MASATOSHI FUJISAKI

On the uniqueness of optimal controls

Séminaire de probabilités (Strasbourg), tome 13 (1979), p. 548-556
<http://www.numdam.org/item?id=SPS_1979__13__548_0>

Université de Strasbourg
Séminaire de Probabilités 1977/78

ON THE UNIQUENESS OF OPTIMAL CONTROLS

by Masatoshi FUJISAKI

INTRODUCTION.

There are many works concerning the existence of optimal controls under various conditions. The purpose of this paper is to give some criteria for the uniqueness of optimal controls.

In § 1 we discuss the uniqueness of optimal controls in the completely observable case, under the hypotheses of Ikeda-Watanabe [2]. In this case we can easily give simple criteria for the uniqueness of the optimal control whose existence is proved in [2].

In § 2 we consider the same problem, but in the partially observable case (Fujisaki [1]). The control system is more complicated, and uniqueness becomes more difficult to prove.

§ 1. COMPLETELY OBSERVABLE CASE.

Let T be a positive number. Loosely speaking, the control problem is the following. One considers a process $(X_t)_{0 \leqq t \leqq T}$ with values in $\mathbb{R}^n$ and continuous paths, solution (in the sense of law) of the following differential equation

(1.1) $$dX_t = u_t dt + d\beta_t \qquad X_0 = x$$

where x is a vector in $\mathbb{R}^n$, and $u_t = \psi(t, X_s, 0 \leqq s \leqq t)$, the control, is a given functional of the process X (u_t depends on the information obtained from the data X_s, $s \leqq t$). The equation (1.1) in the sense of law means that the law of the process

$$\beta_t = X_t - x - \int_0^t u_s ds$$

is the same as that of standard n-dimensional brownian motion. Then our optimization problem consists in finding a control u in a suitable class, such that the process X minimizes a cost function

(1.2) $$J = E[\int_0^T f(t, |X_t|) dt]$$

where f(t,x) is a $\mathbb{R}^1_+$-valued function over the product space $[0,T] \times \mathbb{R}^1_+$, which for each t is increasing w.r. to x.

Now we define the class of controls precisely. Since J depends only on the law of the process X, we may assume first that X_t is the coordinate mapping w_t at time t on the Banach space $W = \underline{\underline{C}}^n$ of all $\mathbb{R}^n$-valued continuous functions over $[0,T]$, with uniform norm. $\underline{\underline{B}}^n$ is the topological Borel field of $W=\underline{\underline{C}}^n$ (also generated by all cylinder sets), and $\underline{\underline{B}}_t^n$ is the sub-σ-field generated by the cylinder sets up to time t. Let Ψ be the class of all $\mathbb{R}^n$-valued functions $\psi(t,w)$ over $[0,T]\times W$ satisfying the following three conditions

(Ψ.1) $\psi(t,w)$ is measurable in (t,w)

(Ψ.2) for each t, $w\mapsto\psi(t,w)$ is measurable w.r. to $\underline{\underline{B}}_t^n$

(Ψ.3) $|\psi(t,w)|\leq 1$ for all (t,w)

Ψ will be called the class of all admissible functionals.

By Girsanov's theorem we have the following proposition on the existence of solutions of equation (1.1) :

PROPOSITION 1.1. For any $\psi\in\Psi$, there is a unique law P^ψ on $\underline{\underline{C}}^n=W$ such that under P^ψ

$$w_0=x \text{ a.s.} \qquad w_t-w_0-\int_0^t\psi(s,w)ds \text{ is a standard n-dimensional brownian motion}$$

There are several ways of constructing the law P^ψ. The first one is the following (a second one will be given later). Start with a Wiener probability law P on W, that is, the unique law under which $(w_t)_{0\leq t\leq T}$ is a standard n-dimensional Brownian motion, and assume the stochastic differential equation on W

$$dX_t = \psi(t,X)dt + dw_t \qquad X_0=x$$

has a unique solution in the pathwise sense. Then the law of the process X is equal to P^ψ .

Since there is a unique law P^ψ associated to ψ, we may also define the cost associated to ψ

$$J(\psi) = E^\psi\left[\int_0^T f(t,|w_t|)dt\right]$$

We say that an admissible functional ψ^o is optimal if

$$J(\psi^o) = \inf_{\psi\in\Psi} J(\psi) \tag{1.3}$$

The following theorem relative to the existence of optimal functionals is due to Ikeda-Watanabe [2]

THEOREM 1.2. There exists an admissible optimal functional ψ^o and moreover it can be written in the following explicit form

$$\psi^o(t,w)= U(w_t) \tag{1.4}$$

$$U(x)= -x/|x| \text{ for } x\neq 0 \;,\quad U(x)=0 \text{ for } x=0 \tag{1.5}$$

Besides that, the law P^{ψ^o} can be constructed by the above method : the stochastic differential equation :

(1.6) $$dX_t = U(X_t)dt + dw_t \qquad X_0=x$$

(w.r. to Wiener measure) has a unique solution in the pathwise sense.

Let us now return to the general case : any process $(X_t)_{0\leq t\leq T}$ with values in $\mathbb{R}^n$ and continuous paths, over some probability space $(\Omega,\underline{\underline{F}},P)$, defines a mapping from Ω to $W=\underline{\underline{C}}^n$, still denoted by X : $X(\omega)$ is the path $t\mapsto X_t(\omega)$. This mapping is measurable. A <u>control</u> for X is a process $(u_t)_{0\leq t\leq T}$ which can be written as $u_t(\omega)=\psi(t,X(\omega))$ for some $\psi\in\Psi$. We say that a pair (X,ψ) is an <u>admissible system</u> , or that the control (u_t) is an <u>admissible control</u> if X and u satisfy the stochastic differential equation (1.1), or equivalently, if the image law of P under the mapping X is the law P^ψ.

In the canonical situation ($\Omega=W$, $X_t=w_t$), X is the identity mapping, and controls are the same as functionals.

Before we study uniqueness, we need some preliminaries and notations. First of all, let us describe the second way of constructing the law P^ψ, on $W=\underline{\underline{C}}^n$, for given $\psi\in\Psi$. We start again with a Wiener probability measure P on W, but this time such that $w_0=x$ P-a.s.. Then we set

(1.7) $$\rho_t(\psi) = \exp\left(\int_0^t \psi(s,w)dw_s - \frac{1}{2}\int_0^t |\psi(s,w)|^2 ds \right)$$

Since $\psi(t,w)$ is bounded, it is well known that $(\rho_t(\psi), \underline{\underline{B}}^n_t , P)$ is a uniformly integrable martingale for $0\leq t\leq T$. If we define a measure P^ψ by

(1.8) $$dP^\psi = \rho_T(\psi)dP$$

then P^ψ is the same measure as in proposition 1.1. Note that, since ρ is a martingale, the density of P^ψ w.r. to P over $\underline{\underline{B}}^n_t$ is $\rho_t(\psi)$.

To see this, we remark that from Girsanov's theorem, the stochastic process given by

(1.9) $$\tilde{w}_t = w_t - w_0 - \int_0^t \psi(t,w)dt$$

is a standard brownian motion under the law (1.8). Therefore equation (1.1) is satisfied and we know it has a unique solution in the sense of law.

Therefore we also have the following result which shows that $J(\psi)$ depends on $\rho_T(\psi)$ only

(1.10) $$J(\psi) = E^\psi\left[\int_0^T f(t,|w_t|)dt\right] = E\left[\rho_T(\psi)\int_0^T f(t,|w_t|)dt\right]$$

Let $F(t,w)$ be a non-negative Borel function defined on $[0,T]\times\underline{\underline{C}}^n$ which is increasing in the following sense : for any t, $0\leq t\leq T$, (1.11) If w^1 and w^2 belong to $\underline{\underline{C}}^n$ and $|w^1(s)|\leq|w^2(s)|$ for all $s\in[0,t]$, then $F(t,w^1) \leq F(t,w^2)$.

Then, by Ikeda-Watanabe [2], as a corollary to Theorem 1.2, we have the following result. If ψ^o is the optimal functional (1.4), and if (X,ψ) is any admissible system on any probability space Ω, then

(1.12) $$E^{\psi^o}[F(t,w)] \leq E[F(t,X(\omega))], \quad 0\leq t\leq T$$

This can be translated into a result which involves only Wiener measure, namely that, for any $\psi\in\Psi$ and any $t\leq T$

(1.13) $$E[\rho_t(\psi^o)F(t,w)] \leq E[\rho_t(\psi)F(t,w)] \,.$$

For each t, $0\leq t\leq T$, denote by $\tilde{\underline{\underline{B}}}{}^n_t$ the sub-σ-field of $\underline{\underline{B}}^n_t$, which is generated by the random variables $|w_s|$, $s\leq t$. Then we can give very easily the following characterization of the functional ψ^o :

THEOREM 1.3. The functional ψ^o satisfies the following properties
(U1) Inequality (1.13) holds for any admissible functional ψ
(U2) For each t, $\rho_t(\psi^o)$ is $\tilde{\underline{\underline{B}}}{}^n_t$-measurable
If ψ' is any admissible functional satisfying these conditions, then $\psi'(t,w)=\psi^o(t,w)$ a.s. with respect to the product measure dtdP.

PROOF. We already know that (U1) is satisfied. By (1.7)

$$\rho_t(\psi^o) = \exp\left(- \int_0^t \frac{(w_s,dw_s)}{|w_s|} - \frac{1}{2}t \right)$$

where $(w_s,dw_s) = \sum_{i=1}^n w_s^i dw_s^i$. On the other hand, $|w_s|^2 = \sum_{i=1}^n |w_s^i|^2$, so that $d(|w_s|^2) = 2(w_s,dw_s)+ n\,dt$, and $\int_0^t \frac{(w_s,dw_s)}{|w_s|} = \frac{1}{2}\int_0^t \frac{1}{|w_s|}(d|w_s|^2 - n\,dt)$ is $\tilde{\underline{\underline{B}}}{}^n_t$-measurable . It is shown in [4] that this process is a 1-dimensional brownian motion.

To prove uniqueness, we remark that if ψ' is another functional with the same properties, we must have for any t

$$E[\rho_t(\psi^o)F(t,w)] = E[\rho_t(\psi')F(t,w)]$$

Taking $F(t,w)$ to be $k_1(|X_{t_1}(w)|)\ldots k_m(|X_{t_m}(w)|)$, where $k_1,\ldots,k_m$ are increasing functions[1] on $[0,\infty[$ and $t_1\leq t,\ldots,t_m\leq t$, we deduce very easily that $\rho_t(\psi^o) = \rho_t(\psi')$ a.s. for each t. Since these processes are continuous, they a.s. have the same paths.

1. Looking a little more closely at the proof, we may restrict the class of functions $F(s,w)$ to those which are adapted to the family $(\underline{\underline{B}}^n_s)$.

On the other hand, equality of these processes implies that of the processes

$$\int_0^t \psi^o(s,w)dw_s = \int_0^t \frac{d\rho_s(\psi^o)}{\rho_s(\psi^o)} \quad , \quad \int_0^t \psi'(s,w)dw_s = \int_0^t \frac{d\rho_s(\psi')}{\rho_s(\psi')}$$

The square integrable martingale $\int_0^t (\psi^o(s,w)-\psi'(s,w))dw_s$ then is equal to 0, and so is its increasing process $\int_0^t (\psi^o(s,w)-\psi'(s,w))^2 ds$. The theorem follows at once.

§ 2. PARTIALLY OBSERVABLE PROBLEM

In this section we intend to apply the preceding results to the partially observable case. Then the situation is more complicated so that there are few works relative to the existence theorem of optimal controls, except linear ones. Here we shall adopt a linear control system as formulated by Fujisaki [1]. First we describe this problem and the previously known results, and next we give some uniqueness results.

Here equation (1.1) is replaced by a system

$$(2.1) \qquad \begin{aligned} d\theta_t &= u_t dt + d\beta_t & \theta_0 &= \text{a given random variable} \\ d\zeta_t &= a_t\theta_t dt + d\gamma_t & \zeta_0 &= 0 \end{aligned}$$

to be solved in the law sense. Here (β_t) and (γ_t), $0 \leqq t \leqq T$, are independent m-dimensional and n-dimensional brownian motions respectively, with $\beta_0=0$, $\gamma_0=0$[1], (θ_t) is an $\mathbb{R}^m$-valued process called the <u>state of channel</u> , (ζ_t) is an $\mathbb{R}^n$-valued process called the <u>output</u> , a_t is an (n,m) matrix (non random, measurable and bounded as a function of t) such that $a_t^* a_t = cI_m$, where c is a positive constant and I_m is the m-identity matrix (we denote by $*$ the transpose of any matrix or vector ; vectors are always meant to be column ones). Finally, $u=(u_t)$ is the <u>control</u> , which may depend upon the a priori distribution of θ_0 and the information obtained on $\{\zeta_s, s \leqq t\}$, but not on $\{\theta_s , s \leqq t\}$ - this is why the problem is called partially observable. The cost function will depend only on the process (θ_t) and will have the following form

$$(2.2) \qquad J(u) = E[\int_0^T f(\theta_t)dt]$$

where $f(x)$ is a non-negative function on $\mathbb{R}^m$, for instance

(1) $f(x) = |x|^2$ (2) $f(x)=0$ $(|x| \leqq H)$, $=1$ $(|x|>H)$

1. In the paper [1] $d\beta_t$ and $d\gamma_t$ are replaced by $B_t d\beta_t$ and $b_t d\gamma_t$ with (non random) orthogonal matrices B_t and b_t. If the system is to be solved in the sense of law, this makes no difference.

where H is a positive constant. Especially, in the 1-dimensional case, $f(x)$ is always taken as an increasing function of $|x|$.

Next we define the class of admissible controls in the same manner as in the completely observable case. Define $\underline{\underline{C}}^n$, $\underline{\underline{C}}^m$, $\underline{\underline{C}}^{n+m}$ as in the preceding section (i.e. Banach spaces of continuous functions over $[0,T]$, with the uniform norm), and the obvious notation for the corresponding σ-fields. Denote by Φ the class of $\mathbb{R}^m$-valued functions $\varphi(t,w)$ over $[0,T]\times\underline{\underline{C}}^n$ satisfying the same three conditions as Ψ in section 1, except for "$\mathbb{R}^n$-valued" . For simplicity we shall work only on the canonical space $W=\underline{\underline{C}}^{m+n}=\underline{\underline{C}}^m\times\underline{\underline{C}}^n$, denoting by Θ the projection map of $\underline{\underline{C}}^{m+n}$ onto $\underline{\underline{C}}^m$, by ζ the corresponding map onto $\underline{\underline{C}}^n$, and by $\Theta_t(w)$, $\zeta_t(w)$ the corresponding values at time t. Then the control $u_t(w)$ is equal to $\varphi(t,\zeta(w))$, and we can again identify controls and functionals.

We again have a proposition similar to proposition 1.1 :

PROPOSITION 2.1. Given any $\varphi\in\Phi$ and any probability law μ on $\mathbb{R}^m$, there exists a unique law on $W=\underline{\underline{C}}^{m+n}$ such that

1) $\Theta_t(w)-\Theta_0(w) - \int_0^t \varphi(s,\zeta(w))ds$ and $\zeta_t(w)-\zeta_0(w)-\int_0^t a_s\Theta_s(w)ds$ are independent m-dimensional and n-dimensional brownian motions

2) $\zeta_0= 0$ a.s. and the law of Θ_0 is μ .

That is, equation (2.1) is satisfied in the sense of law. We shall denote this law by $P^{\varphi,\mu}$ or simply P^{φ} if no confusion can arise.

The cost corresponding to this law can be denoted by $J(\varphi,\mu)$ or $J(\varphi)$ if no confusion can arise. A functional φ^o, or the corresponding control $u_t^o(w)=\varphi^o(t,\zeta(w))$, is said to be <u>optimal</u> (for a given μ) if

$$J(\varphi^o,\mu) = \inf_{\varphi\in\Phi} J(\varphi,\mu) \tag{2.3}$$

For convenient choices of the cost function (for instance $f(x)$ given by (2) after (2.2)) and of the measure μ, one can again give an explicit description of an optimal control.

THEOREM 2.2 (Fujisaki [1]). If the initial distribution μ of Θ_0 is normal $N(m,\sigma^2)$, where m is an m-vector and σ^2 is an (m,m)-matrix of the type cI_m , $c>0$, then there exists some optimal $\varphi\in\Phi$ such that the control $u_t=\varphi(t,\zeta)$ can be represented as

$$u_t = U(m_t) \tag{2.4}$$

where $U(x)$, $x\in\mathbb{R}^m$ is given by (1.5), and $m_t=E^{\varphi}[\Theta_t\,|\,\zeta_s,s\leq t]$.

We are going now to apply section 1 to the partially observable problem. Since there are essential computational difficulties in the multidimensional case, we are going to assume that m=n=1. Let $\underline{\underline{G}}$ be the class of all real valued functions $g(x,\alpha)$ over $\mathbb{R}^1\times\mathbb{R}^1$ which satisfy the following condition : if $\eta(x,\alpha,\sigma^2)$ is the normal density with mean α and variance σ^2, then

(2.5) $$\tilde{g}(\sigma^2,\alpha) = \int g(x,\alpha)\eta(x,\alpha,\sigma^2)dx$$

depends only on $\sigma^2, |\alpha|$, is non-negative and increases with $|\alpha|$ (this implies that g itself is non-negative ; functions $g(x,\alpha)=f(|x|)$, where f is a non-negative increasing function on $[0,\infty[$, obviously belong to the class $\underline{\underline{G}}$).

Consider now the minimizing problem with the same control system as in theorem 2.2. , but involving a cost function of the type

(2.6) $$J(u) = E[\int_0^T g(\theta_t,m_t)dt] \qquad g\in\underline{\underline{G}}$$

We shall use the following results from [1]. For simplicity we assume that a=1. It can be shown that the conditional distribution of θ_t w.r. to $\underline{\underline{F}}_t = \sigma\{\zeta_s, s\leq t\}$ is Gaussian $N(m_t, \sigma_t^2)$, where m_t is the conditional mean as in theorem (2.2), and $\sigma_t^2=E[(\theta_t-m_t)^2|\underline{\underline{F}}_t]$, that is, the conditional variance. Furthermore, they satisfy the following equations

(2.7) $$\frac{d\sigma_t^2}{dt} = 1-(\sigma_t^2)^2 \quad , \sigma_0^2=\sigma^2$$

(2.8) $$dm_t = u_t dt + \sigma_t^2 d\nu_t \ , \ m_0=m$$

where (ν_t) is a $(\underline{\underline{F}}_t)$-adapted 1-dimensional brownian motion. Note that σ_t^2 is a non random, well determined function of t only.

On the other hand, the cost function J(u) can be written as

(2.9) $$J(u)= E[\int_0^T \tilde{g}(t,m_t)dt]$$

In (2.8), the control u_t is a functional $\varphi(t,\zeta)$ adapted to $(\underline{\underline{F}}_t)$, while the process (m_t) is adapted to $(\underline{\underline{F}}_t)$, but may generate smaller σ-fields. If it turns out that the control u_t can be written as a functional $\overline{\varphi}(t,m)$, then looking at (2.8) and (2.9) only we have a completely observable control problem, entirely similar to that of section 1 since $\tilde{g}(t,x)$ is, for fixed t, an increasing function of $|x|$ only. The only difference lies in the fact that we have, in (2.8), $\sigma_t^2 d\nu_t$ instead of simply $d\nu_t$.

From now on, when we lock for uniqueness criteria, we shall restrict ourselves to the following subclass $\hat{\Phi}$ of Φ , which consists of those φ for which the control u_t depends only on the conditional mean $E^{\varphi}[\theta_t|\underline{\underline{F}}_t]=m_t$. We know from theorem 2.2 that this class contains optimal controls. It is easy to see that the control can be written as $\overline{\varphi}(t,m)$ with some (possibly different) $\overline{\varphi}$.

Let $\varphi(t,\zeta)=\overline{\varphi}(t,m)$ be a control of the class $\hat{\Phi}$. It is easy to construct on $\underline{\underline{C}}^1$ the law $P^{\overline{\varphi}}$ of the conditional mean process m . Let P be the fixed law on $\underline{\underline{C}}^1$ under which $\int_0^t dw_s/\sigma_s^2$ is a standard brownian motion, and such that $w_0=m$ a.s.. Then according to Girsanov's theorem, $P^{\overline{\varphi}}$ is absolutely continuous w.r. to P, with density $\rho_T(\overline{\varphi})$, given by

(2.10) $$\rho_t(\overline{\varphi}) = \exp\Big(\int_0^t (\sigma_s^2)^{-2}\overline{\varphi}(s,w)dw_s - \tfrac{1}{2}\int_0^t (\sigma_s^2)^{-2}\overline{\varphi}^2(s,w)ds \Big).$$

Since we are reduced to a completely observable problem, we may apply now the method that leads to theorem 1.3, with very small changes, to get the uniqueness results given below. However, the statements concern the functional $\overline{\varphi}$ associated with φ rather than φ itself, and so the conditions are difficult to verify.

Let $\tilde{\underline{\underline{B}}}{}^1_t$ be the σ-field generated by the random variables $|w_s|$, $s\leqq t$. Then we have the following

PROPOSITION 2.3. Let $\overline{\varphi}^o(t,w) = U(w_t)$, where U(x) is given by (1.5). Then $\overline{\varphi}^o$ satisfies the following properties :

1) For arbitrary n (n=1,2,...), for any $g_i \in \underline{\underline{G}}$ (i=1,2,...,n) and subdivision $0\leqq t_1<t_2\dots<t_n\leqq T$ of [0,T]

$$E^{\overline{\varphi}^o}[E^{\overline{\varphi}^o}[g_1(\theta^o_{t_1},m^o_{t_1})|\underline{\underline{F}}^o_{t_1}]\times\dots\times E^{\overline{\varphi}^o}[g_n(\theta^o_{t_n},m^o_{t_n})|\underline{\underline{F}}^o_{t_n}]]$$
$$\leqq E^{\overline{\varphi}}[E^{\overline{\varphi}}[g_1(\theta_{t_1},m_{t_1})|\underline{\underline{F}}_{t_1}]\times\dots\times E^{\overline{\varphi}}[g_n(\theta_{t_n},m_{t_n})|\underline{\underline{F}}_{t_n}]]$$

for any $\overline{\varphi}\in\hat{\Phi}$.

2) For any t, $\rho_t(\overline{\varphi}^o)$ is $\tilde{\underline{\underline{B}}}{}^1$-measurable.

Moreover, if another $\overline{\varphi}\in\hat{\Phi}$ satisfies these properties then it holds that $\overline{\varphi}^o(t,w)=\overline{\varphi}(t,w)$ a.e. (dtdP).

The proof depends on the preceding discussions and the following lemma :

LEMMA 2.4. Let $\overline{\varphi}$ and $\overline{\varphi}'$ be in $\hat{\Phi}$ and satisfy the formula

$$E^{\overline{\varphi}}[\int_0^T g(\theta_t,m_t)dt] = E^{\overline{\varphi}'}[\int_0^T g(\theta'_t,m'_t)dt]$$

for all g in $\underline{\underline{G}}$, where m_t and m'_t are the conditional expectations of θ_t and θ'_t corresponding to $\overline{\varphi}$ and $\overline{\varphi}'$ respectively. Then for any $t\in[0,T]$

(2.11) $$E[\rho_t(\overline{\varphi}) \mid \sigma(|w_t|)] = E[\rho_t(\overline{\varphi}') \mid \sigma(|w_t|)] \quad \text{a.e.}(P),$$

where $\sigma(|w_t|)$ is the σ-field generated by the sets $\{|w_t|<a\}$, $a>0$.

PROOF. To see that such g generate $\sigma(|w_t|)$, it is enough to take $g(x,\alpha) = g_\varepsilon(x,\alpha) = I_{(a,\infty)}(|\varepsilon x+(1-\varepsilon)\alpha|)$, $\varepsilon>0$, $\alpha>0$.

REFERENCES

[1] M. FUJISAKI. On stochastic control of a Wiener process. J. Math. Kyoto University. 18-2 (1978) p. 229-238.

[2] N. IKEDA and S. WATANABE. A comparison theorem for solutions of stochastic differential equations and its applications. Osaka J. Math. 14, 1977, p. 619-633.

[3] R.S. LIPTZER and A.N. SHIRYAEV. Statistics of stochastic processes. Izd. Nauka, 1975.

[4] M. YOR. Les filtrations de certaines martingales du mouvement brownien dans $\mathbb{R}^n$. Dans ce volume.

M. Fujisaki
Kobe University of Commerce
Kobe, Japon .