+1} \\ & = \left(\int_{t_i}^{t_{i+1}} C(r, Y_{t_i})\,dr\right)v_{i+1}\, i = n-1, n-2, \ldots, 1; v_n = f \end{aligned}\right\} \tag{6.11}$$

On réécrit la première partie de (6.11) sous la forme abrégée:

$$v_{i+1} - v_i + \mathscr{L}^i v_i \Delta t + \mathscr{A}^i v_{i+1}\Delta t + \mathscr{B}^i v_{i+1}\Delta\beta_i = 0$$

Il résulte de (6.4) que pour Δt suffisamment petit,

$$I - \int_{t_i}^{t_{i+1}} L(r, Y_{t_i})\,dr$$

est un opérateur coercif, et que (6.11) définit alors une suite de v.a. v^i, $\bar{\mathscr{F}}_t^{t_i}$-mesurable à valeurs dans H^1. On définit:

$$v^n(\theta, x) = v_i(x), si \theta \in [t_i, t_{i+1}[\tag{6.12}$$

$$\rho_R = 1_{\{\sup_{s \leq \theta \leq t} |Y_\theta| \leq R\}} \tag{6.13}$$

On a alors le:

LEMME 6.2 *La suite* $(\{\rho_R v^n, n \in \mathbb{N}\}$ *reste dans un borné de* $L^2(\Omega; L^2(s, t; H^1)) \cap L^\infty(s, t; L^2(R^N))$.

Il existe une sous-suite v^m telle que:

$$\rho_R v^m(s) \rightarrow \rho_R v(s)$$

dans $L^2(\Omega \times R^N)$ faible

Preuve Posons

$$Y_\theta^R = Y_\theta \frac{|Y_\theta| \wedge R}{|Y_\theta|}$$

$$\mathscr{A}^{R_i} = \int_{t_i}^{t_{i+1}} \frac{Y_\theta^R}{\theta} d\theta \left(\int_{t_i}^{t_{i+1}} B(\theta, Y_{t_{i+1}}) \frac{d\theta}{t} \right).$$

On définit la suite v_i^R par:

$$\begin{cases} v_{i+1}^R - v_i^R + \mathscr{L}^i v_i^R \Delta t + \mathscr{A}^{R_i} v_{i+1}^R \Delta t = \mathscr{B}^i v_{i+1}^R \Delta \beta_i \\ i = n-1, n-2, \ldots, 1 \qquad\qquad v_n^R = f \end{cases}$$

On pose $v^{Rn}(\theta, x) = v_i^R(\theta, x)$, $si\,\theta \in [t_i, t_{i+1}[$

On montre, comme dans [14, II° partie, théorème 3.1], que v^{Rn} reste dans un borné de $L^2(\Omega; L^2(0, T; H^1) \cap L^\infty(0, T; L^2(R^N)))$, et qu'il existe une sous-suite $v^{Rm}(s)$ qui converge vers $v^R(s)$ dans $L^2(\Omega \times R^N)$ faible. v^R est la solution de:

$$dv^R(s) + L_s v^R(s)\, ds + B_s v^R(s)\, d\beta_s$$

$$+ \frac{Y_s^R}{s} B_s v^R(s)\, ds = C_s v^R(s)\, ds, \quad 0 \leqq s \leqq t$$

$$v^R(t) = f$$

Le lemme résulte alors de ce que $\rho_R v^n \equiv \rho_R v^{Rn}$ et $\rho_R v \equiv \rho_R v^R$.

LEMME 6.3 *Supposons, outre les hypothèses* (6.2)...(6.7), *que tous les coefficients qui apparaissent ont leurs dérivées de tous ordres par rapport à x bornées, et continues en* (x, y) *uniformément sur tout compact de* $R_+ \times R^N \times R^D$. *On suppose en outre que* $f \in C_0^\infty(R^N)$.

Alors la suite $\{\rho_R v^n, n \in \mathbb{N}\}$ *reste dans un borné de* $L^2(\Omega, L^\infty(0, T; C_b^2(R^N)))$.

Preuve Grâce aux hypothèses faites sur les coefficients, les dérivées partielles en x de tous ordres de $\{v_i^R; i=0,1,\ldots,n\}$ satisfont un système d'équations du type (6.14). On montre alors, grâce au lemme 6.2, que la suite v^{Rn} reste dans un borné de $L^2(\Omega; L^\infty(0,T;H^k))$, $\forall k \in \mathbb{N}$, où

$$H^k \triangleq \left\{ u \in L^2(R^N) : \frac{\partial_u^\alpha}{\partial x_1^{\alpha_1} \ldots \partial x_N^{\alpha_N}} \in L^2(R^N) \, \forall |\alpha| \leqq k \right\}$$

Or si $k > N/2+2$, H^k s'injecte continûment dans l'espace des fonctions de classe C^2, nulles à l'infini ainsi que leurs dérivées jusqu'à l'ordre 2—voir Lions [11].

6.3 Equations du filtrage et de la prédiction

Nous allons maintenant généraliser les résultats des théorèmes 3.1 et 3.2.

Il ne semble pas que les démonstrations de Krylov–Rosovskii [8] puissent s'adapter à la situation de ce parahraphe. Nous allons reprendre, en les adaptant, les démonstrations de [14].

PROPOSITION 6.4 *Outre* (6.2)...(6.7), *on suppose vérifiées les hypothèses du lemme* 6.3. *Alors* $\forall s \in]0,t], \forall x \in \mathbb{R}^N$,

$$v(s,x) = \mathring{E}_{sx}[f(X_t)Z_t^s \,|\, \mathscr{F}_t^s]$$

Preuve On définit ρ_R par (6.13). Soit $s = t_1 < t_2 < \ldots < t^n = t$, avec $t_{i+1} - t_i = (t-s)/n$. On pose:

$$\chi_i^n = Z_{t_{i+1}}^s v^n(t_{i+1}, X_{t_{i+1}}) - Z_{t_i}^s v^n(t_i, X_{t_i})$$

où v^n est défini par (6.11)-(6.12).

$$\rho_R E_{sx}^{\mathscr{F}_t^s}\left[\sum_{i=1}^{n-1} \chi_i^n\right] = \rho_R\{E_{sx}^{\mathscr{F}_t^s}[Z_t^s f(X_t)] - v^n(s,x)\}$$

D'après le lemme 6.2, $\forall R > 0$, la quantité ci-dessus converge vers

$$\rho_R\{E_{sx}^{\mathscr{F}_t^s}[Z_t^s f(X_t)] - v(s,x)\}$$

dans $L^2(\Omega \times R^N)$ faible. Nous allons maintenant établir que $\forall x \in R^N$, $\forall R > 0$, la même quantité tend vers 0 dans $L^1(\Omega, \mathscr{G}^s, \mathring{P}_{sx})$.

Ceci suffira à démontrer la proposition.

$$\chi_i^n = (Z_{t_{i+1}}^s - Z_{t_i}^s)[v^n(t_{i+1}, X_{t_{i+1}}) - v^n(t_{i+1}, X_{t_i})]$$

$$+ (Z_{t_{i+1}}^s - Z_{t_i}^s) v^n(t_{i+1}, X_{t_i})$$

$$+ Z_{t_i}^s [v^n(t_{i+1}, X_{t_{i+1}}) - v^n(t_{i+1}, X_{t_i})]$$

$$+ Z_{t_i}^s [v^n(t_{i+1}, X_{t_i}) - v^n(t_i, X_{t_i})]$$

$$= [v^n(t_{i+1}, X_{t_{i+1}}) - v^n(t_{i+1}, X_{t_i})] \int_{t_i}^{t_{i+1}} Z_r^s h(X_r, Y_r)\, dY_r$$

$$+ v^n(t_{i+1}, X_{t_i}) \int_{t_i}^{t_{i+1}} Z_r^s h(X_r, Y_r)\, dY_r$$

$$+ Z_{t_i}^s \nabla v^n(t_{i+1}, X_{t_i}) \cdot (X_{t_{i+1}} - X_{t_i}) + \tfrac{1}{2} Z_{t_i}^s T_r \left\{ \frac{\partial^2 v^n}{\partial x^2} (t_{i+1}, H_i)(X_{t_{i+1}} - X_{t_i})^{\otimes 2} \right\}.$$

$$- Z_{t_i}^s \mathscr{L}^i v^n(t_i, X_{t_i}) \Delta t + Z_{t_i}^s \mathscr{C}^i v^n(t_{i+1}, X_{t_i}) \Delta t - Z_{t_i}^s \mathscr{B}^i v^n(t_{i+1}, X_{t_i}) \Delta Y_i$$

où H_i appartient au segment de droite joignant X_{t_i} et $X_{t_{i+1}}$, et

$$\mathscr{C}^i = \int_{t_i}^{t_{i+1}} C(s, Y_{t_i}) \frac{ds}{\Delta t}.$$

En utilisant (6.8) on obtient:

$$\chi_i^n = A_i^n + B_i^n + C_i^n + D_i^n + E_i^n, \text{ avec:}$$

$$A_i^n = [v^n(t_{i+1}, X_{t_{i+1}}) - v^n(t_{i+1}, X_{t_i})] \int_{t_i}^{t_{i+1}} Z_r^s h(X_r, Y_r)\, dY_r$$

$$- Z_{t_i}^s \nabla v^n(t_{i+1}, X_{t_i}) \int_{t_i}^{t_{i+1}} c^* h(X_r, Y_r)\, dr$$

$$B_i^n = v^n(t_{i+1}, X_{t_i}) \int_{t_i}^{t_{i+1}} Z_r^s h(X_r, Y_r)\, dY_r$$

$$- Z_{t_i}^s v^n(t_{i+1}, X_{t_i}) \left[h^i(X_{t_i}, Y_{t_{i+1}}) \Delta Y_i - \sum_{k=1}^{D} \frac{\partial h_k^i}{\partial y_k} (X_{t_i}, Y_{t_i}) \Delta t \right]$$

$$C_i^n = Z_{t_i}^s \nabla v^n(t_{i+1}, X_{t_i}) \int_{t_i}^{t_{i+1}} b(X_r, Y_r)\, dr$$

$$- Z_{t_i}^s \nabla v^n(t_i, X_{t_i}) \int_{t_i}^{t_{i+1}} b(X_{t_i}, Y_{t_i})\, dr$$

$$D_i^n = Z_{t_i}^s \nabla v^n(t_{i+1}, X_{t_i}) \left[\int_{t_i}^{t_{i+1}} c^*(X_r, Y_r)\, dY_r + \int_{t_i}^{t_{i+1}} \tilde{c}\, d\tilde{Y}_r \right]$$

$$- Z_{t_i}^s \nabla v^n(t_{i+1}, X_{t_i}) \left[c^i(X_{t_i}, Y_{t_{i+1}}) \Delta Y_i \right.$$

$$\left. - \sum_{k=1}^{D} \frac{\partial c^i}{\partial y_k}(X_{t_i}, Y_{t_{i+1}}) \Delta t \right]$$

$$E_i^n = \tfrac{1}{2} Z_{t_i}^s Tr \left\{ \frac{\partial^2 v^n}{\partial x^2}(t_{i+1}, H_i)(X_{t_{i+1}} - X_{t_i})^{\otimes 2} \right\}$$

$$- \tfrac{1}{2} Z_{t_i}^s \sum_{j,l} \int_{t_i}^{t_{i+1}} a_{jl}(s, X_{t_i}, Y_{t_i}) \cdot ds \frac{\partial^2 v^n}{\partial x_j \partial x_l}(t_i, X_{t_i})$$

où

$$h^i(x, y) = \int_{t_i}^{t_{i+1}} h(s, x, y) \frac{ds}{\Delta t}, \; c^i(x, y) = \int_{t_i}^{t_{i+1}} c(s, x, y) \frac{ds}{\Delta t},$$

$$\Delta Y_i = Y_{t_{i+1}} - Y_{t_i}$$

On montre aisément que

$$\sum_1^n A_i^n, \sum_1^n C_i^n \text{ et } \sum_1^n E_i^n$$

tendent vers 0 dans $L^1(\Omega)$, quand $\Delta t \to 0$. Nous allons traiter l'expression

$$\sum_1^n B_i^n,$$

pour laquelle la démonstration est un peu plus délicate. Le terme

$$\sum_1^n D_i^n$$

se traite de façon similaire, après avoir remarqué que $\mathring{E}_{sx}(./\bar{\mathscr{F}}_t^s)$ de l'intégrale par rapport à $d\tilde{Y}$ est nulle.

L'expression suivante diffère de B_i^n par un terme en $o(\Delta t)$:

$$\bar{B}_i^n = v(t_{i+1}, X_{t_i}) \int_{t_i}^{t_{i+1}} [Z_r^s h(X_r, Y_r) - Z_{t_i}^s h^i(X_{t_i}, Y_{t_i})]\, dY_r$$

En appliquant la formule de Ito à l'intégrand ci-dessus, on obtient:

$$\bar{B}_i^n = v(_{t_{i+1}}, X_{t_i}) \int_{t_i}^{t_{i+1}} dY_r \left[\int_{t_i}^r \bar{a}_\mu\, d\mu + \int_{t_i}^r \bar{b}_\mu\, dY_\mu + \int_{t_i}^r \bar{c}_\mu\, d\tilde{Y}_\mu \right]$$

Le premier terme dans $\bar{B}_i^n$ est d'ordre de grandeur $o(\Delta t)$, et $\mathring{E}_{sx}(./\bar{\mathscr{F}}_t^s)$ du dernier terme est nulle. Il reste à considérer:

$$v(t_{i+1}, X_{t_i}) \left\{ \int_{t_i}^{t_{i+1}} dY_r \left(\int_{t_i}^r \bar{b}_\mu\, dY_\mu \right) \right\} = v(t_{i+1}, X_{t_i}) \left\{ \bar{b}_{t_i} [\Delta Y_i^2 - \Delta t] \right.$$

$$\left. + \int_{t_i}^{t_{i+1}} dY_r \left[\int_{t_i}^r (\bar{b}_\mu - \bar{b}_{t_i})\, dY_\mu \right] \right\}$$

$$E \left| v(t_{i+1}, X_{t_i}) \int_{y_i}^{t_{i+1}} dY_r \int_{t_i}^r (\bar{b}_\mu - \bar{b}_{t_i})\, dY_\mu \right|$$

$$\leqq \sqrt{E(|v(t_{i+1}, X_{t_i})|^2)} \times \sqrt{\int_{t_i}^{t_{i+1}} dr \int_{t_i}^r E(|\bar{b}_\mu - \bar{b}_{t_i}|^2)\, d\mu}$$

Ce terme est d'ordre de grandeur $o(\Delta t)$, car $\bar{b}_r$ est continue en moyenne quadratique.

Finalement, par un raisonnement classique:

$$\sum_{i=1}^n v(t_{i+1}, X_{t_i}) \bar{b}_{t_i} (\Delta Y_i^2 - \Delta t) \to 0 \quad \text{dans} \quad L^1(\Omega) \qquad \blacksquare$$

On a alors le:

THÉORÈME 6.5 *Sous les hypothèses du* §6.1 *et* (6.10), $\forall s \in]0, t]$, on a l'égalité suivante, $d\mathring{P} \times dx$ p.p.:

$$v(s, x) = \mathring{E}_{sx}[f(X_t) Z_t^s / \bar{\mathscr{F}}_t^s]$$

Preuve Supposons tout d'abord $f \in C_0^\infty(R^N)$.

Soit b^n, σ^n, h^n, g^n, $\tilde{g}^n$ une suite de coefficients qui vérifie les hypothèses de la proposition 6.4. On suppose que ces coefficients vérifient (6.2)...(6.7) avec des bornes et des constantes α et β indépendantes de n. On suppose que ces coefficients convergent respectivement vers b, σ, h, g, et $\tilde{g}$, uniformément en (t, x, y), et qu'il en est de même de

$$\frac{\partial \sigma_{ij}^n}{\partial x_i}, \sum_1^D \frac{\partial h_k^n}{\partial y_k} \quad \text{et} \quad \sum_1^D \frac{\partial c_{ik}^n}{\partial y_k}.$$

Alors $\mathring{P}_{sx}^n$ converge étroitement vers $\mathring{P}_{sx}$ (cf. Stroock–Varadhan [18]). Remarquons que les restrictions de $\mathring{P}_{sx}^n$ et $\mathring{P}_{sx}$ à $\bar{\mathscr{F}}_t^s$ coïncident avec la restriction à $\bar{\mathscr{F}}_t^s$ de $\mathring{P}$.

On montre par un raisonnement classique que $\rho_R v^n(s) \to \rho_R v(s)$ dans $L^2(\Omega \times R^N; d\mathring{P} \times dx)$.

Soit ξ une application continue de Ω dans R, $\bar{\mathscr{F}}_t^s$ mesurable, qui s'annule en dehors de l'ensemble

$$\left\{ \sup_{s \leq r \leq t} |Y_r| \leq R \right\}.$$

D'après la Proposition 6.4,

$$\mathring{E}[v^n(s, x)\xi] = \mathring{E}_{sx}^n[f(X_t)\,{}^n Z_t^s \xi]$$

$$= E_{sx}^n[f(X_t)\xi]$$

D'après ce qui précède, on peut passer à la limite, pour $n \to \infty$, dans l'égalité ci-dessus. Donc:

$$\mathring{E}[v(s, x)\xi] = \mathring{E}_{sx}[f(X_t) Z_t^s \xi] \; dx\text{-p.p.}$$

Donc, grâce à la latitude de choix de ξ et de R,

$$v(s, x) = \mathring{E}_{sx}[f(X_t) Z_t^s \mid \bar{\mathscr{F}}_t^s]$$

$d\mathring{P} \times d \times$ p.p., lorsque f est régulière. Le résultat général s'en déduit par passage à la limite. □

On introduit alors l'E.D.P.S. progressive:

$$\left.\begin{aligned} &dp(s) = L_s^* p(s)\, ds + B_s^* p(s)\, dY_s \\ &p(0) = p_0 \end{aligned}\right\} \tag{6.15}$$

Soient $s = t_0 < t_1 < \ldots < t_n = t$, comme ci-dessus. On considère le schéma d'approximation de (6.15):

$$\left.\begin{aligned} &p_{i+1} - p_i = \left(\int_{t_{i+1}}^{t_{i+2}} L^*(r, Y_{t_{i+1}})\, dr\right) p_{i+1} + \left(\int_{t_1}^{t_{i+1}} B^*(r, Y_{t_i}) \frac{dr}{\Delta t}\right) p_i \Delta Y_i \\ &p_1 = \int_{s-\Delta t}^{s} p(r) \frac{dr}{\Delta t} \end{aligned}\right\} \tag{6.16}$$

(6.16) définit une suite de v.a. p_i, $\mathscr{F}_{t_i}$ mesurable a valeurs dans H^1. On pose:

$$p^n(r) = p_i, \quad \text{si} \quad r \in [t_i, t_{i+1}[$$

On peut alors montrer, par la méthode du Lemme 6.2:

LEMME 6.6 *p^n reste dans un borné de $L^2(\Omega; L^2(s, t; H^1))$. Il existe une sous suite $p^m(t)$ de $p^n(t)$, telle que:*

$$p^m(t) \to p(t)$$

dans $L^2(\Omega \times R^N)$ faible.

Multipliant scalairement (6.11) par p_i, et (6.16) par v_{i+1}, et itérant de $i = 1$ à $n-1$, on obtient:

$$(p_n, v_n) - \left\langle \left(\int_t^{t+\Delta t} L(r, Y_t)\, dr\right) v_n, p_n \right\rangle = (p_1, v_1)$$

$$- \left\langle \left(\int_s^{s+\Delta t} L(r, Y_s)\, dr\right) v_1, p_1 \right\rangle + \sum_{i=1}^{n-1} \left\langle \left[\left(\int_{t_i}^{t_{i+1}} B(s, Y_{t_i}) \frac{ds}{\Delta t}\right) \Delta Y_i\right.\right.$$

$$\left.\left. + \left(\int_{t_i}^{t_{i+1}} C(s, Y_{t_i})\, ds\right) - \left(\int_{t_i}^{t_{i+1}} B(s, Y_{t_{i+1}}) \frac{ds}{\Delta t}\right) \Delta Y_i \right] v_{i+1}, p_i \right\rangle$$

Grâce aux lemmes 6.2 et 6.6, on peut passer à la limite dans l'égalité ci-dessus, d'où l'on tire:

$$(p(t), f) = (p(s), v(s)) \quad \text{p.s.,} \quad \forall s \in]0, t]$$

On a montré le:

THÉORÈME 6.7 *Le processus* $(p(s), v(s))$, $s \in]0, t]$, *est p.s. constant.*

On peut maintenant établir le:

THÉORÈME 6.8 $\forall t > 0, p(t, x)(p(t), 1)^{-1}$ *est la densité de la loi conditionnelle de* X_t, *sachant* $\mathscr{F}_t$.

Preuve Il suffit de nouveau de montrer que $\forall t \geqq 0$, $\forall f \in C_1(R^N)$,

$$\mathring{E}[f(X_t)Z_t | \mathscr{F}_t] = (p(t), f) \tag{6.17}$$

Fixons $f \in C_1(R^N)$. On notera $v(s, x)$ la solution correspondante de l'équation (6.9). Fixons alors $t > 0$, et soit $\varepsilon \in]0, t[$. Par un argument similaire à celui de Lemme 3.9,

$$\mathring{E}(f(X_t)Z_t^\varepsilon / \mathscr{F}_t) = \mathring{E}^{\mathscr{F}_t}[\mathring{E}^{\mathring{\mathscr{F}}_t^\varepsilon}_{\varepsilon, X_\varepsilon, Y_\varepsilon}(f(X_t)Z_t^\varepsilon)] = \mathring{E}^{\mathscr{F}_t}[\mathring{E}^{\mathscr{F}_t^\varepsilon}_{\varepsilon, X_\varepsilon}(f(X_t)Z_t^\varepsilon)]$$

Il résulte alors du théorème 6.2:

$$\mathring{E}(f(X_t)Z_t^\varepsilon / \mathscr{F}_t) = \mathring{E}(v(\varepsilon, X_\varepsilon) / \mathscr{F}_t) \tag{6.18}$$

Soient P^ε et $\mathring{P}^\varepsilon$ des mesures de probabilité définies comme P et $\mathring{P}$, avec les coefficients h, g et $\tilde{g}$ remplacés par:

$$h_\varepsilon = 1_{\{s \geqq \varepsilon\}} h; \, g_\varepsilon = 1_{\{s \geqq \varepsilon\}} g; \, \tilde{g}_\varepsilon = 1_{\{s < \varepsilon\}} I + 1_{\{s \geqq \varepsilon\}} \tilde{g}$$

Soit alors p_ε la solution de:

$$\left. \begin{array}{l} dp_\varepsilon(s) = L_s^* p_\varepsilon(s)\, ds + 1_{\{s \geqq \varepsilon\}} B_s^* p_\varepsilon(s) . \, dY_s \\ p_\varepsilon(0) = p_0 \end{array} \right\}$$

On vérifie aisément que, sous la loi $\mathring{P}^\varepsilon$, $p_\varepsilon(\varepsilon)$ est la densité de la loi conditionnelle de X_ε, sachant $\mathscr{F}_\varepsilon$. Mais (6.11) est encore vraie si l'on remplace $\mathring{E}$ par $\mathring{E}^\varepsilon$. Donc, par un argument similaire à celui de Lemme 3.10,

$$\mathring{E}^\varepsilon(f(X_t)Z_t^\varepsilon / \mathscr{F}_t) = (p_\varepsilon(\varepsilon), v(\varepsilon)).$$

De façon analogue au théorème 6.7, on vérifie que le processus $(p_\varepsilon(s), v(s))$ est p.s. constant sur l'intervalle $[\varepsilon, t]$.
Donc:

$$\mathring{E}^\varepsilon(f(X_t)Z_t^\varepsilon/\mathscr{F}_t) = (p_\varepsilon(t), f) \tag{6.19}$$

Il reste à passer à la limite dans (6.19). Soit $\varphi \in C_b(\Omega, R)$, $\mathscr{F}_t$ mesurable. Il résulte de (6.19):

$$\mathring{E}^\varepsilon(f(X_t)Z_t^\varepsilon\varphi) = \mathring{E}^\varepsilon[(p_\varepsilon(t), f)\varphi]$$

Soit:

$$E^\varepsilon[f(X_t)\varphi] = \mathring{E}[(p_\varepsilon(t), f)\varphi] \tag{6.20}$$

Mais P^ε converge étroitement vers P, lorsque $\varepsilon \to 0$-voir Stroock–Varadhan [18]. De plus, on montre par un raisonnement classique:

$$p_\varepsilon(t) \to p(t) \quad \text{dans} \quad L^2(\Omega, \mathscr{F}_t, \mathring{P}) \quad \text{quand} \quad \varepsilon \to 0.$$

On peut donc passer à la limite dans (6.20), quand $\varepsilon \to 0$, d'où:

$$\mathring{E}[f(X_t)Z_t\varphi] = \mathring{E}[(p(t), f)\varphi]$$

(6.17) résulte alors de la latitude de choix de φ. □

La solution du problème de prédiction résulte de celle du problème de filtrage, par l'argument déjà utilisé au §3.

6.4 Equations du lissage

Soient p et v les solutions des E.D.P.S:

$$\left.\begin{aligned} &dp(r) = L_r^* p(r)\,dr + B_r^* p(r)\,dY_r, r \geqq 0 \\ &p(0) = p_0 \end{aligned}\right\} \tag{6.21}$$

$$\left.\begin{aligned} &dv(r) + L_r v(r)\,dr + B_r v(r) \oplus d\beta_r + \frac{Y_r}{r} B_r v(r)\,dr = C_r v(r)\,dr, \quad 0 \leqq r \leqq t \\ &v(t) = 1 \end{aligned}\right\} \tag{6.22}$$

En combinant les raisonnements du §6.2 et du §3.3, on obtient que l'équation (6.22) a une solution unique $v(r)\mathscr{F}_t^r$ adaptée, avec:

$$v \in L^2(\varepsilon, t; H_\rho^1) \cap C([\varepsilon, t]; L_\rho^2) \text{ p.s., } \forall \varepsilon > 0$$

THÉORÈME 6.9 $\forall s \in]0,t]$, $p(s,x)v(s,x)(p(s),v(s))^{-1}$ *est la densité de la loi conditionnelle de* X_s, *sachant* $\mathscr{F}_t$.

Preuve: En adaptant les démonstrations des Lemmes 3.9 et 3.10, en utilisant le théorème 6.5 avec $f=1$ et (6.17), ainsi que l'égalité $\mathring{E}_{sX_sY_s}^{\mathscr{F}_t^s} = \mathring{E}_{sX_s}^{\mathscr{F}_t^s}$, on montre que $\forall f \in C_1(R^N)$,

$$\mathring{E}(f(X_s)Z_t \mid \mathscr{F}_t) = \mathring{E}^{\mathscr{F}_t}(Z_s f(X_s) Z_t^s)$$

$$= \mathring{E}^{\mathscr{F}_t}[Z_s f(X_s) \mathring{E}_{sX_sY_s}^{\mathscr{F}_t^s}(Z_t^s)]$$

$$= \mathring{E}^{\mathscr{F}_t}[Z_s f(X_s) v(s, X_s)]$$

$$= (p(s), v(s) f)$$

Mais alors $p(s,x)v(s,x) \geqq 0$ p.p. et p.s., donc par convergence monotone:

$$\mathring{E}(Z_t / \mathscr{F}_t) = (p(s), v(s)) \tag{6.17}$$

Ici, $(.,.)$ désigne le produit de dualité entre $L^1(R^N)$ et $L^\infty(R^N)$. On montre en effet à partir de (6.17) et du Théorème 6.5:

$$p(s) \in L^1(R^N), v(s) \in L^\infty(R^N), \quad \text{p.s.}$$

Le Théorème découle alors du Lemme 2.1. ■

Remarque 6.6 On peut montrer que quand $s \downarrow 0, v(s,x)$ converge vers $v(0,x) = \mathring{E}_{0x}(Z_t / \mathscr{F}_t)$. Le résultat du Théorème 6.9 s'étend alors au cas $s = 0$. ■

Bibliographie

[1] A. Bensoussan and J. L. Lions, *Application des inéquations variationnelles en contrôle stochastique*. Dunod (1978).

[2] J. M. C. Clark, The design of robust approximations to the stochastic differential equations of non linear filtering in *Communication Systems and Random Process Theory*, Ed. J. Skwirzynski, Sijthoff and Noordhoof (1978).

[3] M. H. A. Davis, On a multiplicative functional transformation arising in non-linear filtering theory, *Z. Wahrscheinlichkeitstheorie ver. Geb.* **54** (1981), 125–139.

[4] H. Doss, Liens entre équations différentielles stochastiques et ordinaires, *Ann. Inst. H. Poincaré* **13** (1977), 99–125.

[5] W. Fleming and E. Pardoux, Optimal control for partially observed diffusions, *Soumis à Siam J. Control.*

[6] K. Ito, Extension of stochastic integrals. *in Proc. Internat. Symp. Stoch Diff. Equ.*, Kyoto 1976, Ed. K. Ito, Wiley (1978).

[7] N. Krylov and B. Rozovskii, On conditional distributions of diffusion processes. *Math USSR Izvestija* **12,** (1978), 336–356.

[8] N. Krylov and B. Rozovskii, On the first integrals and Liouville equations for diffusion processes. to appear in Proc. IFIP Conf. on Stoch. Diff. Equ. held at Visegrad (1980).

[9] H. Kunita, Non linear filtering for the system with general noise. in *Stoch. Cont. Theory and Stoch. Diff. Syst.*, Eds M. Kohlmann and W. Vogel, Lecture Notes in Control and Inf. Sciences, 16 Springer (1979).

[10] H. Kunita, Stochastic differential equations arising from non linear filtering. Preprint.

[11] J. L. Lions, *Problèmes aux limites dans les E.D.P.* Presses de l'Univ. de Montréal (1965).

[12] R. Liptzer and A. Shiryaev, *Statistics of random processes*. Springer-Verlag, Berlin-Heidelberg-New York, 1977.

[13] R. Liptzer and A. Shiryaev, Non linear interpolation of components of Markov processes. *Th. Prob. Appl.* **13,** (1968).

[14] E. Pardoux, Stochastic partial differential equations and filtering of diffusion processes, *Stochastics* **3,** (1979), 127–167.

[15] E. Pardoux, Backward and forward stochastic partial differential equations associated with a non linear filtering problem. in *Proc. IEEE 18th Conf on Dec and Cont.* (1979).

[16] E. Pardoux, Non linear filtering, prediction and smoothing. in *Stochastic Systems*, Ed. M. Hazewinkel and J. C. Willems, D. Reidel, Dordrecht, 1981.

[17] B. Rozovskii, Equations différentielles stochastiques aux dérivées partielles, intervenant dans les problèmes de filtrage non-linéaire, (*en Russe*) *UMNXXVII*, **3** (1972), 213–214.

[18] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*.Springer-Verlag, Berlin-Heidelberg-New York, 1979.

[19] H. Sussmann, On the gap between deterministic and stochastic ordinary differential equations. *Ann. of Prob.* **6,** (1978), 19–41.

[20] J. Szpirglas and G. Mazziotto, Modèle général de filtrage non linéaire et équations différentielles stochastiques associées. *C. R. Acad. Sciences* Paris t. 286, 1067–1070 (1978).

[21] M. Zakai, On the optimal filtering of dissusion processes. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **11** (1969), 230–243.