EVOLUTION OF SOME PROBLEMS OF STOCHASTIC CONTROL
WHEN THE DISCOUNT VANISHES

J.M. LASRY

CNRS et Université de Paris IX Dauphine Mathématiques
de la Décision, 75775 Paris

H. HOWARD [7] , D. BLACKWELL [3] , and other authors have studied the evolution of the Markovian problems of control (P_τ) involving discounted cost functions, when the discount parameter $\tau < 1$ tends to 1. The main result is that there is a limit problem (P_1) such that the optimal policy P_1 of (P_1) is the limit of optimal policies p_τ of the problems (P_τ) ($\tau < 1$).

These results concern general Markovian processes with discrete time $t \in \mathbb{N}$, but might perhaps be extended to general Markovian processes with continuous time $t \in [0, +\infty[$.

In this article we study the case in which the Markovian processes are diffusions ξ , which are solutions of a differential stochastic equation : $d\xi = a(t)dt + dW$, where W is a brownian motion and where the control is the non-anticipative stochastic function a .

For this class of Markovian processes the results of convergence when the discount goes to one are mainly the same. The difference is that in the case of diffusion processes, the general frame of BELMAN-HALMITON-JACOBI theory [2] leads to an elliptic quasi-linear partial differential equation : H. FLEMING [4,5,...] has shown that this equation has a regular solution from which the optimal control can be computed. We shall recall (theorem 1) H. FLEMING's result which was not stated exactly in the same context; then give a similar analytical treatment for the limit (i.e : no-discounted) problem (theorem 2); and finally give the convergence results (theorem 3). The proofs of these three theorems are in [8,9] . Here we shall emphasize on the analytic treatment of the "ergodic" (limit) problem and give a proof of theorem 2 different from that of [8,9] .

In $\mathbb{R}^n$, $n \geq 1$, such results have been found by P. MANDL [11] using ordinary differential equations methods.

Discussions with A. BENSOUSSAN and L. TARTAR were helpful for this work.

§ 1. RESULTS

a) We note f a twice continuously differentiable function from $\mathbb{R}^n \times \mathbb{R}^n$ into $\mathbb{R}$ such that

$$(1) \quad f(x+m, y) = f(x, y) \quad \forall m \in \mathbb{N}; \forall x, y \in \mathbb{R}^n$$

$$(2) \quad \exists \alpha > 0 \text{ such that for all } x, y, z \in \mathbb{R}^n$$

$$\sum_{i,j=1}^n \frac{\partial^2}{\partial y_i \partial y_j} f(x, y) z_i z_j \geq \alpha z^2$$

$$(3) \quad \exists c > 0 \text{ such that for all } x, y \in \mathbb{R}^n :$$

$$\left\| \frac{\partial}{\partial x} f(x, y) \right\| + \left\| \frac{\partial}{\partial y} f(x, y) \right\| \leq c f(x, y)$$

Convexity hypothesis (1) is natural if one wants the existence of a regular control (see cost function below). Periodicity hypothesis (2) is a way to introduce the compactness (i.e. $\mathbb{R}^n / \mathbb{Z}^n$ is compact) we need for "ergodic" results. Growth hypothesis (3) is not very strong : For example $f(x, y) = \exp(x+y)$ satisfy (3). In application it can be useful to replace f by $f + \text{constant}$ (which does not change the control problem below) in order to verify (3).

We note $f^*(x, \cdot)$ the conjugate function of $f(x, \cdot)$, i.e. (ROCKAFELLAR [13]):

$$(4) \quad f^*(x, p) = \sup \{ py - f(x, y) \mid y \in \mathbb{R}^n \}$$

One can verify that $f^*: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is twice continuously differentiable, and that

$$(5) \quad \varphi(x, p) = \frac{\partial}{\partial p} f^*(x, p)$$

is the point where the function $y \rightarrow py - f(x, y)$ is minimum.

b) We denote W a standart brownian motion; we note E the expectation. Let A be the set of stochastic functions $a: \mathbb{R}_+ \rightarrow \mathbb{R}^n$ which are bounded non anticipative functionnals of W . For all $x \in \mathbb{R}^n$, $s > 0$ and $a \in A$, define $I(s, x, a)$ and $\lambda(x, a)$ by :

$$(6) \quad I(s, x, a) = E \int_0^{+\infty} f(\xi(t), a(t)) e^{-st} dt$$

$$(7) \quad \lambda(x, a) = \lim_{T \rightarrow +\infty} \inf E \frac{1}{T} \int_0^T f(\xi(t), a(t)) dt$$

where ξ is the stochastic process define by

$$(8) \quad \xi(0) = x, \quad d\xi = a(t)dt + \sqrt{2} dW$$

($\sqrt{2}$ in order to simplify formulas).

Let H be the set of all bounded lipschitzian functions $h: \mathbb{R}^n \rightarrow \mathbb{R}^n$. For all $h \in H$, $x \in \mathbb{R}^n$, $s > 0$, let

$$(9) \quad J(s, x, h) = E \int_0^{+\infty} e^{-st} f(\eta(t), a(t)) dt$$

$$(10) \quad \nu(x, h) = \lim_{T \rightarrow +\infty} E \frac{1}{T} \int_0^T f(\eta(t), a(t)) dt$$

where η is defined by the following stochastic differential equation :

$$(11) \quad \eta(o) = x, \quad d\eta = h(\eta) dt + \sqrt{2} dw$$

and where $a \in A$ is defined by

$$(12) \quad a(t) = h(\eta(t))$$

Take $x_o \in \mathbb{R}^n$ and $h_o \in H$.

Let η_o be defined by (11); then let $a_o \in A$ be defined by (12). Then we can compute ξ_o by (8) and we find $\xi_o = \eta_o$ by (11). So we have

$$(13) \quad I(x_o, s, a_o) = J(x_o, s, h_o) \quad \text{and} \quad \lambda(x_o, a_o) = \gamma(x_o, h_o)$$

Elements of H are called closed loop controls and elements of A are called open loop controls. By (13), it seems that closed loop controls are less efficient than the open ones. Precisely we have :

$$(14) \quad \inf_{h \in H} J(x, s, h) \geq \inf_{a \in A} I(x, s, a)$$

$$(15) \quad \inf_{h \in H} \gamma(x, h) = \inf_{a \in A} \lambda(x, a)$$

Actually theorem 1 and 2 below will imply that there is equality in (14) and (15), and that the optimal control is a closed loop control.

THEOREM 1 Let $s > 0$. The function u_s defined by

$$(16) \quad u_s(x) = \inf \{ I(x, s, a) \mid a \in A, x \in \mathbb{R}^n \} \text{ is three time continuously differentiable, periodic (i.e. : } u_s(x+m) = u_s(x), \forall x \in \mathbb{R}^n, \forall m \in \mathbb{Z}^n \}$$

This function is the unique periodic solution of the following PDE :

$$(17) \quad -\Delta u_s = s u_s(x) + f^*(x, -\nabla u_s(x)) = 0, \quad \forall x \in \mathbb{R}^n$$

(where : Δ = Laplacian, ∇ = gradient)

Let $h_s \in H$ be defined by (see(5)) :

$$(18) \quad h_s(x) = \varphi(x, -\nabla u_s(x)), \quad \forall x \in \mathbb{R}^n$$

Then we have

$$(19) \quad u_s(x) = J(x, s, h_s) = \inf \{ J(x, s, h) \mid h \in H \}$$

which means in particular that h_s is the optimal control in the problem

$$(20) \quad \text{Minimize } J(x, s, h) \\ h \in H$$

THEOREM 2. The function λ defined by

$$(21) \quad \lambda(x) = \inf \{ \lambda(x, a) \mid a \in A \}, \quad \forall x \in \mathbb{R}^n$$

$$(22) \quad \text{is constant : } \lambda(x) = \lambda_o, \quad \forall x \in \mathbb{R}^n.$$

There exists a function $v_o : \mathbb{R}^n \rightarrow \mathbb{R}$ periodic, i.e.:

$$v_o(x+m) = v_o(x), \quad \forall m \in \mathbb{Z}^n, \quad \forall x \in \mathbb{R}^n,$$

three time continuously differentiable, such that

$$(23) \quad -\Delta v_o(x) + \lambda_o + f^*(x, -\nabla v_o(x)) = 0, \quad \forall x \in \mathbb{R}^n$$

Let $h_o \in H$ be defined by (see 15)):

$$(24) \quad h_0(x) = \varphi(x, -\nabla v_0(x)), \quad \forall x \in \mathbb{R}^n$$

we have, for all $x \in \mathbb{R}^n$:

$$(25) \quad \lambda_0 = \rho(x, h_0) = \inf \{ \rho(x, h) \mid h \in L \}$$

Which means in particular that h_0 is the solution of the following problem of stochastic control

$$(26) \quad \text{Minimize}_{h \in H} \rho(x, h)$$

Moreover, if (λ', h') is another solution of (23) then $\lambda' = \lambda_0$ and $h' = h_0 + \text{constant}$.

THEOREM 3 . Let u_s, h_s, λ_0, v_0 be defined as in theorem 1 and 2. Then when $s > 0$ converges to 0, we have

$$(27) \quad \|u_s - \lambda_0\| \rightarrow 0$$

$$(28) \quad \|\nabla u_s - \nabla v_0\| \rightarrow 0$$

$$(29) \quad \|h_s - h_0\| \rightarrow 0$$

$$(30) \quad \text{with } \|g\| = \sup \{ \|g(x)\| + \|\nabla g(x)\| \mid x \in \mathbb{R}^n \}$$

§ 2. PROOF OF THEOREM 2

Proofs of theorem 1 and 3, and of unicity in theorem 2 can be found in [8, 9].

Let $C^{k, \alpha}(\mathbb{R}^n)$ be the space of functions $u : \mathbb{R}^n \rightarrow \mathbb{R}$ which are k -time differentiable with α -Holder-continuous derivatives, (see [10]). Let $C_*^{k, \alpha}$ be the subspace of $C^{k, \alpha}(\mathbb{R}^n)$ of functions u such that

$$(1) \quad u(x+m) = u(x) \quad \forall x \in \mathbb{R}^n, \forall m \in \mathbb{Z}^n$$

$$(2) \quad u(0) = 0$$

From the classical results about Laplacian (Schauder, see [10], Agmon-Douglis-Nirenberg [1]), one can deduce the following lemma

Lemma 1 If $g \in C^0$, there exists a unique $u \in C_*^1$ and a unique $\lambda_0 \in \mathbb{R}$ such that (in a weak sens):

$$(3) \quad -\Delta u + \lambda_0 = g \quad \text{in } \mathbb{R}^n$$

Moreover : $u \in C_*^{1, \alpha}$ for all $\alpha \in]0, 1[$, and there exists a constant C_α , not depending on g , such that

$$(4) \quad \|u\|_{1, \alpha} \leq C_\alpha \|g\|_0$$

(where $\|\cdot\|_{k, \alpha}$ the norm in $C^{k, \alpha}$)

If $g \in C_*^{k, \alpha}$ then $u \in C_*^{k+2, \alpha}$ and

$$(5) \quad \|u\|_{k+2,\alpha} \leq b_\alpha \|g\|_{k,\alpha}$$

where b_α is a constant which does not depend on g

Définition 2 : According to lemma 1 one can define an operator K from $C^1_\cdot$ into itself by $v = K(u)$ iff there exists $\lambda_0 \in \mathbb{R}$ such that

$$(6) \quad -\Delta v + \lambda_0 = -f^*(\cdot, -\nabla u(\cdot))$$

Remark that (6) and periodicity imply that

$$(7) \quad \lambda_0 = - \int_{[0,1]^n} f^*(x, -\nabla u(x)) dx$$

Lemma 3 : K is continuous and compact.

Proof : For $r \geq 0$, let

$$(8) \quad \theta(r) = \sup \{ \|f^*(x, y)\| \mid x, y \in \mathbb{R}^n, \|y\| \leq r \}$$

Then by (4) and (6) we have

$$(9) \quad \|u\|_1 \leq r \Rightarrow \|K_u\|_{1,\alpha} \leq c_\alpha \theta(r)$$

As the injection $C^{1,\alpha}_\cdot \rightarrow C^1$ is compact this proves that K is compact.

Now let $\|u_n - u\|_1 \rightarrow 0$ and let $v_n = K_{u_n}$. By compactity there exists $v \in C^1$ and a subsequence v_{n_k} such that $\|v_{n_k} - v\|_1 \rightarrow 0$.

From $-\Delta v_n + \lambda_n = -f^*(\cdot, -\nabla u_n(\cdot))$, from the continuity of f^* and from (7), we deduce that $-\Delta v + \lambda = -f^*(\cdot, u(\cdot))$, with $\lambda = \lim \lambda_{n_k}$.

So $v = Ku$. This shows the continuity of K .

Lemma 4 ^(*). There exists a constant $c_3 \in \mathbb{R}$ such that if $u \in C^1_\cdot$, if $\alpha \in [0, 1]$, and if $u = \alpha Ku$, then $\|u\|_1 \leq c_3$.

proof Let $u \in C^1_\cdot$, $u = \alpha Ku$, $0 < \alpha \leq 1$ ($\alpha = 0$ is trivial). Then by definition 2, we have

$$(10) \quad -\Delta u + \alpha \lambda = -\alpha f^*(x, -\nabla u(x)), \quad \forall x \in \mathbb{R}^n$$

Let us first show that $u \in C^3_\cdot$, by a standard a posteriori regularity argument. As $u \in C^1_\cdot$, we see that $f^*(\cdot, -\nabla u(\cdot))$ belongs to $C^0_\cdot$. From lemma 1 and (10) we deduce that $u \in C^{1,\alpha}_\cdot$. Now we see that $f^*(\cdot, -\nabla u(\cdot))$ belongs to $C^{0,\alpha}_\cdot$. Again from (10) and lemma 1, this implies $u \in C^{2,\alpha}_\cdot$. Now we see that $f^*(\cdot, -\nabla u(\cdot)) \in C^{1,\alpha}_\cdot$, and again by (10) and lemma 1 we have $u \in C^{3,\alpha}_\cdot$. That should be enough.

As u is continuous and periodic there exists $x_0 \in \mathbb{R}^n$ such that $u(x_0) = \max u$. As u is twice continuously differentiable, this implies $\nabla u(x_0) = 0$ and $\Delta u(x_0) \leq 0$. So from (10) we get :

$$\alpha \lambda \leq f^*(x_0, 0)$$

By the same argument on the minimum of u we find

$$(11) \quad |\lambda| \leq c_1 = \sup \{ \|f^*(x, 0)\| \mid x \in \mathbb{R}^n \}$$

^(*)Remark The proof of lemma 4 given here was kindly communicated by L. TARTAR. The initial probabilistic proof can be found in [8,9].

Now let y belongs to R^n , and $\|y\| = 1$. Put

$$(12) \quad v(x) = y \cdot \nabla u(x), \quad \forall x \in R^n$$

By derivation we deduce from (10) that

$$(13) \quad -\Delta v(x) = -g_1(x, -\nabla u(x)) + g_2(x, -\nabla u(x)) \cdot \nabla u(x)$$

where, for all x, p in R^n :

$$(14) \quad g_1(x, p) = y \cdot \frac{\partial}{\partial x} f^*(x, p) \text{ and } g_2(x, p) = \frac{\partial}{\partial p} f^*(x, p)$$

let $w = c(u - \min u + c_1) - v$, where c_1 come from (11) and c come from (3) of § 1. Then from (10) and (13) one deduces

$$(15) \quad -\Delta w + \alpha w - \alpha a \cdot \nabla w = b \quad \text{with}$$

$$(16) \quad \begin{cases} a(x) = g_2(x, -\nabla u(x)) \text{ and} \\ b(x) = c\alpha [-\lambda + c_1 - f^*(x, -\nabla u(x)) + u(x) - \min u \\ \quad - a(x) \cdot \nabla u(x)] - \alpha [v(x) - g_1(x, -\nabla u(x))] \end{cases}$$

From the definition § 1 (4) of f^* one can deduce the usual relations of convex analysis (see Rockafellar [13]).

$$(17) \quad \begin{aligned} f(x, y) &= \sup \{ y \cdot p - f^*(x, p) \mid p \in R^n \} \\ &= y \cdot \bar{p} - f^*(x, \bar{p}) \end{aligned}$$

where $\bar{p}$ is defined by one of the equivalent following relations.

$$(18) \quad y = \frac{\partial}{\partial p} f^*(x, p) \Leftrightarrow p = \frac{\partial}{\partial y} f(x, y)$$

One can also show that

$$(19) \quad \frac{\partial}{\partial x} f(x, y) = - \frac{\partial}{\partial x} f^*(x, p), \text{ when } y, p \text{ satisfy (18)}$$

Then from (17, 18, 19) and from § 1.3 one deduces

$$(20) \quad c \left[p \cdot \frac{\partial}{\partial p} f^*(x, p) - f^*(x, p) \right] \geq \left\| \frac{\partial}{\partial x} f^*(x, p) \right\| + \|p\|$$

From (20) with $p = -\nabla u(x)$, from the definition of a (see (16)), from $\|v(x)\| = \|y \cdot \nabla u(x)\| \leq u(x)$ (because $\|y\| = 1$), and from the definition (14) of g , one deduces:

$$(21) \quad c [-a(x) \cdot \nabla u(x) - f^*(x, -\nabla u(x))] - v(x) + g_1(x, -\nabla u(x)) \geq 0$$

From $u(x) - \min u \geq 0$, from $c_1 - \lambda \geq 0$ (see (11)), from (16, 21) we deduce

$$(22) \quad b(x) \geq 0, \quad \forall x \in R^n$$

As w is continuous and periodic there exists $x_0 \in R^n$ such that $w(x_0) = \min w$. As w is twice continuously differentiable, one has $\nabla w(x_0) = 0$ and $\Delta w(x_0) \geq 0$. So from (15) and (22) one gets $\alpha w(x_0) \geq b(x_0) \geq 0$. So $w \geq \min w = w(x_0) \geq 0$. From the definitions of v and w one deduces

$$(23) \quad y \cdot \nabla u(x) \leq c [u(x) - \min u + c_1], \quad \forall x \in R^n$$

But (23) is true for all y such that $\|y\| = 1$. So we have

$$(24) \quad \|\nabla u(x)\| \leq c [u(x) - \min u + c_1], \quad \forall x \in R^n$$

Let $x_0 \in R^n$ be such that $u(x_0) = \min u$. From (24) we get by Gronwall lemma:

$$(25) \quad u(x) - \min u + c_1 \leq c_1 \exp(c \|x - x_0\|)$$

Periodicity allows us to suppose $\|x - x_0\| \leq \sqrt{n}$. So from (24) and (25) we deduce

$$(26) \quad \|\nabla u(x)\| \leq c_2 = c c_1 \exp(c \sqrt{n})$$

As $u \in C^1_*$ imply $u(0) = 0$, one deduces $\|u\|_1 \leq c_3$ from (26) with c_3 depending on c_2 .

Lemma 5. There exists a function $v_0 \in C^3_*$ and a real λ_0 such that

$$(27) \quad -\Delta v_0(x) + \lambda_0 + f^*(x, -\nabla v_0(x)) = 0, \quad \forall x \in \mathbb{R}^n$$

Proof Thanks to lemma 3 and 4, and to the Schaefer fixed point theorem (Schaefer [14]) there exists $v_0 \in C^1_*$ such that $v_0 = K v_0$. So by an argument given at the beginning of the proof of lemma 4, one has $v_0 \in C^3_*$ and (27).

Lemma 6. Let $0 \in R$ and $v_0 \in C^3_*$ satisfy (27). Then

$$(28) \quad v_0(x) \leq \lambda(x, a) \quad \text{for all } a \in A$$

$$(29) \quad v_0(x) \leq \rho(x, h) \quad \text{for all } h \in H$$

$$(30) \quad v_0(x) = \rho(x, h_0) \quad \text{where } h_0 \text{ is defined by}$$

$$h_0(x) = g_2(x, -\nabla v_0(x)) \quad \forall x \in \mathbb{R}^n \quad (\text{see (14)})$$

Proof Let us compute $E(v(\xi(t)))$ where ξ is defined by : $\xi(0) = x$, $d\xi = a(t) + \sqrt{2}dW$

By Ito's lemma (Gihman - Skorohod [6]) we have

$$(31) \quad E(v(\xi(t))) - E(v(\xi(0))) =$$

$$E \int_0^T \Delta v(\xi(t)) + a(t) \cdot \nabla v(\xi(t)) dt + \sqrt{2} \int_0^T \nabla v(\xi(t)) dW$$

As $t \rightarrow \infty$ $\nabla v(\xi(t))$ is non-anticipative we have $E \int_0^T \nabla v(\xi(t)) dW = 0$. So we deduce from (17, 27, 31) that

$$(31) \quad E(v(\xi(t)) - v(x)) \geq E \int_0^T [\lambda_0 - f(\xi(t), a(t))] dt$$

Divide each side of (31) by T , and let $T \rightarrow +\infty$. We find (28) and from the definition (§1,7) of λ . From (28) and from (§1,13) we deduce (29). If $h = h_0$ we deduce from (17) that there is equality in (31), which leads to (30).

REFERENCES

- (1) S. AGMON, A. DOUGLIS, L. NIREMBERG
Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, I, Comm. pure and appl. Math., t.12, 1959, p. 623 - 727.
- (2) R. BELMAN,
Dynamic programming, Princeton, 1957
- (3) D. BLACKWELL,
Discret dynamic programming,
Ann. Math. Stat., t.33 - p. 719 à 726, 1962
- (4) W.H. FLEMING,
The Cauchy problem for a non linear first order partial differential equation,
J. of diff. eq., t.5 - 1969.
- (5) W.H. FLEMING,
Some markovian optimization problems, J. of Math. Mech., 12, n°1 - 1963,
p. 131-140.
- (6) I. GIHMAN A.V. SKOROHOD,
Stochastic differential equations, Springer Verlag , 1972.
- (7) R. HOWARD,
Dynamic programming and Markov processes, Technology Presse and Wiley, N.Y..
- (8) J.M. LASRY,
Comptes rendus du congrès de contrôle optimal de Juin 1974 à l'I.R.I.A.,
lecture notes - 1975, econ. math. systems., n° 107.
- (9) J.M. LASRY,
Thèse, Université de Paris IX Dauphine, secrétariat de Mathématiques de la
Décision, 75775 PARIS CEDEX 16.
- (10) C.A. LADYZENSKAJA et N.N. URAL'CEVA,
Equations aux dérivées partielles de type elliptique, Nanka,
Moscou 1964 ; traduction française ; Dunod - 1968.
- (11) P. MANDL,
Analytic Treatment of one dimensionnal Markov Processes,
ch. VI, Springer Verlag, N.Y. - 1968.

- (12) H. MOULIN,
Prolongements des jeux à deux personnes de somme nulle,
Thèse, S.M.F., 1976.

- (13) T. ROCKAFELLAR,
Convex Analysis
Princeton University Press - 1970.

- (14) H. SCHAEFFER,
Über die Methode der a priori Schranken, Math. ann.,
t.279, 1955, p. 415-416.