ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE

PIERRE-LOUIS LIONS

Le problème de Cauchy pour les équations de Hamilton-Jacobi-Bellman

Annales de la faculté des sciences de Toulouse 5e série, tome 3, nº 1 (1981), p. 59-68

<http://www.numdam.org/item?id=AFST_1981_5_3_1_59_0>

Annales Faculté des Sciences Toulouse
Vol III, 1981, P. 59 à 68

LE PROBLEME DE CAUCHY POUR LES EQUATIONS DE HAMILTON-JACOBI-BELLMAN

Pierre-Louis Lions (1)

(1) Laboratoire d'Analyse Numérique, Université Pierre et Marie Curie, Tour 55-65, 4 Place Jussieu 75230 Paris Cédex 05.

Résumé : Dans cet article est résolu le problème de Cauchy pour les équations de Hamilton-Jacobi-Bellman, dans le cas uniformément elliptique. La démonstration utilise des techniques d'estimations a priori basées sur le principe du maximum ainsi que des arguments probabilistes.

Summary : In this paper the Cauchy problem for Hamilton-Jacobi-Bellman equations is solved in the uniformly elliptic case. The proof uses a priori estimates techniques based upon maximum principle and probabilistic arguments.

INTRODUCTION

Le but de cet article est la résolution du problème de Cauchy suivant :

$$
(1) \quad \left\{ \begin{array}{l} \dfrac{du}{dt} + \sup\limits_{v \in V} \left\{ A(t,v)\, u - f(t,v) \right\} = 0 \quad \text{p.p. dans } \mathcal{O} \times \,]0,T[\\[2ex] u(t,x) = 0 \text{ sur } \partial\mathcal{O} ;\; u(o,x) = u_0(x) \,; \end{array} \right.
$$

où les opérateurs $A(t,v)$ sont des opérateurs elliptiques du 2ème ordre dépendant de la variable t et d'un paramètre v.

Le problème (1) intervient dans le contrôle optimal de diffusions non-homogènes. Nous détaillons dans la section I le problème de minimisation stochastique associé à (1).

La version elliptique du problème (1) a fait l'objet d'une étude précédente de l'auteur [7] , et nous montrons dans la section II comment les arguments de [7] peuvent être adaptés et modifiés pour traiter le cas parabolique.

Enfin remarquons que certaines équations paraboliques du type Hamilton-Jacobi-Bellman sont étudiées dans [4] , [13] , un problème «parabolique-elliptique» étant traité dans [10] .

I - LE PROBLEME DE CONTROLE STOCHASTIQUE

I.1 - Hypothèses, notations

Soit $\mathcal{O}$ un ouvert borné régulier de $\mathbb{R}^N$ et soit $T > 0$. Soit V un ensemble convexe fermé de $\mathbb{R}^p$. Enfin, soient $\sigma_{ij}(t,x,v)$, $b_i(t,x,v)$, $c(t,x,v)$, $f(t,x,v)$ des fonctions de $[0,T] \times \bar{\mathcal{O}} \times V$ vérifiant

$$(2)\quad \begin{cases} \varphi(t,x,v) \text{ (resp. } f(t,x,v)) \text{ demeure dans un borné de } C^{2,1}(\mathcal{O} \times [0,T]) \\ \text{(resp. } W^{2,1,\infty}(\mathcal{O} \times]0,T[\,)) \text{ quand v décrit V ; où } \varphi = \sigma_{ij}, b_i, c, \end{cases}$$

$$(2')\quad \begin{cases} \exists \rho \in C(V,\mathbb{R}_+)\,,\ \rho(o) = 0 \text{ telle que} \\ |\varphi(t,x,v) - \varphi(t,x,v')| \leqslant \rho(|v-v'|)\ \forall (t,x) \in [0,T] \times \mathcal{O},\ \forall v, v' \in V, \\ \qquad\qquad \text{où } \varphi = \sigma_{ij}, b_i, c, f\,; \end{cases}$$

$$(3)\quad \begin{cases} \exists \nu > 0\ \forall (t,x,v) \in [0,T] \times \mathcal{O} \times V,\ \sum\limits_{i,j,k} \frac{1}{2} \sigma_{ik}\, \sigma_{jk}(t,x,v)\, \xi_i \xi_j \geqslant \nu\, |\xi|^2, \\ \qquad\qquad \text{pour tout } \xi \in \mathbb{R}^N. \end{cases}$$

Nous poserons $a(t,x,v) = \frac{1}{2}\ \sigma(t,x,v)\ .\ \sigma^T(t,x,v)$, où σ^T désigne la transposée de σ. Enfin nous supposons la donnée initiale dans $W^{2,\infty}(\mathcal{O})$:

$$(4)\qquad u_o \in W^{2,\infty}(\mathcal{O}) \cap W_o^{1,\infty}(\mathcal{O})\,.$$

Le problème que nous allons considérer n'est pas (1) mais

$$(1') \quad \begin{cases} -\dfrac{du}{dt} + \sup\limits_{v \in V} \{A(t,v)u - f(t,v)\} = 0 \quad \text{p.p. dans } \Theta \times \,]0,T[\\[2ex] u(t,x) = 0 \text{ sur } \partial\Theta;\ u(T,x) = u_0(x)\ ; \end{cases}$$

l'opérateur $A(t,v)$ étant donné par

$$A(t,v) = \sum_{i,j} -a_{ij}(t,x,v)\frac{\partial^2}{\partial x_i \partial x_j} - \sum_i b_i(t,x,v)\frac{\partial}{\partial x_i} + c(t,x,v)\,.$$

Il est clair que les problèmes (1) et (1') sont équivalents à un renversement du temps près : $\varphi(t,x) \to \Psi(t,x) = \varphi(T-t,x)$ pour $\varphi = u, a, b, c, f$.

I.2 - Le problème de contrôle stochastique

On définit un système admissible $\mathcal{A}$ comme étant la donnée de

$$\mathcal{A} = (\Omega, F, F_t, P, W_t, v(t), y_{x,s}(t))$$

où i) (Ω, F, F_t, P, W_t) est un espace de Wiener, ii) $v(t)$ est un processus non anticipatif à valeurs dans V, iii) $(y_{s,x}(t))$ est la solution de

$$(6) \quad \begin{cases} dy_{s,x}(t) = \sigma(t,y_{s,x}(t), v(t))\, dW_t + b(t,y_{s,x}(t), v(t))dt \quad \text{si } t \geq s\,. \\[2ex] y_{s,x}(s) = x\,. \end{cases}$$

A tout système admissible, on associe la fonction coût suivante :

$$J(t,x,\mathcal{A}) = E\left[\int_t^{\tau_{t,x} \wedge T} f(s,y_{t,x}(s), v(s)) \exp\left(-\int_0^s c(\lambda,y_{t,x}(\lambda),v(\lambda))d\lambda\right) ds\right]$$

où $\tau_{t,x}$ est le temps de sortie de Θ du processus $y_{t,x}(t)$.

Le problème de minimisation est alors :

$$(6) \qquad u(t,x) = \inf_{\mathcal{A} \text{ admissible}} J(t,x,\mathcal{A}) .$$

Heuristiquement, le raisonnement de la programmation dynamique indique que u doit être solution de (1'). Le raisonnement est valide si on sait a priori que u est régulière (voir [4]).

Notre démarche est la suivante : nous montrons tout d'abord que (1') admet une solution dans $W^{2,1,\infty}(\mathcal{O} \times]0,T[)$, puis que toute solution de (1') coïncide avec la fonction coût optimum. Aussi montrons-nous que le problème (1') admet une unique solution dans $W^{2,1,\infty}(\mathcal{O} \times]0,T[)$, et que celle-ci coïncide avec (6).

II - RESOLUTION DU PROBLEME DE CAUCHY

II.1 - Le résultat de résolution

THEOREME II.1. *Sous les hypothèses (1), (2), (3), (4), il existe une unique solution* u(x) *dans* $W^{2,1,\infty}(\mathcal{O} \times]0,T[)$ *de*

$$(1') \quad \begin{cases} -\dfrac{du}{dt} + \sup\limits_{v \in V} \{A(t,v)u - f(t,v)\} = 0 \text{ p.p. dans } \mathcal{O} \times]0,T[\\[2ex] u(t,x) = 0 \text{ sur } \partial\mathcal{O},\ u(T,x) = u_0(x) . \end{cases}$$

De plus u(t,x) *est donnée par*

$$u(t,x) = \inf_{\mathcal{A} \text{ admissible}} J(t,x,\mathcal{A}) .$$

Remarque II.1. Le résultat est conservé si au lieu de considérer un ensemble de contrôle V continu, on considère une famille dénombrable : on supprime alors l'hypothèse (2) et tout le reste est conservé.

Remarque II.2. Le théorème reste exact si on suppose $\mathcal{O}$ non borné. Il convient alors de remplacer dans (2) $C^{2,1}(\overline{\mathcal{O}} \times [0,T])$ par $C_b^{2,1}(\overline{\mathcal{O}} \times [0,T])$ (espace des fonctions bornées admettant des dérivées secondes bornées en x et la dérivée en t bornée).

Remarque II.3. En utilisant la méthode de N.V. Krylov [5] ainsi que l'estimation correspondant au cas parabolique de [6] , on peut montrer que l'unicité a lieu dans la classe $W^{2,1,N+1}_{loc}(\mathcal{O} \times]0,T[) \cap C(\overline{\mathcal{O}} \times [0,T])$.

Remarque II.4. Nous avons de plus l'estimation suivante

$$\| u \|_{W^{2,1,\infty}} \leqslant C \left\{ \sup_{v \in V} \| f(v) \|_{W^{2,1,\infty}} + \| u_o \|_{W^{2,\infty}} \right\}$$

où C ne dépend ni de u_o, ni des $f(v)$.

II.2. Démonstration

La démonstration suit celle de P.L. Lions [7] . Aussi n'allons nous présenter que les modifications qui surviennent dans le cas parabolique.

Rappelons que la méthode est de considérer (pour m fixé et $v_1 \dots v_m$ fixés dans V) le système pénalisé suivant : soit $\epsilon > 0$ (on convient que $1 \equiv m + 1$)

$$\begin{cases} -\dfrac{du^i_\epsilon}{dt} + A^i u^i_\epsilon + \beta_\epsilon(u^i_\epsilon - u^{i+1}_\epsilon) = f^i \quad \text{p.p. dans } \mathcal{O} \times [0,T[\\ u^i_\epsilon(t,x) = 0 \text{ sur } \partial\mathcal{O};\ u^i_\epsilon(T,x) = u_o(x)\ ; \end{cases}$$

où $A^i = A(v_i)$, $f^i = f(v_i)$ et β_ϵ est une fonction satisfaisant à : β_ϵ est croissante, convexe, de classe C^∞, $\beta_\epsilon(t) = 0$ si $t \leqslant 0$, $\beta_\epsilon(t)$ croît vers $+\infty$ quand $\epsilon \to 0$ si $t > 0$.

Nous distinguerons deux étapes dans la démonstration :

i) l'estimation $\| u^i_\epsilon \|_{W^{2,1,\infty}}$ uniforme en ϵ et m,

ii) le passage à la limite et l'unicité.

i) La partie estimation de la démonstration est pratiquement inchangée par rapport à [7] .

En suivant la démonstration donnée dans [7] , on constate que l'estimation $L^\infty(\mathcal{O} \times]0,T[)$ est conservée ainsi que :

$$| \nabla u^i_\epsilon(t,x) | \leqslant \text{Const.}, \quad \forall x \in \partial\mathcal{O}.$$

En admettant provisoirement le fait que

$$(7) \qquad \left| \frac{\partial u^i_\epsilon}{\partial t}(0,x) \right| \leqslant \text{Const.}, \quad \forall\, x \in \partial\mathcal{O}'$$

il est facile de déduire comme dans [7], en considérant la fonction auxiliaire

$$w = \left(\frac{\partial u}{\partial t}\right)^2 + |\nabla u|^2,$$

que nous avons

$$\| u^i_\epsilon \|_{W^{1,\infty}(\mathcal{O}' \times]0,T[)} \leqslant \text{Const.}.$$

Ensuite le reste de la démonstration s'applique et nous obtenons comme en [7] (l'hypothèse des coefficients d'ordre zéro assez grands devenant inutile dans le cas parabolique) que

$$\| u_\epsilon(t) \|_{W^{2,\infty}(\mathcal{O}')} \leqslant \text{Const.}, \quad \forall\, t \in [0,T],$$

ce qui implique l'estimation $W^{2,1,\infty}(\mathcal{O}' \times]0,T[)$ annoncée.

Il ne nous reste donc qu'à montrer (7) : or si nous posons $\overline{w} = u_0(x) + \lambda(T-t)$, alors pour λ assez grand

$$-\frac{\partial \overline{w}}{\partial t} + A^i\, \overline{w} \geqslant f^i \text{ sur } \overline{\mathcal{O}'} \times [0,T]$$

$$\overline{w}(t,x)\,|_{\partial\mathcal{O}'} \geqslant 0, \quad \overline{w}(T,x) = u_0(x).$$

Donc par une application élémentaire du principe du maximum (comme dans [7]) nous en déduisons

$$u^i_\epsilon(t,x) \leqslant \overline{w}(t,x).$$

De même pour $\underline{w} = u_0(x) - \lambda(T-t)$, avons-nous $u^i_\epsilon(t,x) \geqslant \underline{w}(t,x)$.

Et de ces inégalités découle : $\left| \dfrac{\partial u^i_\epsilon}{\partial t}(T,x) \right| \leqslant \lambda.$

Ainsi nous avons montré

$$(8) \qquad \| u^i_\epsilon \|_{W^{2,1,\infty}(\mathcal{O}' \times]0,T[)} \leqslant \text{Const. (indépendante de } \epsilon \text{ et de m).}$$

ii) Du fait que $\beta_\epsilon \geqslant \beta_\eta$, si $\epsilon \leqslant \eta$, on déduit que u^i_ϵ décroit quand ϵ décroit vers 0. Donc $u^i_\epsilon(t,x) \downarrow u^i(t,x) \in W^{2,1,\infty}(\Theta' \times]0,T[)$.

De plus de par le choix de $\beta_\epsilon(\beta_\epsilon(t) \uparrow +\infty$ si $t \geqslant 0)$, nous avons $u^i(x) = u(x)$ (indép. de i). Nous noterons $u(t,x) = u_m(t,x)$ (pour indiquer la dépendance par rapport à $(v_1,\ldots,v_m)$).

Comme d'autre part $\beta_\epsilon(t) \geqslant 0$, nous avons bien sûr

$$-\frac{du_m}{dt} + A^i u_m \leqslant f^i \quad \text{sur} \quad \Theta' \times]0,T[\,.$$

Maintenant on considère (v_m), suite dense dans V, et pour tout m, u_m correspondra au choix des m premiers termes de la suite. D'après (8) u_m est dans un borné de $W^{2,1,\infty}(\Theta' \times]0,T[)$ donc il existe u et une sous-suite encore notée u_m telles que $u \in W^{2,1,\infty}(\Theta' \times]0,T[)$ et $u_m \to u$ dans $C(\overline{\Theta}' \times [0,T])$ (ou dans $W^{2,1,\infty}(\Theta' \times]0,T[)$ faible).

En résumé, nous avons donc :

$$\overline{u}(t,x) = \lim_m u_m(t,x) = \lim_m \lim_\epsilon \downarrow u^i_{\epsilon,m}(t,x)\,,$$

$$\overline{u} \in W^{2,1,\infty}(\Theta' \times]0,T[),\ \overline{u}\,|_{\partial\Theta'} = 0,\ \overline{u}(T,x) = u_o(x)$$

$$\forall\, v \in V \quad -\frac{\partial\overline{u}}{\partial t} + A(v)\overline{u} \leqslant f(v) \ \text{ dans } \Theta' \times]0,T[\,.$$

Pour achever la démonstration nous allons raisonner de la façon suivante : (a) montrer que $\overline{u} = u$ où u est donnée par (6), (b) en déduire que $\overline{u}$ satisfait (1'), (c) en déduire que toute solution de (1') coïncide avec u.

(a) Comme $\overline{u} \in W^{2,1,\infty}(\Theta' \times]0,T[)$, il est facile de montrer par un argument de classe monotone que la formule de Itô est valable pour $\overline{u}(s,y_{t,x}(s))$ et on obtient :

$$\overline{u}(t,x) \leqslant J(t,x,\mathcal{A}) \text{ pour tout système admissible } \mathcal{A}\,.$$

Donc $\overline{u} \leqslant u$.

Maintenant, nous introduisons $(u^i_{k,m})$ solution du système d'inéquations quasi-variationnelles suivant (voir [3], [12]) avec $k > 0$ fixé.

$$(9) \quad \begin{cases} -\dfrac{\partial u_{k,m}}{\partial t} + A^i u^i_{k,m} \leqslant f^i \\ \qquad u^i_{k,m} \leqslant k + u^{i+1}_{k,m} \quad \text{p.p. sur } \Theta' \times]0,T[\\ \left(-\dfrac{\partial u_{k,m}}{\partial t} + A^i u^i_{k,m} - f^i\right)\left(u^i_{k,m} - k - u^{i+1}_{k,m}\right) = 0\,, \\ u^i_{k,m}(t,x) = 0 \ \text{ sur } \ \partial\Theta',\ \ u^i_{k,m}(T,x) = u_o(x). \end{cases}$$

Alors l'interprétation stochastique de $u^i_{k,m}$ (qui s'obtient comme dans [3] et [12]) montre que $u^i_{k,m}$ est une fonction coût optimum portant sur une classe de systèmes admissibles incluse dans celle considérée en (6).

Donc nous avons :

$$u(t,x) \leqslant u^i_{k,m}(t,x) \qquad (\forall\ i,k,m)\,.$$

D'autre part il est standard (voir [12]) que $u^i_{k,m} = \lim\limits_{\epsilon \downarrow 0} \downarrow u^i_{\epsilon,k,m}$ où $(u^i_{\epsilon,k,m})$ est solution du système pénalisé avec $\beta_\epsilon(u^i_\epsilon - u^{i+1}_\epsilon)$ remplacé par $\beta_\epsilon(u^i_\epsilon - k - u^{i+1}_\epsilon)$.

Donc nous avons

$$u_m(t,x) = \lim_{\epsilon \downarrow 0} \downarrow u^i_{\epsilon,m}(t,x) = \lim_{\epsilon \downarrow 0} \downarrow \lim_{k \downarrow 0} \downarrow u^i_{\epsilon,k,m}(t,x)\ ;$$

d'où

$$u_m(t,x) = \lim_{k \downarrow 0} \downarrow \lim_{\epsilon \downarrow 0} \downarrow u^i_{\epsilon,k,m}(t,x) = \lim_{k \downarrow 0} u^i_{k,m}(t,x) \geqslant u(t,x)\,.$$

En conclusion $\overline{u}(t,x) \geqslant u(t,x)$ et finalement

$$\overline{u}(t,x) \equiv u(t,x)\,.$$

(b) Nous savons donc que $u(t,x) \in W^{2,1,\infty}(\mathcal{O} \times]0,T[\,)$. Nous pouvons alors appliquer la méthode probabiliste de [8] pour montrer que (1') est satisfaite. La démonstration s'adapte telle quelle et nous obtenons

$$(1')\quad \begin{cases} -\dfrac{\partial u}{\partial t} + \sup\limits_{v \in V} \{A(t,v)u - f(t,v)\} = 0 \quad \text{p.p. dans } \mathcal{O} \times]0,T[\\ \\ u(t,x) = 0 \text{ sur } \partial\mathcal{O},\ u(T,x) = u_o(x)\,. \end{cases}$$

(c) L'unicité découle immédiatement d'un résultat de N.V. Krylov [6] .

En effet si u désigne une autre solution dans $W^{2,1,\infty}(\mathcal{O} \times]0,T[\,)$ de (1'), il existe $\alpha_{ij}(t,x)$, $\beta_i(t,x)$, $\gamma(t,x)$, bornées mesurables sur $\mathcal{O} \times]0,T[$ telles que (α_{ij}) soit définie positive uniformément en (t,x) dans $\mathcal{O}\times]0,T[$ et

$$\begin{cases} -\dfrac{\partial(u-\tilde{u})}{\partial t} - \sum\limits_{i,j} \alpha_{ij}(t,x) \dfrac{\partial^2(u-\tilde{u})}{\partial x_i\, \partial x_j} + \sum\limits_i \beta_i(t,x) \dfrac{\partial(u-\tilde{u})}{\partial x_i} + \gamma(t,x)(u-\tilde{u}) \leqslant 0 \\ \qquad\qquad \text{p.p. dans } \mathcal{O} \times]0,T[\\ (u-\tilde{u})(t,x) = 0 \ \text{ sur } \partial\mathcal{O},\ (u-\tilde{u})(T,x) = 0. \end{cases}$$

Donc d'après N.V. Krylov [6] , on en déduit : $u \leqslant \tilde{u}$ p.p. Mais d'après la formule de Itô, comme en (a), on a : $\tilde{u} \leqslant u$ p.p.
Ceci achève la démonstration du théorème.

Remarque II.5. Il serait intéressant de montrer directement le passage à la limite par une méthode analytique, c'est-à-dire de montrer que $u(t,x)$ satisfait (1') sans utiliser la fonction coût optimum. Dans le cas uniformément elliptique ceci se fait (voir [3] , [7]) grâce à une méthode de L.C. Evans [2] , mais cette technique ne semble pas permettre de traiter le cas parabolique en toute généralité. On peut néanmoins avec une extension de la technique de L.C. Evans arriver à passer à la limite (voir P.L. Lions [11]), mais ceci nécessite une régularité en t des coefficients et des données supérieures à celle que nous avons supposée ici.

Au contraire la méthode présentée ici s'adapte au cas elliptique, et permet également de traiter des problèmes dégénérés (voir P.L. Lions [9]).

Remarque II.6. La partie (a) de (ii) peut être légèrement simplifiée en adaptant au cas parabolique les résultats de [1] .

En effet si u_ϵ^i désigne la solution du système pénalisé (avec $\beta_\epsilon(t) = \frac{1}{\epsilon} t^+$) alors d'après l'interprétation stochastique ([1]) des u_ϵ , nous avons immédiatement $u_\epsilon^i \geqslant u$ (en effet u_ϵ est la fonction coût optimum sur une classe de systèmes admissibles incluse dans celle de la section I.2). Néanmoins, cette démonstration nécessitant l'introduction un peu lourde des évolutions aléatoires (voir [1]), nous avons préféré utiliser le système (9).

REFERENCES

[1] A. BENSOUSSAN and P.L. LIONS. *«Control of random evolutions»*. A paraître dans Stochastics.

[2] L.C. EVANS. *«A convergence theorem for solutions of nonlinear second order elliptic equations»*. Ind. Univ. Math. J. 27 (1978), p. 875-887.

[3] L.C. EVANS and A. FRIEDMAN. *«Optimal stochastic switching and the Dirichlet problem for the Bellman equations»*. Trans. Amer. Math. Soc. 253 (1979), p. 365-389.

[4] W.H. FLEMING and R. RISHEL. *«Deterministic and stochastic optimal control»*. Springer-Verlag, (1975) New-York.

[5] N.V. KRYLOV. *«On the uniqueness of a solution of Bellman's equation»*. Math. U.S.S.R. Izv. 5 (1971), p. 1387-1398.

[6] N.V. KRYLOV. *«Sequences of convex functions and estimates of the maximum of the solution of a parabolic equation»*. Sib. Math. Journal. Vol. 17, n° 2 (1976), p. 226-236.

[7] P.L. LIONS. *«Résolution de problèmes généraux de Bellman Dirichlet»*. Comptes Rendus Paris, Série A 287 (1978), p. 747-750, et Acta Mathematica (1981).

[8] P.L. LIONS. *«Contrôle de diffusions dans $\mathbb{R}^N$»*. Comptes Rendus Paris, Série A 287 (1979) p. 339-342, et Comm. Pure Appl. Math. 34 (1981), p. 121-147.

[9] P.L. LIONS. *«Equations de Hamilton-Jacobi-Bellman dégénérées»*. Comptes Rendus Paris, Série A 289 (1979), p. 329-332.

[10] P.L. LIONS. *«Some problems related to the Bellman-Dirichlet equation for two elliptic operator»*. Comm. P.D.E. 5(7) (1980), p. 753-771.

[11] P.L. LIONS. A paraître.

[12] P.L. LIONS et J.L. MENALDI. *«Problèmes de Bellman avec le contrôle dans les coefficients du plus haut degré»*. Comptes Rendus Paris, Série A 287 (1978), p. 409-412, et à paraître au SIAM J. Control Optim.

[13] M. NISIO. *«Some remarks on stochastic optimal controls»*. Japan J. Maths. Vol. 1, n° 1 (1975), p. 159-183.

(Manuscrit reçu le 29 septembre 1980)