W. H. Fleming and M. Nisio
Nagoya Math. J.
Vol. 93 (1984), 71–108

ON STOCHASTIC RELAXED CONTROL FOR PARTIALLY OBSERVED DIFFUSIONS

W. H. FLEMING AND M. NISIO

§ 1. Introduction

In this paper we are concerned with stochastic relaxed control problems of the following kind. Let $X(t)$, $t \geq 0$, denote the state of a process being controlled, $Y(t)$, $t \geq 0$, the observation process and $p(t, \cdot)$ a relaxed control, that is a process with values probability measures on the control region Γ. The state and observation processes are governed by stochastic differential equations

$$(1.1)\qquad \begin{cases} dX(t) = \alpha(X(t))dB(t) + \int_\Gamma \gamma(X(t), u)p(t, du)dt \\ X(0) = \xi \end{cases}$$

and

$$(1.2)\qquad \begin{cases} dY(t) = h(X(t))dt + dW(t) \\ Y(0) = 0 \end{cases}$$

where B and W are independent Brownian motions with values in R^n and R^m respectively, (put $m = 1$ for simplicity).

The problem is to maximize a criterion of the form

$$J = Ef(X(T))$$

by a suitable choice of admissible relaxed control p. In a customary version of stochastic control under partial observation, $p(t, \cdot)$ is measurable with respect to σ-field generated by the observation process $Y(s)$, $s \leq t$. Instead of discussing the problem of this type, we treat some wider class of admissible relaxed controls (see § 2), inspired by Fleming-Pardoux [8]. Roughly speaking, our problem is the following; Let

Received September 13, 1982.
This research was supported in part by the National Science Foundation under contract MCS 79-03554 and in part by the Air Force Office of Scientific Research under contract AF-AFOSR 81-0116.

$$L(T) = \exp\left\{\int_0^T h(X(s))dY(s) - \frac{1}{2}\int_0^T |h(X(s))|^2 ds\right\}. \tag{1.3}$$

Then B and Y turn out as independent Brownian motions under a new probability $\mathring{P}$, defined by

$$\frac{d\mathring{P}}{dP} = L^{-1}(T) \tag{1.4}$$

appealing to the so-called Girsanov transformation. For admissibility we merely require that $p(t, \cdot)$ is independent of future increments of $Y(\theta) - Y(s)$, θ, $s \geq t$, and B, with respect to $\mathring{P}$. Moreover we are concerned with $q(dt, du)$ instead of $p(t, du)dt$. (see Definition 1). Thus the criterion J can be expressed as

$$J = \mathring{E}f(X(T))L(T) \tag{1.5}$$

where $\mathring{E}$ stands for the expectation with respect to $\mathring{P}$, and $X(t)$ is a solution of the following system equation;

$$dX(t) = \alpha(X(t))dB(t) + \int_\Gamma \gamma(X(t), u)q(dt, du). \tag{1.6}$$

Under Lipschitz continuity and boundedness of α and γ, (1.6) has a unique solution (Theorem 1).

In Section 2 we introduce some metric spaces which are appropriate to our optimization problems. In Section 3 we prove the compactness of spaces of solutions and relaxed controls q. This guarantees the existence of optimal one (Theorem 3).

In the latter half we treat a nonlinear semigroup associated with relaxed control under partial observation. In this case we regard, as the state space, the unnormalized conditional distribution $\Lambda(t)$ of $X(t)$ given past observation and control. Hence $\Lambda(t)$ is a process valued in measures on R^n and satisfying the Zakai equation. Thus our problems turn out as optimization problems of measure valued processes. After we prove the continuity of $\Lambda(t)$ with respect to initial distribution $X(0)$ and data of past observation and control, (see Theorem 5), we construct a nonlinear semigroup $S(t)$, $t \geq 0$, on a Banach lattice of bounded and uniformly continuous functions, defined on the space of measures (Theorem 7). Following Fleming [6], we show that the generator of $S(t)$ relates to a dynamic programming equation, so-called Mortensen's equation.

§ 2. Notations and preliminaries

Let $(\Omega, F, \mathring{P})$ be a probability space. Let B and Y be n-dimensional and 1-dimensional Brownian motions, defined on $(\Omega, F, \mathring{P})$ respectively. Γ is a given convex compact subset of R^k, called a control region. $M(\Gamma)$ denotes the totality of positive finite measures defined on $\boldsymbol{B}(\Gamma)$(=Borel field of Γ). By $\hat{M}([0\ T] \times \Gamma)$ we denote the set of all mappings λ: $[0\ T] \times \boldsymbol{B}(\Gamma) \to [0\ T]$ such that

0) $\lambda(0, A) = 0 \quad {}^\forall A \in \boldsymbol{B}(\Gamma)$

i) $\lambda(t, \Gamma) = t \quad {}^\forall t \geq 0$

ii) $\lambda(t, \cdot) \in M(\Gamma)$ for all $t > 0$

iii) $\lambda(t, A)$ is increasing in t for all $A \in \boldsymbol{B}(\Gamma)$

iv) $\sup_{A \in \boldsymbol{B}(\Gamma)} |\lambda(s, A) - \lambda(t, A)| = |t - s|$.

From (ii) and (iii), λ determines a measure on $[0\ T] \times \Gamma$ and $\lambda([s, t], \cdot) \equiv \lambda(t, \cdot) - \lambda(s, \cdot) \in M(\Gamma)$, if $t > s$, and $\lambda([s, t], \Gamma) = t - s$.

Let q be a mapping; $[0\ T] \times \boldsymbol{B}(\Gamma) \times \Omega \to [0, 1]$, such that

v) for all $A \in \boldsymbol{B}(\Gamma)$, $q(\cdot, A, \cdot)$ is $\boldsymbol{B}[0\ T] \times F$-measurable

vi) $q \in \hat{M}([0\ T] \times \Gamma)$ $\mathring{P}$-almost surely.

Definition 1. $\mathscr{A} = (\Omega, F, \mathring{P}, \xi, B, Y, q)$ is called an admissible (relaxed) system, if ξ is an n-random vector on $(\Omega, F, \mathring{P})$, which is independent of (B, Y, q), B and $(Y\ q)$ are independent and the increments $(Y(t) - Y(s), t \geq s)$ are independent of $\sigma_s(Y, q)$(= σ-field generated by $Y(\theta)$, $\theta \leq s$ and $q(\theta, A)$, $\theta \leq s$, $A \in \boldsymbol{B}(\Gamma)$).

Definition 2. The component q of $\mathscr{A}$ is called a relaxed control, and we denote it by $q_{\mathscr{A}}$ when $\mathscr{A}$ is stressed. $\lambda \in \hat{M}([0\ T] \times \Gamma)$ can be regarded as a relaxed control. $\mathfrak{A}$ denotes the totality of admissible systems.

Let $\alpha(x)$ be a symmetric $n \times n$ matrix valued function on R^n and γ an n-vector continuous function on $R^n \times \Gamma$. We assume the following conditions

(A1) $|g(x, u)| \leq b, \quad {}^\forall x \in R^n,\ u \in \Gamma \ \ g = \alpha, \gamma$

(A2) $|g(x, u) - g(x', u)| \leq K|x - x'|, \ {}^\forall x, x' \in R^n\ u \in \Gamma, g = \alpha, \gamma.$

For an admissible system $\mathscr{A} = (\Omega, F, \mathring{P}, \xi, B, Y, q)$ we consider the stochastic differential equation (SDE in short)

$$
(2.1) \qquad \begin{cases} dX(t) = \alpha(X(t)dB(t) + \int_\Gamma \gamma(X(t), u)q(dt, du) \\ X(0) = \xi . \end{cases}
$$

THEOREM 1. *There exists a unique solution X of (2.1) which is $\sigma_t(\xi, B, q)$-progressively measurable and has continuous paths.*

Proof. We apply a usual successive approximation method. We define X_n in the following way

$$X_0(t) = \xi$$

$$X_{n+1}(t) = \xi + \int_0^t \alpha(X_n(s))dB(s) + \int_0^t \int_\Gamma \gamma(X_n(s), u)q(ds, du) \tag{2.2}$$
$$n = 0, 1, 2, \cdots .$$

Then, X_n is $\sigma_t(\xi, B, q)$-progressively measurable and has continuous paths by (iv) and (A1).

$$X_{n+1}(t) - X_n(t) = \int_0^t (\alpha(X_n(s)) - \alpha(X_{n-1}(s))dB(s)) + \int_0^t \int_\Gamma (\gamma(X_n(s), u) - \gamma(X_{n-1}(s), u))q(ds, du)$$

So, using (A1) and (A2) we see

$$\begin{aligned}
&\left(\int_0^t \int_\Gamma |\gamma(X_n(s), u) - \gamma(X_{n-1}(s), u)|\, q(ds, du)\right)^2 \\
&\quad \leq \int_0^t \int_\Gamma |\gamma(X_n(s), u) - \gamma(X_{n-1}(s), u)|^2 q(ds, du) q(t, \Gamma) \\
&\quad \leq Kt \int_0^t \int_\Gamma |X_n(s) - X_{n-1}(s)|^2 q(ds, du) \\
&\quad = Kt \int_0^t |X_n(s) - X_{n-1}(s)|^2 ds .
\end{aligned}$$

Putting $\rho_n(t) = E|X_{n+1}(t) - X_n(t)|^2$, we have

$$\rho_n(t) \leq K_1 \int_0^t \rho_{n-1}(s)ds , \qquad \text{for } \forall t \leq T,$$

with some $K_1 = K_1(T)$. This implies

$$\rho_n(t) \leq \frac{t^{n-1}K_1^{n-1}}{(n-1)!} \mathring{E}|\xi|^2 . \tag{2.3}$$

Therefore

$$\sum_n \mathring{E}|X_{n+1}(t) - X_n(t)| \leq \sum \sqrt{\rho_n(t)} < \infty .$$

This implies that $X_n(t)$ converges $\mathring{P}$-almost surely. Hence $X(t) = \lim_{n\to\infty} X_n(t)$

can be regarded as $\sigma_t(\xi, B, q)$-progressively measurable and moreover a martingale inequality tells us that, as $n \to \infty$,

$$\int_0^t \alpha(X_n(s))dB(s) \to \int_0^t \alpha(X(s))dB(s) \quad \text{uniformly in } t \in [0, T], \tag{2.4}$$

$\mathring{P}$-almost surely. On the other hand

$$\begin{aligned} \sup_{t \le T} &\left| \int_0^t \int_\Gamma \gamma(X_n(s), u) - \gamma(X(s), u) q(ds, du) \right. \\ &\le \int_0^T \int_\Gamma |\gamma(X_n(s), u) - \gamma(X(s), u)| \, q(ds, du) \\ &\le K \int_0^T \min(|X_n(s) - X(s)|, 2b) ds . \end{aligned}$$

By virtue of the convergence theorem we get, as $n \to \infty$,

$$\int_0^t \int_\Gamma \gamma(X_n(s), u) q(ds, du) \to \int_0^t \int_\Gamma \gamma(X(s), u) q(ds, du), \tag{2.5}$$

uniformly in $t \in [0, T]$, $\mathring{P}$-almost surely.

Combining (2.4) and (2.5) with (2.2), X turns out as a solution of (2.1) and $X_n(t)$ converges to $X(t)$ uniformly in $t \in [0 \ T]$ $\mathring{P}$-almost surely. Hence X has continuous paths.

Let Y be a $\sigma_t(\xi, B, q)$-progressively measurable solution of (2.1). Then, applying a routine method, we can easily see for $\forall t$,

$$X(t) = Y(t) \qquad \mathring{P}\text{-almost surely.}$$

This completes the proof of Proposition 2.1.

L denotes the Prohorov metric for probability measures. That is following, [11]. Let ε_{21} be the infinimum of ε such that

$$\mu_1(F) \le \mu_2(U_\varepsilon(F)) + \varepsilon \qquad \text{for all closed subset } F,$$

where $U_\varepsilon(F)$ is the ε-neighbourhood of F. ε_{12} is defined by switching μ_1 and μ_2. Set

$$L(\mu_1, \mu_2) = \max(\varepsilon_{12}, \varepsilon_{21}) . \tag{2.6}$$

Put

$$M(\Gamma, t) = \{\lambda \in M(\Gamma); \lambda(\Gamma) = t\}, \ t > 0 . \tag{2.7}$$

Define a metric ρ_t as follows

$$\rho_t(\lambda_1, \lambda_2) = L\left(\frac{\lambda_1}{t}, \frac{\lambda_2}{t}\right), \qquad \lambda_i \in M(\Gamma, t) . \tag{2.8}$$

Since Γ is compact, $(M(\Gamma, t), \rho_t)$ is a compact metric space. Put $D = \{r_i,$ rational $\in [0, T]$, $i = 1, 2, \cdots\}$ and

$$\tilde{M}_T = \prod_{i=1}^{\infty} M(\Gamma, r_i) . \tag{2.9}$$

We endow $\tilde{M}_T$ with a metric $\tilde{d}_T$ such that

$$\tilde{d}_T(\tilde{\lambda}, \tilde{\mu}) = \sum_{n=1}^{\infty} \frac{1}{2^n} \rho_{r_n}(\lambda_n, \mu_n) \tag{2.10}$$

where $\tilde{\lambda} = (\lambda_1, \lambda_2, \cdots)$ and $\tilde{\mu} = (\mu_1, \mu_2, \cdots)$. Hence $\tilde{\lambda}_k$, $k = 1, 2, \cdots$ is a $\tilde{d}_T$-Cauchy sequence, iff $\lambda_{k,i}$, $k = 1, 2, \cdots$ is a ρ_{r_i}-Cauchy sequence for any component i. Therefore again $(\tilde{M}_T, \tilde{d}_T)$ is a compact metric space.

Since $\lambda \in \tilde{M}([0\ T] \times \Gamma)$ is determined by $\tilde{\lambda} = (\lambda(r_1), \lambda(r_2), \cdots) \in \tilde{M}_T$, we define a metric $\hat{d}_T$ on $\hat{M}([0\ T] \times \Gamma)$ by

$$\hat{d}_T(\lambda, \mu) = \tilde{d}_T(\tilde{\lambda}, \tilde{\mu}) . \tag{2.11}$$

PROPOSITION 2.1. *$\hat{M}([0\ T] \times \Gamma, \hat{d}_T)$ is a compact metric space.*

Proof. Let $\lambda_k(r_i, \cdot)$ converge to $\lambda_{(i)}$ in ρ_{r_i}. Then $\lambda_{(i)} \in M(\Gamma, r_i)$ and for $g \in C_b(\Gamma)$(=bounded continuous function on Γ).

$$\int_\Gamma g(u)\lambda_k(r_i, du) \to \int g(u)\lambda_{(i)}(du) , \qquad \text{as } k \to \infty . \tag{2.12}$$

Define $\lambda(r_i, A)$ by $\lambda(r_i, A) = \lambda_{(i)}(A)$. Then putting $g = 1$ in (2.12), we see

$$\lambda(r_i, \Gamma) = r_i.$$

Let $r_i > r_j$ and set $R(A) = \lambda(r_i, A) - \lambda(r_j, A)$. Then R is a signed measure on Γ. Since $\lambda_k(r_i, \cdot) - \lambda_k(r_j, \cdot) \in M(\Gamma, r_i - r_j)$ and $\lambda_k(r_i, \cdot) - \lambda_k(r_j, \cdot)$ converges to R weakly by (2.12). R turns out as a positive measure and $R \in M(\Gamma, r_i - r_j)$. This means

$$\lambda(r, A) \text{ is increasing in rational } r \tag{2.13}$$

and

$$|\lambda(r_i, A) - \lambda(r_j, A)| \leq |\lambda(r_i, \Gamma) - \lambda(r_j, \Gamma)| = r_i - r_j . \tag{2.14}$$

Now we will construct λ which corresponds to $(\lambda(r_i, \cdot), i = 1, 2, \cdots) \in \tilde{M}_T$, in the following way,

$$\lambda(t, A) = \lim_{r_i \uparrow t} \lambda(r_i, A) . \tag{2.15}$$

Then λ clearly satisfies the conditions (0) $\sim$ (iv), namely $\lambda \in M([0\ T] \times \Gamma)$ and $\hat{d}_T(\lambda_k, \lambda) \to 0$, as $k \to \infty$.

Remark. $\rho_t(\lambda_k(t), \lambda(t)) \le (4/r_i)|t - r_i| + \rho_{r_i}(\lambda_k(r_i), \lambda(r_i))$ by condition (iv). Hence

$$\lambda_k(t) \to \lambda(t) \qquad \text{in } \rho_t .$$

For $g \in C([0\ T] \times \Gamma)$ ($=$ continuous function on $[0\ T] \times \Gamma$)

$$\int_0^t g(s, u)\lambda_k(ds, du) \to \int_0^t g(s, u)\lambda(ds, du) , \qquad \text{as } k \to \infty . \tag{2.16}$$

Proof of (2.16). Since g is uniformly continuous, $g([2^n t]/2^n, u)$ converges to $g(t, u)$ uniformly on $[0\ T] \times \Gamma$, where $[c]$ is the largest integer $\le c$. Suppose $\sup_{s,u}|g([2^n s]/2^n, u) - g(s, u)| < \varepsilon$. Then

$$\begin{aligned}
&\left|\int_0^t\int_\Gamma g(s, u)\lambda_k(ds, du) - \int_0^t\int_\Gamma g(s, u)\lambda(ds, du)\right| \\
&\quad \le \int_0^t\int_\Gamma \left|g(s, u) - g\left(\frac{[2^n s]}{2^n}, u\right)\right|(\lambda_k(ds, du) + \lambda(ds, du)) \\
&\quad\quad + \left|\int_0^t\int_\Gamma g\left(\frac{[2^n s]}{2^n}, u\right)\lambda_k(ds, du) - \int_0^t\int_\Gamma g\left(\frac{[2^n s]}{2^n}, u\right)\lambda(ds, du)\right| \\
&\quad \le 2\varepsilon t + \text{2nd term.}
\end{aligned}$$

Appealing to (2.15), we see that the 2nd term tends to 0, as $k \to \infty$, for any n. Hence we can conclude (2.16).

Let ζ_i, $i = 1, 2$, be $\tilde{M}_T$(or $\hat{M}([0\ T] \times \Gamma)$)-valued random variables, which may be defined on different probability spaces. ν_i denotes the probability distribution of ζ_i. So, ν_i is a probability on $(\tilde{M}_T, \tilde{d}_T)$ (or $(\hat{M}([0\ T] \times \Gamma), \hat{d}_T)$ respectively). Let $\tilde{m}_T$(or $\hat{m}_T$) denote the totality of $\tilde{M}_T$(or $\hat{M}([0\ T] \times \Gamma)$)-valued random variables. We endow the following Prohorov metric $\tilde{D}_T$(or $\hat{D}_T$) on $\tilde{m}_T$(or $\hat{m}_T$ resp.),

$$\begin{aligned} \tilde{D}_T(\zeta_1, \zeta_2) &= L(\nu_1, \nu_2) \\ \hat{D}_T(\zeta_1, \zeta_2) &= L(\nu_1, \nu_2) . \end{aligned} \tag{2.17}$$

Since $\tilde{M}_T$ and $\hat{M}([0\ T] \times \Gamma)$ are compact metric spaces, $(\tilde{m}_T, \tilde{D}_T)$ and $(\hat{m}_T, \hat{D}_T)$ are also compact spaces.

For $\mathscr{A} = (\Omega, F, \mathring{P}, \xi, B, Y, q)$, we sometimes denote ξ by $\xi_{\mathscr{A}}$ and so on, when any confusion might occur. Let $X(= X_{\mathscr{A}})$ be a solution of (2.1) for $\mathscr{A}$. Then (X, ξ, B, Y, q) becomes a $M_T = C([0, T] \to R^n) \times R^n \times C([0\ T] \to R^n) \times ([0\ T] \to R^1) \times \hat{M}([0\ T] \times \Gamma)$ valued random variable. Endowing M_T

with a usual metric $d_T(=$ sum of metric of each component) M_T turns out as a complete separable metric space. Let m_T denote the totality of M_T-valued random variables and D_T the Prohorov metric on m_T. Hereafter we denote (X, ξ, B, Y, q) by $(X, \mathscr{A})$ for simplicity if no confusion occurs. We also say that $\mathscr{A}_n \to \mathscr{A}$ (in Prohorov topology), if $(\xi_n, B_n, Y_n, q_n) \to (\xi, B, Y, q)$ in Prohorov topology. Ω_n, F_n, P_n can also depend on n. $\xi_{\mathscr{A}}$, $B_{\mathscr{A}}$ and $(Y_{\mathscr{A}}, q_{\mathscr{A}})$ are independent for any $\mathscr{A} \in \mathfrak{A}$ and $B_{\mathscr{A}}$ is a Brownian motion. Therefore we have

PROPOSITION 2.2. *$\mathscr{A}_n \to \mathscr{A}$, iff $\xi_n \to \xi$ in law and $(Y_n, q_n) \to (Y, q)$ in Prohorov topology.*

Now we put the set $\mathscr{P}_\mu =$ totality of probability distributions of $(Y_{\mathscr{A}}, q_{\mathscr{A}})$, $\mathscr{A} \in \mathfrak{A}(\mu)$, $\mathfrak{A}(\mu)$ defined later (3.8). Since $\xi_{\mathscr{A}}$, $B_{\mathscr{A}}$ and $(Y_{\mathscr{A}}, q_{\mathscr{A}})$ are independent for $\mathscr{A} \in \mathfrak{A}$, $\mathscr{P}_\mu$ does not depend on μ, say $\mathscr{P}$. Moreover $\pi \in \mathscr{P}$, iff π is a probability on $C([0\ T] \to R^1) \times \hat{M}([0\ T] \times \Gamma)$ such that the first component y is a Brownian motion under π and its increments $y(t) - y(s)$ is independent of $\sigma_s(y, \lambda)$ for $t > s$, where λ is the second component, (see § 2 of Fleming-Pardoux [8]). Since $C([0\ T] \to R^1) \times \hat{M}([0\ T] \times \Gamma)$ is a product of complete separable metric space, it becomes a complete separable metric space. So we introduce the Prohorov topology on $\mathscr{P}$. Then $\mathscr{P}$ is a compact metric space, because the first component is a Browian motion and $\hat{M}([0\ T] \times \Gamma)$ is a compact metric space. Now we have

PROPOSITION 2.3.

i) *$\mathscr{P} =$ totality of probability distribution of $(Y_{\mathscr{A}}, q_{\mathscr{A}})$, $\mathscr{A} \in \mathfrak{A}$, is a compact metric space with Prohorov metric.*

ii) *$\mathscr{P} =$ totality of probability distribution of $(Y_{\mathscr{A}}, q_{\mathscr{A}})$, $\mathscr{A} \in \mathfrak{A}(\mu)$, for any μ.*

iii) *For $\mathscr{A}, \mathscr{A}' \in \mathfrak{A}(\mu)$, $D_T((X_{\mathscr{A}}, \mathscr{A}), (X_{\mathscr{A}'}, \mathscr{A}')) = 0$ iff probability distribution of $(Y_{\mathscr{A}}, q_{\mathscr{A}}) =$ probability distribution of $(Y_{\mathscr{A}'}, q_{\mathscr{A}'})$.*

§ 3. Existence of optimal relaxed control

Let N be a compact subset of probability measures on R^n with Prohorov metric. Put $P_\eta =$ Probability law of η and

$$\mathfrak{N} = \{(X, \mathscr{A});\ P_{\xi_{\mathscr{A}}} \in N\}.$$

PROPOSITION 3.1. *$\mathfrak{N}$ is a compact subset of (m_T, D_T).*

Proof. By the condition (A1), $\{X_{\mathscr{A}};\ P_{\xi_{\mathscr{A}}}\in N\}$ is totally bounded in Prohorov topology. $B_{\mathscr{A}}$ and $Y_{\mathscr{A}}$ are Brownian motions for any $\mathscr{A}\in\mathfrak{A}$. Since $\hat{m}_T$ is compact, $\{q_{\mathscr{A}};\ \mathscr{A}\in\mathfrak{A}\}$ is totality bounded. Therefore

$$\mathfrak{N}=\{(X,\mathscr{A});\ P_{\xi_{\mathscr{A}}}\in N\}$$

is a totally bounded subset of (m_T, D_T).

Now we will show that $\mathfrak{N}$ is closed. Let $(X_k, \mathscr{A}_k)$, $k=1,2,\cdots$ be a Cauchy sequence. Using Skorobod's theorem, we can construct $(X_k^*, \xi_k^*, B_k^*, Y_k^*, q_k^*)$ and $(X^*, \xi^*, B^*, Y^*, q^*)$ on a probability space $(\Omega^*, F^*, \mathring{P}^*)$, so that

(3.1) $(X_k^*, \xi_k^*, B_k^*, Y_k^*, q_k^*)$ has the same probability law as $(X_k, \xi_k, B_k, Y_k, q_k)$, $k=1,2,\cdots$.

(3.2) As $k\to\infty$, $(X_k^*, \xi_k^*, B_k^*, Y_k^*, q_k^*)$ converges to $(X^*, \xi^*, B^*, Y^*, q^*)$ in d_T metric, $\mathring{P}^*$-almost surely.

Hence ξ^*, B^* and (Y^*, q^*) are independent and B^* and Y^* are Brownian motions. Moreover we see that, for a.a. $\omega(\mathring{P}^*)$, $q^*(\cdot,\omega)\in\hat{M}([0\ T]\times\Gamma)$ and $\hat{d}_T(q_k^*(\cdot,\omega), q^*(\cdot,\omega))$ tends to 0 as $k\to\infty$, by virtue of Proposition 2.2. On the other hand (2.15) implies that $q^*(t, A, \cdot)$ is F^*-measurable. Since $q(t, A, \omega)$ is continuous in t, $q^*(\cdot, A, \cdot)$ is $\boldsymbol{B}_1[0\ T]\times F^*$-measurable. Namely q^* satisfies the conditions (v) and (vi). (3.2) again tells us that $Y^*(t)-Y^*(r_i)$ is independent of $\sigma_s(Y^*, q^*)$ whenever $s\le r_i\le t$. Since Y^* has continuous paths, this implies that $Y^*(t)-Y^*(s)$ is independent of $\sigma_s(Y^*, q^*)$, Therefore $\mathscr{A}^*=(\Omega^*, F^*, \mathring{P}^*, \xi^*, B^*, Y^*, q^*)\in\mathfrak{A}$.

Next we will show that X^* is a solution of (2.1) for $\mathscr{A}^*$.

$$(3.3)\quad \left|\int_0^t\int_\Gamma \gamma(X_k^*(s,\omega),u)q_k^*(ds,du,\omega)-\int_0^t\int_\Gamma \gamma(X^*(s,\omega),u)q^*(ds,du,\omega)\right|$$
$$\le\int_0^t\int_\Gamma|\gamma(X_k^*(s,\omega),u)-\gamma(X^*(s,\omega),u)|q_k^*(ds,du,\omega)$$
$$+\left|\int_0^t\int_\Gamma \gamma(X^*(s,\omega),u)(q_k^*(ds,du,\omega)-q^*(ds,du,\omega))\right.$$

$$\text{the 1st term}\quad \le K\int_0^t|X_k^*(s,\omega)-X^*(s,\omega)|\,q_k^*(ds,\Gamma,\omega)$$
$$=K\int_0^t|X_k^*(s,\omega)-X^*(s,\omega)|\,ds\,.$$

Since $X_k^*(\cdot,\omega)$ converges to $X^*(\cdot,\omega)$ uniformly in $[0,T]$, the 1st term converges to 0, as $k \to \infty$. Appealing to Remark of Proposition 2.1, the 2nd term converges to 0, as $k \to \infty$. So, we have

$$\int_0^t \int_\Gamma \gamma(X_k^*(s, \omega), u) q_k^*(ds, du, \omega) \longrightarrow \int_0^t \int_\Gamma \gamma(X^*(s, \omega) q^*(ds, du, \omega). \tag{3.4}$$

Using a routine method we get

$$\int_0^t \alpha(X_k^*(s) dB_k^*(s)) \longrightarrow \int_0^t \alpha(X^*(s) dB^*(s)) \qquad \text{in proba } (\mathring{P}^*). \tag{3.5}$$

From (3.4) and (3.5), we conclude that X^* is a solution of (2.1) for $\mathscr{A}^*$. This completes the proof of Proposition 3.1.

COROLLARY. *If $\mathscr{A}_k \to \mathscr{A}$, then $(X_k, \mathscr{A}_k) \to (X, \mathscr{A})$ in D_T.*

Let f and h be bounded and uniformly continuous functions on R^n. Define a pay-off function $J(\mathscr{A})$ as follows,

$$J(\mathscr{A}) = \mathring{E} f(X_{\mathscr{A}}(T)) L(T, \mathscr{A}) \tag{3.6}$$

where $\mathring{E}$ stands for the expectation in $(\Omega, F, \mathring{P})$, and

$$L(T, \mathscr{A}) = \exp\left(\int_0^T h(X(s)) dY(s) - \frac{1}{2} \int_0^T |h(X(s))|^2 ds\right) \tag{3.7}$$

where $X = X_{\mathscr{A}}$ and $Y = Y_{\mathscr{A}}$.

For a probability measure μ, we denote

$$\mathfrak{A}(\mu) = \{\mathscr{A} \in \mathfrak{A};\ P_{\xi_{\mathscr{A}}} = \mu\} \tag{3.8}$$

i.e. the set of all admissible system where initial distribution equals to μ.

For a given μ we want to maximize $J(\mathscr{A})$ by a suitable choice of $\mathscr{A} \in \mathfrak{A}(\mu)$.

THEOREM 2. *There exists an optimal admissible system $\tilde{\mathscr{A}} \in \mathfrak{A}(\mu)$, that is*

$$\sup_{\mathscr{A} \in \mathfrak{A}(\mu)} J(\mathscr{A}) = J(\tilde{\mathscr{A}}). \tag{3.9}$$

Proof. Let $\mathscr{A}_k \in \mathfrak{A}(\mu)$ be approximately optimal, i.e.

$$\lim_{k \to \infty} J(\mathscr{A}_k) = \sup_{\mathscr{A} \in \mathfrak{A}(\mu)} J(\mathscr{A}). \tag{3.10}$$

By virtue of Proposition 3.1, some subsequence $(X_{k_i}, \mathscr{A}_{k_i})$ converges to $(X, \mathscr{A})$ in Prohorov topology. For simplicity we may assume $(X_k, \mathscr{A}_k) \to$

$(X, \mathscr{A})$ as $k \to \infty$. Again Skorobod's theorem tells us that their suitable version satisfy (3.1) and (3.2). So we again assume that $(X_k, \mathscr{A}_k)$ and $(X, \mathscr{A})$ satisfy (3.1) and (3.2), since $J(\mathscr{A})$ depends on only probability law.

From boundedness of f and h, we have

$$(3.11) \qquad \mathring{E}(f(X_k(T))L(T, \mathscr{A}_k))^2 \leq \|f\|^2 e^{2T\|h\|}, \qquad k = 1, 2, \cdots .$$

Hence $\{f(X_k(T))L(T, \mathscr{A}_k), k = 1, 2, \cdots\}$ is uniformly integrable. On the other hand $L(T, \mathscr{A}_k)$ tends to $L(T, \mathscr{A})$ in proba. Appealing to the convergence theorem we get

$$(3.12) \qquad \lim_{k \to \infty} J(\mathscr{A}_k) = J(\mathscr{A}) .$$

Combining (3.12) with (3.10), we complete the proof.

Remark. Appealing to Corollary of Proposition 3.1, we see that if $\mathscr{A}_k \to \mathscr{A}$, then $J(\mathscr{A}_k) \to J(\mathscr{A})$.

Now we treat the following case; $r(x, u) = b_1(x) + b_2(x)u$ where $b_2(x)$ is $n \times k$ matrix.

THEOREM 3. *If $r(x, u) = b_1(x) + b_2(x)u$, then $q = q_{\mathscr{A}}$ can be replaced by a Γ-valued $\sigma_t(q)$-progressively measurable process U* (i.e. *usual admissible control under partial observation*). *That is, $X = X_{\mathscr{A}}$ is a unique solution of the following S.D.E.*

$$(3.13) \qquad \begin{cases} dX(t) = \alpha(X(t))dB(t) + \gamma(X(t), U(t))dt \\ X(0) = \xi \end{cases}$$

where $B = B_{\mathscr{A}}$ and $\xi = \xi_{\mathscr{A}}$.

Proof. Our required U is obtained by the following lemma.

LEMMA. *There exists a $\sigma_t(q)$-progressively measurable Γ-valued process U such that*

$$(3.14) \qquad \int_0^t \int_\Gamma uq(ds, du) = \int_0^t U(s)ds, \qquad \textit{for } \forall t \leq T.$$

Proof. Define U_k as follows

$$(3.15) \qquad U_k(t, \omega) = \begin{cases} 2^k \int_\Gamma u(q(t, du, \omega) - q(t - 2^{-k}, du, \omega)), & \text{for } t \geq 2^{-k} \\ \dfrac{1}{t} \int_\Gamma uq(t, du, \omega), & \text{for } 0 < t < 2^{-k}. \end{cases}$$

Since Γ is convex compact, $U_k(t, \omega) \in \Gamma$ and $\sigma_t(q)$-progressively measurable. Moreover the compactness of Γ tells us that $\{U_k, k = 1, 2, \cdots\}$ is weakly totally bounded in $L^2([0\ T] \times \Omega)$. Hence some subsequence converges weakly and their suitable convex combinations converge strongly, say $\sum_{p=\ell}^{N\ell} \sigma_p^\ell U_{n_p}(t, \omega)$ converges to $U(t, \omega)$ in $L^2([0\ T] \times \Omega)$, as $\ell \to \infty$. So U is a Γ-valued $\sigma_t(q)$-progressively measurable process.

On the other hand the definition of U_k implies

$$\int_0^t U_k(s, \omega) ds \to \int_\Gamma uq(t, du, \omega)\left(= \int_0^t \int_\Gamma uq(ds, du, \omega)\right). \tag{3.16}$$

as $k \to \infty$. Taking the convex combination of U_{n_p}, we can conclude that U satisfies (3.14) by bounded convergence theorem.

Now we return to the proof of Theorem 3. Since ξ, B and U are independent, (3.13) has a unique solution. So it is enough to show that $X = X_{\mathscr{A}}$ satisfies (3.13). By the Lemma we can see

$$\int_0^t b_2(X(s)) \int_\Gamma uq(ds, du) = \int_0^t b_2(X(s)) U(s) ds. \tag{3.17}$$

Using "$\int_0^t b_1(X(s)) q(ds, du) = \int_0^t b_1(X(s)) ds$", we have

$$\int_0^t \int_\Gamma \gamma(X(s), u) q(ds, du) = \int_0^t \gamma(X(s), U(s)) ds. \tag{3.18}$$

This completes the proof of Theorem 3.

Definition 3. A Γ-valued process U is called an admissible control under partial observation, if ξ, B and (Y, U) are independent and $Y(t) - Y(s)$ is independent of $\sigma_s(Y, U)$. Precisely speaking $\mathscr{A}_U = (\Omega, F, \mathring{P}, \xi, B, Y, U)$ is called an admissible usual system.

An admissible control U can be regarded as the following relaxed control q,

$$q(t, A, \omega) = \int_0^t \delta_{U(s,\omega)}(A) ds = |s \leq t;\ U(s, \omega) \in A| \tag{3.19}$$

where δ_a is the δ-measure at a. Appealing to Theorems 2 and 3, we can derived,

Corollary. *If $r(x, u) = b_1(x) + b_2(x)u$, then there exists an optimal admissible usual system $\tilde{\mathscr{A}}_U$. That is,*

$$J(\tilde{\mathscr{A}}_U) = \sup_{\mathscr{A}_U:\text{ad. usual sys}} J(\mathscr{A}_U) = \sup_{\mathscr{A} \in \mathfrak{A}(u)} J(\mathscr{A}). \tag{3.20}$$

This fact was directly proved by Haussmann [9] and in a slightly different form by Fleming-Pardoux [8].

§4. Approximation by usual controls

For $\mathscr{A} = (\Omega, F, \mathring{P}, \xi, B, Y, q)$ we define P_n by

$$P_n(t, A, \omega) = \begin{cases} \dfrac{1}{t} q(t, A, \omega) & 0 < t < 2^{-n} \\ (q(t, A, \omega) - q(t - 2^{-n}, A, \omega))2^n, & 2^{-n} \leq t, \end{cases} \tag{4.1}$$

namely P_n is an approximate time derivative of q. $P_n(\cdot, A, \cdot)$ is $\sigma_t(q)$-progressively measurable and $P_n(t, \cdot, \omega)$ is a probability on Γ. Define q_n by

$$q_n(t, A, \omega) = \int_0^t P_n(s, A, \omega) ds. \tag{4.2}$$

Then q_n satisfies the conditions (v) and (vi) and $\mathscr{A}_n = (\Omega, F, \mathring{P}, \xi, B, Y, q_n) \in \mathfrak{A}$. Since we have

$$|q_n(t, A, \omega) - q(t, A, \omega)| \leq 2^{-n} + 2^n \int_{t-2^{-n}}^t |q(s, A, \omega) - q(t, A, \omega)| ds, \quad \text{for a.a. } \omega(\mathring{P}), \tag{4.3}$$

the condition (iv) implies, as $n \to \infty$,

$$\sup_A |q_n(t, A, \omega) - q(t, A, \omega)| \to 0, \quad \text{uniformly on } [0, T]. \tag{4.4}$$

and

$$\hat{d}_T(q_n(\cdot, \omega), q(\cdot, \omega)) \longrightarrow 0. \tag{4.5}$$

Fix $u_0 \in \Gamma$ arbitrarily and define $P_{n,k}$ by

$$P_{n,k}(t, A, \omega) = \begin{cases} P_n\left(\dfrac{[2^k t]}{2^k}, A, \omega\right), & \text{for } t \geq 2^{-k} \\ \delta_{u_0}(A) & \text{for } t < 2^{-k}. \end{cases} \tag{4.6}$$

Then $P_{n,k}$ is a step function in the time variable t. Put $q_{n,k}$ as follows.

$$q_{n,k}(t, A, \omega) = \int_0^t P_{n,k}(s, A, \omega) ds. \tag{4.7}$$

We call $q_{n,k}$ a switching relaxed control with interval 2^{-k}. It is clear that $(\Omega, F, \mathring{P}, \xi, B, Y, q_{n,k}) \in \mathfrak{A}$ and

$$|q_{n,k}(t, A, \omega) - q_n(t, A, \omega)| \leq 2^{n-k-1} t + 2^{-k}. \tag{4.8}$$

Therefore we get

$$\lim_{k\to\infty} \sup_{t\le T, A} |q_{n,k}(t, A, \omega) - q_n(t, A, \omega)| = 0\,. \tag{4.9}$$

Now we conclude the following proposition,

PROPOSITION 4.1. *For $\mathscr{A} = (\Omega, F, \mathring{P}, \xi, B, Y, q) \in \mathfrak{A}$, there exists an approximate sequence of switching relaxed control q_k with interval 2^{-k}, such that $q_k(\cdot, A, \cdot)$ is $\sigma_t(q)$-progressively measurable and moreover*

$$\lim_{k\to\infty} \sup_{0\le t\le T, A} |q_k(t, A) - q(t, A)| = 0\,, \qquad \mathring{P}\text{-almost surely} \tag{4.10}$$

and

$$\lim_{k\to\infty} \hat{d}_T(q_k, q) = 0\,, \qquad \mathring{P}\text{-almost surely}. \tag{4.11}$$

Putting $\mathscr{A}_k = (\Omega, F, \mathring{P}, \xi, B, Y, q_k)$, we can see the following corollary, by virtue of the Remark of Theorem 2.

COROLLARY. *There exists an approximate admissible switching system $\mathscr{A}_k$, such that*

$$\mathscr{A}_k \longrightarrow \mathscr{A}\,, \qquad \text{as } k \to \infty\,. \tag{4.12}$$

Hence $J(\mathscr{A}_k)$ converges to $J(\mathscr{A})$.

THEOREM 4. *There exists a Γ-valued $\sigma_t(q)$-progressively measurable process U_k, such that*

$$q_k(t, A, \omega) = \int_0^t \delta_{U_k(s,\omega)}(A) ds \tag{4.13}$$

approximates q in the following sence; $\mathscr{A}_k = (\Omega, F, \mathring{P}, \xi, B, Y, q_k)$ satisfies (4.12).

Proof. By the Corollary of Proposition 4.1, we may assume that $\mathscr{A}$ is an admissible switching system with interval 2^{-N}. Appealing to a Chattering Lemma [5], we will construct our desired U_{k_t} in the following way.

Let $\{u_1, \cdots, u_m\}$ be an ε-net of Γ, and $V_1, \cdots, V_m \in \boldsymbol{B}(\Gamma)$ a partition of Γ such that

$$|u_i - u| < \varepsilon \qquad \text{for } {}^\forall u \in V_i\,. \tag{4.14}$$

Since a given q is a switching relaxed control, it can be written by

$$q(t, A, \omega) = \int_0^t p(s, A, \omega) ds$$

with p of step function in s. Define $\hat{p} = \hat{p}_\varepsilon$ and $\hat{q} = \hat{q}_\varepsilon$ as follows,

$$\hat{p}(s, \{u_i\}, \omega) = p(s, V_i, \omega) \tag{4.15}$$

and

$$\hat{q}(t, \{u_i\}, \omega) = \int_0^t \hat{p}(s, \{u_i\}, \omega) ds . \tag{4.16}$$

Then $\hat{p}(t, \cdot, \omega)$ is a discrete probability on Γ and for $\forall g \in C(\Gamma)$

$$\int_\Gamma g(u)\hat{p}_\varepsilon(t, du, \omega) \longrightarrow \int_\Gamma g(u)p(t, du, \omega) , \quad \text{as} \quad \varepsilon \downarrow 0 . \tag{4.17}$$

Define θ_i, $i = 0, \cdots, m$ as follows: Let $j^{2^{-N}} \leq s_1 < s_2 < (j+1)2^{-N}$,

$$\begin{aligned} &\theta_0(\omega) = s_1 \\ &\theta_i(\omega) = \sum_{\ell=1}^{i} \int_{s_1}^{s_2} \hat{p}(t, \{u_\ell\}, \omega)dt + s_1, \qquad i = 1, \cdots, m . \end{aligned} \tag{4.18}$$

Then $s_1 = \theta_0(\omega) \leq \theta_1(\omega) \leq \cdots \leq \theta_m(\omega) = s_2$ and

$$\begin{aligned} \int_{s_1}^{s_2} \int_\Gamma g(u)\hat{q}(dt, du, \omega) &= \int_{s_1}^{s_2} \int_\Gamma g(u)\hat{p}(t, du, \omega)dt \\ &= \sum_{i=1}^{m} g(u_i) \int_{s_1}^{s_2} \hat{p}(t, \{u_i\}, \omega)dt = \sum_{i=1}^{m} g(u_i)(\theta_i(\omega) - \theta_{i-1}(\omega)) \\ &= \int_{s_1}^{s_2} g(U(t, \omega))dt = \int_{s_1}^{s_2} \int_\Gamma g(u)\delta_{U(t,\omega)}(du)dt \end{aligned} \tag{4.19}$$

where

$$U(t, \omega) = U_{\varepsilon, s_1, s_2}(t, \omega) = u_i \qquad \text{on } [(\theta_{i-1}(\omega), \theta_i(\omega)) . \tag{4.20}$$

Therefore $U_{\varepsilon, s_1, s_2}(t)$ is $\sigma_{j2^{-N}}(q)$-measurable. Putting $\varepsilon = 2^{-k}$, $s_1 = \ell 2^{-k} (k > N)$, we define U_k by

$$\begin{aligned} U_k(t, \omega) &= U_{2^{-k}, \ell 2^{-k}, (\ell+1)2^{-k}}(t, \omega) \\ &\qquad \text{for } \ell 2^{-k} \leq t < (\ell + 1)2^{-k}, \ \ell = 0, 1, 2 \cdots . \end{aligned} \tag{4.21}$$

Consider the SDE

$$\begin{cases} d\xi_k(t) = \alpha(\xi_k(t))dB(t) + \gamma(\xi_k(t), U_k(t))dt \\ \xi_k(0) = \xi . \end{cases} \tag{4.22}$$

If we regard $U_k(t)$ as $\delta_{U_k(t)}$, then ξ_k turns out a solution of (2.1) for $\mathscr{A}_k = (\Omega, F, \mathring{P}, \xi, B, Y, q_k)$ where $q_k(t, A, \omega) = \int_0^t \delta_{U_k(s,\omega)}(A)ds$. Moreover (4.19) means

$$(4.23) \qquad q_k(\ell 2^{-k}, \cdot, \omega) = \hat{q}_{2^{-k}}(\ell 2^{-k}, \cdot, \omega), \qquad \ell = 0, 1, \cdots, [2^\ell T].$$

Hence, combining with (4.17), we can see that, as $k \to \infty$,

$$(4.24) \qquad \mathring{d}_T(q_k, q) \longrightarrow 0, \quad \mathring{P}\text{-almost surely.}$$

Evidently this completes the proof.

§ 5. Continuity of conditional expectation

According to [8] we define L pathwise. Hereafter we assume the following smoothness on h.

(A3) $h, \dfrac{\partial h}{\partial x_i}, \dfrac{\partial^2 h}{\partial x_i \partial x_j}$ $i, j = 1, \cdots, n$ are bounded and uniformly continuous.

Putting $X = X_{\mathscr{A}}$, Ito's formula tells us that

$$\int_0^T h(X(t))dY(t) = h(X(T))Y(T) - \int_0^T Y(t)dh(X(t))$$

$$(5.1) \qquad = h(X(T))Y(T) - \sum_{ij} \int_0^T Y(t)\frac{\partial h}{\partial x_i}(X(t))\alpha_{ij}(X(t))dB_j$$

$$\int_0^T Y(t)A(t, q)h(X(t))dt$$

where

$$(5.2) \qquad A(t, q)h = \frac{1}{2}\sum a_{ij}(x)\frac{\partial^2 h}{\partial x_i \partial x_j} + \sum R_i(t, x, \omega; q)\frac{\partial h}{\partial x_i}$$

with

$$(5.3) \qquad R(t, x, \omega; q) = \overline{\lim_{n\to\infty}}\, 2^n \int_\Gamma \gamma(x, u)(q(t, du, \omega) - q(t - 2^{-n}, du, \omega)).$$

So, R is $\sigma_t(q)$-progressively measurable for any x, and Lipschitz continuous with respect to x. Moreover for any (x, ω),

$$(5.4) \qquad R(t, x, \omega; q) = \frac{\partial}{\partial t}\int_\Gamma \gamma(x, u)q(t, du, \omega) \qquad \text{for a.a. } t.$$

For $y \in C([0\ T] \to R^n)$ and $\lambda \in \hat{M}([0\ T] \times \Gamma)$ we define $\mathscr{L}$ by

$$\mathscr{L}(\theta, \xi, B, y, \lambda)$$

$$(5.5) \qquad = \exp\Big[y(\theta)h(\eta(\theta)) - \int_0^\theta y(s)A(s, \lambda)h(\eta(s))ds - \frac{1}{2}\int_0^\theta |h(\eta(s))|^2 ds$$

$$- \sum_{ij}\int_0^s y(s)\frac{\partial h}{\partial x_i}(\eta(s))\alpha_{ij}(\eta(s))dB_j(s)\Big]$$

where η is a solution of S.D.E.

$$\begin{cases} d\eta(t) = \alpha(\eta(t))dB(t) + \int_\Gamma \gamma(y(t), u)\lambda(dt, du) \\ \eta(0) = \xi . \end{cases} \tag{5.6}$$

Applying a successive approximation, we can see that η is a Borel function of ξ, B and λ. Hence $\mathscr{L}$ is Borel measurable with respect to θ, ξ, B, y, λ. From (A1) and (A3) we have the following evaluation,

$$\begin{aligned} e^{-F(y,\theta)} \exp\Big[-\sum_{ij}\int_0^\theta y(t)\frac{\partial h}{\partial x_i}(\eta(t))\alpha_{ij}(\eta(t))dB_j(t) \\ - \frac{1}{2}\int_0^t \sum_j \Big(y(t)\sum_i \frac{\partial h}{\partial x_i}(\eta(t))\alpha_{ij}(\eta(t))\Big)^2 dt\Big] \\ \leq \mathscr{L}(\theta, \xi, B, y, \lambda) \leq e^{F(y,\theta)} \exp\Big[-\sum_{ij}\int_0^\theta \cdots dB_j(t) - \frac{1}{2}\int_0^t \cdots dt\Big] \end{aligned} \tag{5.7}$$

where $$F(y, \theta) = K_1(\sup_{t\leq\theta}|y(t)| + 1)^2(\theta + 1) .$$

Since $(C[0\ T] \to R^n), \|\ \ \|)$ and $(\hat{M}[0\ T) \times \Gamma, \hat{d}_T)$ are complete separable metric spaces, the regular conditional probablity $\mathring{P}((X, \xi, B, Y, q) \in \cdot / Y = y, q = \lambda)$ exists. This regular conditional probability is nothing but the probability distribution of $(\eta, \xi, B, y, \lambda)$, because (ξ, B) and (Y, q) are independent. Putting $\mu = P_\xi(=$ probability distribution of $\xi)$, we have a version of conditional expectation as follows.

$$\begin{aligned} \mathring{E}(f(X(\theta))L(\theta, \mathscr{A})/Y = y, q = \lambda) &= \mathring{E}f(\eta(\theta))\mathscr{L}(\theta, \xi, B, y, \lambda) \\ &= \int \mathring{E}f(\eta(\theta, x))\mathscr{L}(\theta, x, B, y, \lambda)d\mu(x), \qquad \text{for bounded Borel } f, \end{aligned} \tag{5.8}$$

where $\eta(\theta, x)$ is a solution of (5.6) with $\eta(0, x) = x$. The right side of (5.8) is Borel measurable with respect to θ, y, λ, which depends on f and μ. So we denote the right side of (5.8) by $C(\theta, y, \lambda, \mu, f)$. Moreover $C(\theta, y, \lambda, \mu, f)$ depends on the value of y and λ up to time θ. Stressing μ we denote $\mathring{E}$ by $\mathring{E}_\mu$. That is,

$$\begin{aligned} C(\theta, Y, q, \mu, f) &= \mathring{E}_\mu(f(X(\theta))L(\theta, \mathscr{A})/\sigma_T(Y, q)) \\ &= \mathring{E}_\mu(f(X(\theta))L(\theta, \mathscr{A})/\sigma_\theta(Y, q)), \qquad \mathring{P}\text{-almost surely.} \end{aligned} \tag{5.9}$$

Using (5.7) we have

$$e^{-F(Y,\theta)} \leq C(\theta, Y, q, \mu, 1) \leq e^{F(Y,\theta)} . \tag{5.10}$$

Now we define $C(\theta, Y, q, \nu, f)$ for a positive measure ν as follows

(5.11) $$C(\theta,\ Y,\ q,\ \nu,\ f) = \|\nu\|\, C\Big(\theta,\ Y,\ q,\ \frac{\nu}{\|\nu\|},\ f\Big)$$

where $\|\nu\| = \nu(R^n)$ and we apply the same notations $\mathring{E}_\nu$ for a general positive measure ν, that is

(5.12) $$\mathring{E}_\nu(f(X(\theta))L(\theta,\ \mathscr{A})/\sigma_\theta(Y,\ q)) = C(\theta,\ Y,\ q,\ \nu,\ f)$$

with $\mathscr{A} = (\Omega, F, \mathring{P}, \xi, B, Y, q)$ for $P_\xi = \nu/\|\nu\|$. The left side of (5.12) stands for

(5.13) $$\mathring{E}_\nu(f(X(\theta))L(\theta,\ \mathscr{A})/\sigma_\theta(Y,\ q)) = \|\nu\|\mathring{E}_{\nu/\|\nu\|}(f(X(\theta))L(\theta,\ \mathscr{A})/\sigma_\theta(Y,\ q))\,.$$

Since ξ is independent of $(B,\ Y,\ q)$, the right side of (5.13) does not depend on a special choice of ξ.

Define $\Lambda(\theta, y, \lambda, \nu)(A)$ by

(5.14) $$\Lambda(\theta, y, \lambda, \nu)(A) = C(\theta, y, \lambda, \nu, \chi_A), \quad A \in \boldsymbol{B}_n\,.$$

Then $\Lambda(\theta, y, \lambda, \nu)$ is a positive measure on R^n and for any bounded Borel function f.

(5.15) $$\langle f, \Lambda(\theta, y, \lambda, \nu)\rangle = C(\theta, y, \lambda, \nu, f)$$

where $\langle f, \Lambda\rangle = \int_{R^n} f(x)\Lambda(dx)$. From (5.10) we see

(5.16) $$\|\nu\|e^{-F(y,\theta)} \leq \|\Lambda(\theta, y, \lambda, \nu)\| \leq \|\nu\|e^{F(y,\theta)}\,.$$

On the other hand $\mathscr{L}(\theta, \xi, B, y, \lambda)$ is continuous in θ, $\mathring{P}$-almost surely, and (5.7) implies the uniformly integrability of $\{\mathscr{L}(\theta,\ \xi,\ B,\ y,\ \lambda),\ \theta \in [0,\ T]\}$. Hence $\|\Lambda(\theta, y, \lambda, \nu)\| = \|\nu\|\,\mathring{E}\mathscr{L}(\theta, \xi, B, y, \lambda)$ is continuous in θ.

Define a metric Δ on $M(R^n)$ (= totality of positive measure on R^n) as follows

(5.17) $$\Delta(\mu, \nu) = L\Big(\frac{\mu}{\|\mu\|},\ \frac{\nu}{\|\nu\|}\Big) + |\|\mu\| - \|\nu\|| + \Big|\frac{1}{\|\mu\|} - \frac{1}{\|\nu\|}\Big|\,.$$

Then $(M(R^n), \Delta)$ is a complete separable metric space and ν_k, $k = 1, 2, \cdots$ is a Cauchy sequence, iff $\langle f, \nu_k\rangle$ converges for any $f \in C_b(R^n)$, as $k \to \infty$. and $\lim_{k\to\infty}\langle 1, \nu_k\rangle(=\|\nu_k\|) > 0$. Recalling Prohorov's theorem we have

PROPOSITION 5.1. *$N \supset M(R^n)$ is Δ-totally bounded, iff there exist positive constants c and c' and for $^\forall\varepsilon > 0$ there is a compact subset $K_\varepsilon \subset R^n$ such that*

(5.18) $$c' \leq \|\nu\| \leq c \quad \textit{and} \quad \nu(K_\varepsilon^c) < \varepsilon \quad \textit{for } \nu \in N.$$

Put m = totality of $M(R^n)$-valued random variables, which may be defined on different probability spaces. We endow the Prohorov metric on m, (called δ metric), namely

$$\delta(\zeta_1, \zeta_2) = L(\mu_1, \mu_2)$$

where μ_i is the probability distribution of ζ_i. Then (m, δ) is a complete separable metric space, because $(M(R^n), \Delta)$ is a complete separable metric space.

Concerning the continuity of Λ, we can prove the following theorem.

THEOREM 5. *If $y_k \to y$ in $C([0\ T] \to R^1)$, $\nu_k \to \nu$ in Δ, $\lambda_k \to \lambda$ in $\hat{d}_T$ and $\theta_k \to \theta$, then*

(5.19) $$\Lambda(\theta_k, y_k, \lambda_k, \nu_k) \longrightarrow \Lambda(\theta, y, \lambda, \nu) \quad \textit{in } \Delta.$$

Proof. Firstly we remark that

(5.20) $$\int_0^s R(t, x, \lambda_k)dt \longrightarrow \int_0^s R(t, x, \lambda)\, dt.$$

Recalling the definition of R for $\lambda \in \hat{M}([0\ T] \times \Gamma)$ (see (5.3)), we get, for any x,

(5.21) $$R(t, x, \lambda_k) = \lim_{\lambda \to \infty} \int_\Gamma \gamma(x, u)(\lambda_k(t, du) - \lambda_k(t - 2^{-n}, du))2^n \qquad \text{for a.a. } t.$$

Hence the bounded convergence theorem implies

(5.22) $$\begin{aligned} \int_0^s R(t, x, \lambda_k)dt &= \lim_{n\to\infty} \left[\int_0^s \int_\Gamma 2^n \gamma(x, u)\lambda_k(t, du)dt \right. \\ &\qquad \left. - \int_{2^n}^s \int_\Gamma 2^n \gamma(x, u)\lambda_k(t - 2^{-n}, du)dt \right] \\ &= \int_\Gamma \gamma(x, u)\lambda_k(s, du). \end{aligned}$$

Since $\lambda_k \to \lambda$ in $\hat{d}_T$, (5.22) means (5.20).

Consequently we can easily see

LEMMA. *If $\phi_k \to \phi$ and $\psi_k \to \psi$ in $C([0\ T] \to R^n)$ and $C([0\ T] \to R^1)$ respectively, then*

(5.23) $$\int_0^s R(t, \phi_k(t), \lambda_k)\psi_k(t)dt \longrightarrow \int_0^s R(t, \phi(t), \lambda)\psi(t)dt.$$

Putting $\Lambda_k = \Lambda(\theta_k, y, \lambda_k, \nu_k)$ and $\Lambda = \Lambda(\theta, y, \lambda, \nu)$, we will show

$$\langle f, \Lambda_k \rangle \longrightarrow \langle f, \Lambda \rangle \qquad \text{for } f \in C_b(R^n). \tag{5.24}$$

Consider the SDE on $(\Omega_k, F_k, \mathring{P}_k)$

$$\begin{cases} d\eta_k(t) = \alpha(\eta_k(t))dB_k(t) + \int_\Gamma \gamma(\eta_k(t), u)\lambda_k(dt, du) \\ \eta_k(0) = \xi_k \end{cases} \tag{5.25}$$

and on $(\Omega, F, \mathring{P})$

$$\begin{cases} d\eta(t) = \alpha(\eta(t))dB(t) + \int_\Gamma \gamma(\eta(t), u)\lambda(dt, du) \\ \eta(0) = \xi \end{cases} \tag{5.26}$$

where ξ_k and ξ have probability distributions $\nu_k/\|\nu_k\|$ and $\nu/\|\nu\|$ respectively. Since $\{(\eta_k, \xi_k, B_k), k = 1, 2, \cdots\}$ is totally bounded in Prohorov topology and any convergent subsequence tends to (η, ξ, B) in Prohorov topology, (η_k, ξ_k, B_k), $k = 1, 2, \cdots$ itself converges to (η, ξ, B) in Prohorov topology. Appealing to Skorobod's theorem, we will assume that $\Omega_k = \Omega$, $F_k = F$, $\mathring{P}_k = \mathring{P}$ and $\mathring{P}$-almost surely $\eta_k \to \eta$ in $C([0\ T] \to R^n)$, $B_k \to B$ in $C([0\ T] \to R^n)$ and $\xi_k \to \xi$ in R^n. Therefore the lemma guarantees

$$\int_0^{\theta_k} y_k(s)A(s, \lambda_k)h(\eta_k(s))ds \longrightarrow \int_0^\theta y(s)A(s, \lambda)h(\eta(s))ds, \quad \mathring{P}\text{-almost surely}. \tag{5.27}$$

Furthermore, using a routine method we have

$$\begin{aligned} &\int_0^{\theta_k} y_k(s)\frac{\partial h}{\partial x_i}(\eta_k(s))\alpha_{ij}(\eta_k(s))dB_{k,j}(s) \\ &\qquad \longrightarrow \int_0^\theta y(s)\frac{\partial h}{\partial x_i}(\eta(s))\alpha_{ij}(\eta(s))dB_j(s) \qquad \text{in proba } \mathring{P}. \end{aligned} \tag{5.28}$$

Hence we have

$$\mathscr{L}(\theta_k, \xi_k, B_k, y_k, \lambda_k) \longrightarrow \mathscr{L}(\theta, \xi, B, y, \lambda) \qquad \text{in proba } \mathring{P}. \tag{5.29}$$

Since (5.7) means the uniformly integrability of $\{f(\eta_k(\theta_k)\mathscr{L}(\theta_k. \xi_k, B_k, y_k, \lambda_k),$ $k = 1, 2, \cdots\}$, (5.29) implies (5.24).

By virtue of Proposition 5.1, (5.16) and (5.20) guarantee the totally boundedness of $\{\Lambda_k, k = 1, 2, \cdots\}$. Consequently, again (5.24) tells us that Λ_k converges to Λ in metric Δ. This completes the proof of Theorem 5.

Now we apply this theorem to admissible systems.

THEOREM 6. *Let* $\mathscr{A}_k = (\Omega_k, F_k, \mathring{P}_k, \xi_k, B_k, Y_k, q_k)$ *and* $\mathscr{A} = (\Omega, F, \mathring{P}, \xi, B, Y, q)$ *where* ξ_k *and* ξ *have probability distributions* $\nu_k/\|\nu_k\|$ *and* $\nu/\|\nu\|$ *respectively. If* $(Y_k, q_k) \to (Y, q)$ *in Prohorov topology,* $\nu_k \to \nu$ *in* Δ *and* $\theta_k \to \theta$, *then*

$$\Lambda(\theta_k, Y_k, q_k, \nu_k) \longrightarrow \Lambda(\theta, Y, q, \nu) \qquad \textit{in metric } \delta. \tag{5.30}$$

Proof. By the assumption $\mathscr{A}_k \to \mathscr{A}$ in Prohorov topology. Hence $(X_k, \mathscr{A}_k) \to (X, \mathscr{A})$ in Prohorov metric D_T, by virtue of Corollary of Proposition 3.1. By Skorobod's theorem we can construct copies $(X_k^*, \mathscr{A}_k^*)$ and $(X^*, \mathscr{A}^*)$ of $(X_k, \mathscr{A}_k)$ and $(X, \mathscr{A})$ respectively, so that $\Omega_k^* = \Omega^*$, $F_k^* = F^*$, $\mathring{P}_k^* = \mathring{P}^*$ and $\mathring{P}^*$-almost surely $(X_k^*, \xi_k^*, B_k^*, Y_k^*, q_k^*) \to (X^*, \xi^*, B^*, Y^*, q^*)$ in d_T. For non-exceptional $\omega \in \Omega^*$, we put $\lambda_k = q_k^*(\cdot, \omega)$, $y_k = Y_k^*(\cdot, \omega)$, $\lambda = q^*(\ , \omega)$ and $y = Y^*(\ , \omega)$. Then y_k, λ_k, y and λ satisfy the condition of Theorem 5, Theorefore we have

$$\Lambda(\theta_k, Y_k^*, q_k^*, \nu_k) \longrightarrow \Lambda(\theta, Y^*, q^*, \nu), \qquad \mathring{P}^*\text{-almost surely}. \tag{5.31}$$

On the other hand Theorem 5 tells us that the mapping $\Lambda(\theta, \cdot, \cdot, \nu)$; $C([0\ T] \to R^1) \times (\hat{M}([0\ T] \times \Gamma)) \to M(R^n)$ is continuous. So $\Lambda(\theta, Y, q, \nu)$ is a random variable, i.e. $\Lambda(\theta, Y, q, \nu) \in m$. Consequently (5.31) implies (5.30).

Recalling Corollary of Theorem 4, we get

COROLLARY. *For any* $\mathscr{A} \in \mathfrak{A}(\mu)$, *there exists an approximate admissible switching system* $\mathscr{A}_k \in \mathfrak{A}(\mu)$, *such that*

$$\Lambda(\theta_k, Y_k, q_k, \nu) \longrightarrow \Lambda(\theta, Y, q, \nu) \qquad \textit{in metric } \delta \tag{5.32}$$

where $\mu = \nu/\|\nu\|$.

§ 6. Semigroup

Let C be the Banach lattice of the totality of bounded continuous mappings from $(M(R^n), \Delta)$ into R^1, with supremum norm and the order $\leq$, i.e.

$$\phi \leq \psi \Longleftrightarrow \phi(\nu) \leq \psi(\nu) \qquad \text{for } {}^\forall\nu \in M(R^n). \tag{6.1}$$

For $\nu \in M(R^n)$, $\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|)$ and $\phi \in C$ we define J by

$$J(t, \mathscr{A}, \nu, \phi) = \mathring{E}(\phi(\Lambda(t, Y, q, \nu)) \tag{6.2}$$

$$\begin{aligned} &\mathring{E}[\mathring{E}(\phi(\Lambda(t, Y, q, \nu)/\sigma(Y, q)] \\ &= \int_{C([0\ T] \to R^1) \times \hat{M}([0\ T] \times \Gamma)} \mathring{E}(\phi(\Lambda(t, y, \lambda, \nu))\pi(dy, d\lambda) \end{aligned} \tag{6.3}$$

where π is the probability distribution of (Y, q). Since $\mathring{E}(\phi(\Lambda(t, y, \lambda, \nu))$ depends only on ϕ, ν, t, y and λ, $J(t, \mathscr{A}, \nu, \phi)$ can be denoted by $J(t, \pi, \nu, \phi)$.

Define $S(t)\phi$ by

$$S(t)\phi(\nu) = \sup_{\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|)} J(t, \mathscr{A}, \nu, \phi) . \tag{6.4}$$

Then by Proposition 2.3 (ii) and (6.3) we have

$$S(t)\phi(\nu) = \sup_{\pi \in \mathscr{P}} J(t, \pi, \nu, \phi) . \tag{6.5}$$

PROPOSITION 6.1. *$J(t, \pi, \nu, \phi)$ is continuous in $(t, \pi, \nu) \in [0\ T] \times \mathscr{P} \times M(R^n)$.*

Proof. Let $t_k \to t$, $\pi_k \to \pi$ and $\nu_k \to \nu$ in their topologies. Take $\mathscr{A}_k \in \mathfrak{A}(\nu_k/\|\nu_k\|)$ (and $\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|)$) such that the probability distribution of (y_k, q_k) (and (Y, q)) is π_k (and π respectively). Then $\mathscr{A}_k \to \mathscr{A}$ by Proposition 2.2. Therefore Theorem 6 guarantees that $\Lambda(t_k, \pi_k, \nu_k, \phi) \to \Lambda(t, \pi, \nu, \phi)$ in metric δ that is in the Prohorov topology. By Skorobod's theorem we can take a copy Λ_k^* of $\Lambda(t_k, \pi_k, \nu_k, \phi)$ and Λ^* of $\Lambda(t, \pi, \nu, \phi)$ so that Λ_k^* converges to Λ_k^* almost surely on $(\Omega^*, F^*, \mathring{P}^*)$. Since ϕ is bounded continuous, we see that

$$\begin{cases} J(t_k, \pi_k, \nu_k, \phi) = \mathring{E}_k\phi(\Lambda(t_k, Y_k, q_k, \nu_k) = \mathring{E}^*\phi(\Lambda_k^*) \\ J(t, \pi, \nu, \phi) = \mathring{E}\phi(\Lambda(t, Y, q, \nu) = \mathring{E}^*\phi(\Lambda^*) \\ \mathring{E}^*\phi(\Lambda_k^*) \to \mathring{E}^*\phi(\Lambda^*) . \end{cases} \tag{6.6}$$

This completes the proof of Proposition 6.1.

Since $\mathscr{P}$ is a compact metric space by Proposition 2.3 (i), we can conclude the following proposition.

PROPOSITION 6.2. *$S(t)\phi \in C$ whenever $\phi \in C$. That is, $S(t)$ is a mapping from C into C. Recalling Corollary of Theorem 6, we see*

$$S(t)\phi(\nu) = \sup_{\substack{\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|) \\ \mathscr{A}:\text{swisching syst.}}} J(t, \mathscr{A}, \nu, \phi) . \tag{6.7}$$

THEOREM 7. $S(t + \theta) = S(t)S(\theta)$, $S(0) =$ identity.

Proof. Consider the SDE on $(\Omega, F, \mathring{P})$, for $\lambda \in \hat{M}([0\ T] \times \Gamma)$

$$\begin{cases} d\eta(t) = \alpha(\eta(t))dB(t) + \int_\Gamma \gamma(\eta(t), u)\lambda(dt, du) \\ \eta(0) = \xi . \end{cases} \tag{6.8}$$

Since a solution $\eta(\cdot, \xi, B, \lambda)$ is unique, η satisfies the following relation

$$\eta(t+\theta, \xi, B, \lambda) = \eta(\theta, \eta(t, \xi, B, \lambda), B_t^+, \lambda_t^+) \tag{6.9}$$

where $B_t^+(s) = B(t+s) - B(t)$, $\lambda_t^+(s, A) = \lambda(t+s, A) - \lambda(t, A)$.

Using Ito's formula we get

$$\mathscr{L}(t+s, \xi, B, y, \lambda) = \mathscr{L}(t, \xi, B, y, \lambda)\mathscr{L}(s, \eta(t, \xi, B, \lambda), B_t^+, y_t^+, \lambda_t^+). \tag{6.10}$$

Define v: $[0\ \ T] \times R^n \times C([0\ \ T] \to R^1) \times \hat{M}([0\ \ T] \times \Gamma) \times C_b(R^n) \to R^1$ by

$$v(t, x, y, \lambda, f) = \mathring{E}f(\eta(t, x, B, \lambda))\mathscr{L}(t, x, B, y, \lambda) \tag{6.11}$$

where $\mathring{E}$ of the right side stands for the expectation with respect to B, since the starting point x is not random. From (6.10) and (6.11) we have

$$\begin{aligned} v(t+s, x, y, \lambda, f) &= \mathring{E}f(\eta(t+s, x, B, \lambda))\mathscr{L}(t+s, x, B, y, \lambda) \\ &= \mathring{E}[\mathscr{L}(t, x, B, y, \lambda)\mathring{E}(f(\eta(s, \eta(t, x, B, \lambda), B_t^+, \lambda_t^+) \\ &\qquad \times \mathscr{L}(s, \eta(t, x, B, \lambda), B_t^+, y_t^+, \lambda_t^+/\sigma_t(B))] . \end{aligned} \tag{6.12}$$

Since $\eta(t, x, B, \lambda)$ is $\sigma_t(B)$-measurable, we see

$$\begin{aligned} &\mathring{E}(f(\eta(s, \eta(t, x, B, \lambda), B_t^+, \lambda_t^+)\mathscr{L}(s, \eta(t, x, B, \lambda), B_t^+, y_t^+, \lambda_t^+)/\sigma_t(B))) \\ &\quad = v(s, \eta(t, x, B, \lambda), y_t^+, \lambda_t^+, f) \end{aligned} \tag{6.13}$$

and, combining with (6.12) we get

$$\begin{aligned} v(t+s, x, y, \lambda, f) &= \mathring{E}v(s, \eta(t, x, B, \lambda), y_t^+, \lambda_t^+, f)\mathscr{L}(t, x, B, y, \lambda) \\ &= v(t, x, y, \lambda, v(s, \cdot, y_t^+, \lambda_t^+, f)) . \end{aligned} \tag{6.14}$$

Recalling (5.8) and (5.15) we get

$$\langle f, \Lambda(t, y, \lambda, \nu)\rangle = \langle v(t, \cdot, y, \lambda), \nu\rangle , \qquad f \in C_b(R^n). \tag{6.15}$$

Hence, by (6.14), we have

$$\begin{aligned} \langle f, \Lambda(t+s, y, \lambda, \nu)\rangle &= \langle v(t+s, \cdot, y, \lambda), \nu\rangle \\ &= \langle v(t, \cdot, y, \lambda, v(s, \cdot, y_t^+, \lambda_t^+, f)), \nu\rangle \\ &= \langle v(s, \cdot, y_t^+, \lambda_t^+, f), \Lambda(t, y, \lambda, \nu)\rangle \\ &= \langle f, \Lambda(s, y_t^+, \lambda_t^+, \Lambda(t, y, \lambda, \nu))\rangle , \qquad f \in C_b(R^n). \end{aligned} \tag{6.16}$$

Consequently

$$\Lambda(t+s, y, \lambda, \nu) = \Lambda(s, y_t^+, \lambda_t^+, \Lambda(t, y, \lambda, \nu)). \tag{6.17}$$

Since (6.17) holds for any $y \in C([0\ \ T] \to R^1)$ and $\lambda \in \hat{M}(0\ \ T] \times \Gamma)$, we have, for any $\mathscr{A} = (\Omega, F, \mathring{P}, \xi, B, Y, q) \in \mathfrak{A}(\nu/\|\nu\|)$,

$$(6.18)\qquad \Lambda(t+s, Y, q, \nu) = \Lambda(s, Y_t^+, q_t^+, \Lambda(t, Y, q, \nu)),\ \mathring{P}\text{-almost surely}.$$

This implies

$$(6.19)\qquad \begin{aligned} J(t+s, \mathscr{A}, \nu, \phi) &= \mathring{E}\phi(\Lambda(t+s, Y, q, \nu)) \\ &= \mathring{E}[\mathring{E}(\phi(\Lambda(t+s, Y, q, \nu)/\sigma_T(Y\ q)))] \\ &= \mathring{E}\phi(\Lambda(s, Y_t^+, q_t^+, \Lambda(t, Y, q, \nu))) \\ &= \mathring{E}\mathring{E}(\phi(\Lambda(s, Y_t^+, q_t^+, \Lambda(t, Y, q, \nu))/\sigma_t(Y, q))) . \end{aligned}$$

Under the regular conditional probability $\mathring{P}(\ /\sigma_t(Y, q))$, (B_t^+, Y_t^+) is a $(n+1)$-dimensional Brownian motion, (Y_t^+, q_t^+) independent of (B, ξ) and Y_t^+ is $\sigma_s(Y_t^+, q_t^+)$-Brownian motion, i.e. independent increments. Hence $\eta(t, \xi, B, q)$, B_t^+ and (Y_t^+, q_t^+) are independent under conditional probability $\mathring{P}(\cdot/\sigma_t(Y, q))$, $\mathring{P}$-almost surely, although the probability distribution of q_t^+ might depend on the past value of $(Y(\theta), q(\theta, A), \theta \leq t, A \in \boldsymbol{B}(\Gamma))$. Hence there exists a null set $N \in \sigma_t(Y, q)$, such that for $\omega \notin N$, $(\Omega, F, \mathring{P}(\cdot/\sigma_t(Y, q))(\omega), \eta(t, \xi, B, q), B_t^+, Y_t^+, q_t^+) \in \mathfrak{A}$. Therefore

$$(6.20)\qquad \mathring{E}(\phi(\Lambda(s, Y_t^+, q_t^+, \Lambda(t, Y, q, \nu))/\sigma_t(Y, q))) \leq (S(s)\phi)(\Lambda(t, Y, q, \nu)),\quad \mathring{P}\text{-almost surely}.$$

Combining (6.20) with (6.19), we have

$$(6.21)\qquad J(t+s, \mathscr{A}, \nu, \phi) \leq S(t)(S(s)\phi)(\nu).$$

Taking the supremum with respect to $\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|)$, we have

$$(6.22)\qquad S(t+s)\phi(\nu) \leq S(t)(S(s)\phi)(\nu).$$

For the converse inequality we will show some lemmas

LEMMA 1. *Let $N \subset (M(R^n), \Delta)$ be totally bounded. Then $\{\Lambda(t, Y_{\mathscr{A}}, q_{\mathscr{A}}, \nu);\ \mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|), \nu \in N\}$ is totally bounded in (m, δ).*

Proof. Consider the SDE, for $\mathscr{A} = (\Omega, F, \mathring{P}, \xi, B, Y, q) \in \mathfrak{A}(\nu/\|\nu\|)$.

$$(6.23)\qquad \begin{cases} d\eta(t) = \alpha(\eta(t))dB(t) + \int_\Gamma \gamma(\eta(t), u)\lambda(dt, du) \\ \eta(0) = \xi . \end{cases}$$

Then, using this unique solution $\eta(t) = \eta(t, \xi, B, \lambda)$ we have

$$(6.24)\qquad \Lambda(t, y, \lambda, \nu)(A) = \|\nu\| \mathring{E}\chi_A(\eta(t))\mathscr{L}(t, \xi, B, y, \lambda).$$

Hence, by (5.7)

$$\begin{aligned}\Lambda^2(t, y, \lambda, \nu)(A) &\leq \|\nu\|^2 \mathring{P}(\eta(t) \in A)\mathring{E}\mathscr{L}^2(t, \xi, B, y, \lambda) \\ &\leq \|\nu\|^2 \mathring{P}(\eta(t) \in A) \exp[2K_2(t+1)(\|y\|_t + 1)^2]\end{aligned} \tag{6.25}$$

where $\|y\|_t = \sup_{s\leq t}|y(s)|$. This means

$$\Lambda(t, y, \lambda, \nu)(A) \leq \|\nu\| \sqrt{\mathring{P}(\eta(t) \in A)} \exp[K_2(t+1)(1+a)^2], \quad \text{whenever } \|y\|_t \leq a. \tag{6.26}$$

On the other hand the condition (A1) implies that, for $\varepsilon' > 0$, there exists $b = b(\varepsilon', t, N)$ such that

$$\mathring{P}(|\eta(t, \xi, B, \lambda)| > b) < \varepsilon' \quad \text{for } {}^\forall\lambda \in \hat{M}([0\ T] \times \Gamma),\ \mathscr{A} \in \mathfrak{A}\left(\frac{\nu}{\|\nu\|}\right),\ \nu \in N. \tag{6 27}$$

Since Y is a Brownian motion, for $\varepsilon > 0$ there exists $a = a(\varepsilon)$ such that

$$\mathring{P}(\sup_{s\leq t}|Y(s)| \leq a) > 1 - \varepsilon \qquad \text{for } {}^\forall\mathscr{A} \in \mathfrak{A}. \tag{6.28}$$

Putting $\varepsilon' = \varepsilon^2 e^{-2K_2(1+t)(1+a(\varepsilon))^2}$, (6.26) gives

$$\Lambda(t, y, \lambda, \nu)(K^c) < \varepsilon\|\nu\| \qquad \text{for } \lambda \in \hat{M}([0\ T] \times \Gamma) \tag{6.29}$$

whenever $\|y\|_t < a$, where the compact set K is given by

$$K = \{x \in R^n : |x| \leq b(\varepsilon', t, N)\}. \tag{6.30}$$

Therefore combining (6.28) and (6.29), we see, for $\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|)$,

$$1 - \varepsilon < \mathring{P}(\sup_{s\leq t}|Y(s)| \leq a(\varepsilon)) \leq \mathring{P}(\Lambda(t, Y, q, \nu)(K^c) < \varepsilon\|\nu\|). \tag{6.31}$$

Recalling the condition "$0 < c' \leq \|\nu\| \leq c$ for ${}^\forall\nu \in N$", (6.31) implies Lemma 1 by virtue of Proposition 5.1.

Applying Prohorov's theorem, Lemma 1 gives

LEMMA 2. *For $\varepsilon > 0$ and a totally bounded set $N \subset (M(R^n), \Delta)$ there exists a compact set $\tilde{N} = \tilde{N}(\varepsilon, t, N) \subset (M(R^n), \Delta)$ such that*

$$\mathring{P}(\Lambda(t, Y, q, \nu) \in \tilde{N}) > 1 - \varepsilon \qquad \text{for } \mathscr{A} \in \mathfrak{A}\left(\frac{\nu}{\|\nu\|}\right),\ \nu \in N. \tag{6.32}$$

LEMMA 3. *Suppose that $M(R^n) = M_0 \cup \cdots \cup M_\ell$ is a Borel partition of $M(R^n)$. Let $\nu_i \in M_i$ and $\mathscr{A} = (\Omega_i, F_i, \mathring{P}_i, \xi_i, B_i, Y_i, q_i) \in \mathfrak{A}(\nu_i/\|\nu_i\|)$. For any fixed $\mathscr{A} = (\Omega, F, P, \xi, B, Y, q) \in \mathfrak{A}(\nu/\|\nu\|)$ we define $\tilde{\Omega}, \tilde{F}, \tilde{\mathring{P}}, \tilde{\xi}, \tilde{B}, \tilde{Y}, \tilde{q}$ as follows.*

$$\tilde{\Omega} = \Omega \times \Omega_0 \times \Omega_1 \times \cdots \times \Omega_\ell , \qquad \tilde{F} = F \times F_0 \times \cdots \times F_\ell$$
$$\mathring{\tilde{P}} = \mathring{P} \times \mathring{P}_1 \times \cdots \times \mathring{P}_\ell , \qquad \tilde{\xi} = \xi$$
$$\tilde{B}(\theta) = \begin{cases} B(\theta) , & \theta \le t \\ B(t) + \sum_{i=0}^{\ell} B_i(\theta - t)\chi_{M_i}(\Lambda(t, Y, q, \nu)) , & \theta \ge t . \end{cases}$$

$\tilde{Y}$ is defined in the same way.

$$\tilde{q}(\theta, A) = \begin{cases} q(\theta, A) , & \theta \le t \\ q(t, A) + \sum_{i=0}^{\ell} q_i(\theta_i - t, A)\chi_{M_i}(\Lambda(t, Y, q, \nu)) , & \theta \ge t . \end{cases}$$

Then $\tilde{\mathscr{A}} = (\tilde{\Omega}, \tilde{F}, \mathring{\tilde{P}}, \tilde{\xi}, \tilde{B}, \tilde{Y}, \tilde{q}) \in \mathfrak{A}(\nu/\|\nu\|)$.

Proof. ξ is independent of $\{(B, Y, q), (\xi_i, B_i, Y_i, q_i), i = 0, \cdots, \ell\}$. So $\tilde{\xi}$ is independent of $(\tilde{B}, \tilde{Y}, \tilde{q})$.

$\tilde{B}$ is a Brownian motion, because for $g \in C_b((R^n)^k)$ and $\theta_j \ge t$, $j = 1, \cdots, k$,

$$\begin{aligned} &\mathring{\tilde{E}}(g(\tilde{B}_j(\theta_j) - \tilde{B}(t), j = 1, \cdots, k)/\sigma_t(B, Y, q) \vee \sigma(Y_i, q_i), i = 0, \cdots, \ell) \\ &\qquad = \mathring{E}_i(g(B_i(\theta_j - t), j = 1, \cdots, k) \qquad \text{if } \Lambda(t, Y, q, \nu) \in M_i . \end{aligned} \tag{6.33}$$

Hence $(\tilde{B}(s) - \tilde{B}(t), s \ge t)$ is a Brownian motion which is independent of $\sigma_t(B, Y, q) \vee \sigma(Y_0, q_0) \vee \cdots \vee \sigma(Y_\ell, q_\ell)$, since B_i is a Brownian motion. This implies that $(\tilde{B}(s) - \tilde{B}(t), s \ge t, B(\theta), \theta \le t)$ is independent of $((Y, q), (Y_i, q_i), i = 0, \cdots, \ell)$, since B is independent of $((Y, q), (Y_i, q_i), i = 0, \cdots, \ell)$. Therefore $\tilde{B}$ is independent of $(\tilde{Y}, \tilde{q})$, because $(\tilde{Y}, \tilde{q})$ is measurable with respect to $\sigma(Y, q, Y_0, q_0, \cdots, Y_\ell, q_\ell)$.

Using a similar calculation as (6.33), we see that for $g \in C_b\ (R^k)$ and $\theta_j \ge \theta \ge t, j = 1, \cdots, k$

$$\begin{aligned} &\mathring{\tilde{E}}(g(\tilde{Y}(\theta_j) - \tilde{Y}(\theta), j = 1, \cdots, k/\sigma_t(Y, q) \vee \sigma_{\theta - t}(Y_0, q_0, \cdots, Y_\ell, q_\ell))) \\ &\qquad = \mathring{E}_i(g(Y_i(\theta_j) - Y_i(\theta), j = 1, \cdots, k)) , \qquad \text{if } \Lambda(t, Y, q, \nu) \in M_i . \end{aligned} \tag{6.34}$$

Therefore $(\tilde{Y}(s) - \tilde{Y}(\theta), s \ge \theta)$ is a Brownian motion which is independent of $\sigma_\theta(\tilde{Y}, \tilde{q})$.

It is clear that $\tilde{q}$ satisfies the conditions (v) and (vi), from the definition of $\tilde{q}$. This completes the proof of Lemma 3.

Now we prove the inequality (6.35) for Theorem.

$$S(t + s)\phi(\nu) \ge S(t)(S(s)\phi)(\nu) . \tag{6.35}$$

Since $J(s, \pi, \nu, \phi)$ is continuous in π, ν and $\mathscr{P}$ is compact, $J(s, \pi, \nu, \phi)$ is continuous in ν uniformly in $\pi \in \mathscr{P}$, namely for $\varepsilon > 0$ there exists $\delta(\varepsilon, \nu) = \delta(\varepsilon, \nu, s, \phi)$ such that, if $\Delta(\nu', \nu) > \delta(\varepsilon, \nu)$ then

$$|J(s, \pi, \nu', \phi) - J(s, \pi, \nu, \phi)| < \varepsilon . \tag{6.36}$$

Hence

$$|S(s)\phi(\nu') - S(s)\phi(\nu)| < \varepsilon , \qquad \text{if } \Delta(\nu', \nu) < \delta(\varepsilon, \nu) . \tag{6.37}$$

Applying Lemma 2 for $N = \{\nu\}$, given $\varepsilon > 0$ we can take a compact set $\tilde{N}$ as in (6.32). From (6.36) and (6.37) we can take a Borel partition of $\tilde{N}$, say $\tilde{N} = M_1 \cup \cdots \cup M_\ell$, such that, if $\nu', \nu'' \in M_i$, then

$$|J(s, \pi, \nu', \phi) - J(s, \pi, \nu'', \phi)| < \varepsilon \tag{6.38}$$

and

$$|S(s)\phi(\nu') - S(s)\phi(\nu'')| < \varepsilon . \tag{6.39}$$

Fix $\nu_i \in M_i$, $i = 1, \cdots, \ell$, arbitrarily and take $\mathscr{A}_i \in \mathfrak{A}(\nu_i/\|\nu_i\|)$ such that

$$J(s, \pi_i, \nu_i, \phi) \geq S(s)\phi(\nu_i) - \varepsilon \tag{6.40}$$

where $\pi_i =$ probability distribution of $(Y_{\mathscr{A}i}, q_{\mathscr{A}i})$. Then, by (6.38) $\sim$ (6.40), we see

$$J(s, \pi_i, \nu', \phi) \geq J(s, \pi_i, \nu_i, \phi) - \varepsilon \geq S(s)\phi(\nu_i) - 2\varepsilon \geq S(s)\phi(\nu') - 3\varepsilon \qquad \text{for } \nu' \in M_i . \tag{6.41}$$

Let $\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|)$ and $M_0 = \tilde{N}^c$ and take $\nu_0 \in M_0(R^n)$ and $\mathscr{A}_0 \in \mathfrak{A}(\nu_0/\|\nu_0\|)$ arbitrarily. Then M_i, $i = 0, \cdots, \ell$ is a Borel partition of $M(R^n)$. According to Lemma 3 we have $(\tilde{\Omega}, \tilde{F}, \mathring{\tilde{P}}, \tilde{\xi}, \tilde{B}, \tilde{Y}, \tilde{q}) \in \mathfrak{A}(\nu/\|\nu\|)$. Then

$$\begin{aligned} J(t+s, \tilde{\mathscr{A}}, \nu, \phi) &= \mathring{\tilde{E}}\phi(\Lambda(t+s, \tilde{Y}, \tilde{q}, \nu)) \\ &= \mathring{\tilde{E}}\phi(\Lambda(s, \tilde{Y}_t^+, \tilde{q}_t^+, \Lambda(t, Y, q, \nu)) \\ &= \mathring{\tilde{E}}[\mathring{\tilde{E}}(\phi(\Lambda(s, \tilde{Y}_t^+, \tilde{q}_t^+, \Lambda(t, Y, q, \nu))/\sigma_t(Y, q))]. \end{aligned} \tag{6.42}$$

On the other hand, by (6.41) we have

$$\begin{aligned} &\mathring{\tilde{E}}(\phi(\Lambda(s, \tilde{Y}_t^+, \tilde{q}_t^+, \Lambda(t, Y, q, \nu))/\sigma_t(Y, q))) \\ &\quad = \sum_{i=0}^{\ell} [\mathring{E}_i\phi(\Lambda(s, Y_i, q_i, \Lambda(t, Y, q, \nu))]\chi_{M_i}(\Lambda(t, Y, q, \nu)) \\ &\quad = \sum_{i=0}^{\ell} J(s, \mathscr{A}_i, \Lambda(t, Y, q, \nu))\chi_{M_i}(\Lambda(t, Y, q, \nu)) \end{aligned}$$

$$\geq \sum_{i=0}^{\ell} S(s)\phi(\Lambda(t, Y, q, \nu))\chi_{M_i}(\Lambda(t, Y, q, \nu)) - 3\varepsilon - \|\phi\| \chi_{M_0}(\Lambda(t, Y, q, \nu))$$
$$= S(s)\phi(\Lambda(t, Y, q, \nu)) - 3\varepsilon - \|\phi\| \chi_{M_0}(\Lambda(t, Y, q, \nu)) .$$

Combining with (6.42) we see

$$\text{(6.43)} \quad \begin{aligned} J(t+s, \tilde{\mathscr{A}}, \nu, \phi) &\geq \mathring{E}(S(s)\phi)(\Lambda(t, Y, q, \nu)) - 3\varepsilon - \|\phi\| \mathring{P}(\Lambda(t, Y, q, \nu) \notin \tilde{N}) \\ &\geq \mathring{E}(S(s)\phi)(\Lambda(t, Y, q, \nu)) - 3\varepsilon - \|\phi\| \varepsilon . \end{aligned}$$

Since $\tilde{\mathscr{A}} \in \mathfrak{A}(\nu/\|\nu\|)$ we see

$$\text{(6.44)} \qquad S(t+s)\phi(\nu) \geq \mathring{E}(S(s)\phi)(\Lambda(t, Y, q, \nu)) - \varepsilon(3 + \|\phi\|) .$$

Taking the supremum with respect to $\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|)$, we conclude

$$\text{(6.45)} \qquad S(t+s)\phi(\nu) \geq S(t)(S(s)\phi)(\nu) - \varepsilon(3 + \|\phi\|) .$$

Tending $\varepsilon \downarrow 0$, we get our desired inequality (6.35). This completes the proof of Theorem 7.

§ 7. Generator and properties of $S(t)$.

We can easily see:

PROPOSITION 7.1. *The following properties hold,*

(i) *monotone,* $S(t)\phi \leq S(t)\psi$ *whenever* $\phi \leq \psi$

(ii) *contraction,* $\|S(t)\phi - S(t)\psi\| \leq \|\phi - \psi\|$

(iii) *continuity,* $S(\theta)\phi(\nu) \to S(t)\phi(\nu)$ *as* $\theta \to t$, *uniformly on any compact set of* $M(R^n)$.

That is, $S(t)$ *is a monotone contraction weakly continuous semigroup on* C.

Proof. (i) From the definition of J, (7.1) is clear

$$\text{(7.1)} \qquad J(t, \pi, \nu, \phi) \leq J(t, \pi, \nu, \psi) , \qquad \text{if } \phi \leq \psi .$$

Hence taking the supremum with respect to $\pi \in \mathscr{P}$, we have (i).

(ii) $|S(t)\phi(\nu) - S(t)\psi(\nu)|$
$$\leq \sup_{\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|)} |\mathring{E}\phi(\Lambda(t, Y_{\mathscr{A}}, q_{\mathscr{A}}, \nu)) - \mathring{E}\psi(\Lambda(t, Y_{\mathscr{A}}, q_{\mathscr{A}}, \nu))|$$
$$\leq \|\phi - \psi\| .$$

Hence taking the supremum with respect to $\nu \in M(R^n)$, we have (ii).

(iii) By Proposition 6.1. $J(t, \pi, \nu, \phi)$ is continuous in $(t, \pi, \nu) \in [0\ T] \times \mathscr{P} \times M(R^n)$. Since $\mathscr{P}$ is compact, $S(t)\phi(\nu)$ is continuous in (t, ν). Hence it is uniformly continuous on $[0\ T] \times F$ where F is compact in $M(R^n)$. This implies (iii).

Now we calculate the generator of $S(t)$, according to [6]. We introduce the following set $\mathscr{D}$ of functions ϕ which depend on finitely many scalar products. Fix $H_N \in C^\infty([0\ \infty) \to [0, 1])$ such that $H_N(x) = 1$ on $[0\ N]$, $= 0$ on $[N+1, \infty)$ and decreasing in x.

$$\begin{aligned}\mathscr{D} = \{\phi;\ & M(R^n) \longrightarrow R^1;\ \phi(\nu) = F(\langle f_1, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)H_N(\langle 1, \nu\rangle) \\ & \text{with } F \in C_b^\infty(R^\ell), f_1, \cdots, f_\ell \in C_0^\infty(R^n),\ \ell = 1, 2, \cdots, N = 1, 2, \cdots\}\end{aligned} \tag{7.2}$$

where C_0^∞ denotes the space of C^∞ functions with compact supports and C_b^∞ the space of C^∞ functions with any bounded derivative. Clearly $\mathscr{D} \subset C$. Moreover we have

PROPOSITION 7.2. *For $\phi \in C$ there exists $\Phi_k \in \mathscr{D}$ such that $\Phi_k(\nu) \underset{k\to\infty}{\longrightarrow} \phi(\nu)$ for any $\nu \in M(R^n)$.*

Proof. We can apply the same method as [6]. Let $v(x_i, 2^{-N})$ (= open ball with center x_i, radius 2^{-N}) $i = 1, 2, \cdots$ be a covering of R^n with $\bigcup_{i=1}^{k_N} v(x_i, 2^{-N}) \supset [-N, N]^n$ (say I_N). Let g_ℓ^N, $\ell = 1, 2, \cdots$ be a C^∞-partition of unity such that

$$\operatorname{supp} g_\ell^N \subset v(x_i, 2^{-N}) \text{ for some } i \tag{7.3}$$

$$\sum_{\ell=1}^{p_N} g_\ell = 1 \quad \text{on} \quad I_N\,. \tag{7.4}$$

Take $y_\ell^N \in \operatorname{supp} g_\ell^N \cap I_N$ arbitrarily. Putting $c_\ell^N(\nu) = \langle g_\ell^N, \nu\rangle$, we define ν_N by

$$\nu_N = \sum_{\ell=1}^{p_N} c_\ell^N(\nu)\delta_{y_\ell^N}\,. \tag{7.5}$$

Then $\|\nu\| \geq \|\nu_N\| \geq \nu(I_N)$ and for $g \in C_b(R^n)$

$$\langle g, \nu_N\rangle \to \langle g, \nu\rangle\,, \quad \text{as} \quad N \to \infty\,. \tag{7.6}$$

Hence

$$\Delta(\nu_N, \nu) \to 0\,, \quad \text{as} \quad N \to \infty\,. \tag{7.7}$$

Denote $F_N(z_1, \cdots, z_{p_N} : \phi) = \phi(\sum_{\ell=1}^{p_N} z_\ell \delta_{y_\ell^N})$. That is,

$$\phi(\nu_N) = F_N(\langle g_1^N, \nu\rangle, \cdots, \langle g_{p_N}^N, \nu\rangle : \phi) \tag{7.8}$$

Therefore, by (7.6), we have

$$\phi(\nu_N) \to \phi(\nu) \quad \text{as} \quad N \to \infty\,. \tag{7.9}$$

From the definition of H_N, $\lim_{N\to\infty} H_N(\langle 1, \nu\rangle) = 1$ for any $\nu \in M(R^n)$. Hence

(7.9) gives

$$F_N(\langle g_1^N, \nu\rangle, \cdots, g_{pN}^N, \nu\rangle : \phi)H_N(\langle 1, \nu\rangle) \to \phi(\nu), \quad \text{as} \quad N \to \infty. \tag{7.10}$$

Take $\tilde{F}_N \in C_b^\infty(R^{pN})$ such that

$$\|\tilde{F}_N - F_N\|_{L_\infty(I_N+1)} < 2^{-N}. \tag{7.11}$$

Then we have

$$\begin{aligned}|\tilde{F}_N(\langle g_1^N, \nu\rangle, \cdots, \langle g_{pN}^N, \nu\rangle)H_N(\langle 1, \nu\rangle)\\ - F_N(\langle g_1^N, \nu\rangle, \cdots, \langle g_{pN}^N, \nu\rangle : \phi)H_N(\langle 1, \nu\rangle)| < 2^{-N}.\end{aligned}$$

Combining with (7.10), we complete the proof.

We calculate the generator of $S(t)$, recalling (6.7). For an admissible switching system $\mathscr{A} \in \mathfrak{A}(\nu/\|\nu\|)$, we have the following Zakai equation for $\Lambda(t) = \Lambda(t, Y_{\mathscr{A}}, q_{\mathscr{A}}, \nu)$, (see Theorem 5.2 in [8]),

$$\begin{cases} d\langle f, \Lambda(t)\rangle = \langle A(U(t))f, \Lambda(t)\rangle dt + \langle hf, \Lambda(t)\rangle dY(t) \\ \langle f, \Lambda(0)\rangle = \langle f, \nu\rangle \qquad \text{for } f \in C_b^2(R^n), \end{cases} \tag{7.12}$$

where

$$Y = Y_{\mathscr{A}}, \qquad A(u) = \sum_{ij} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_i \gamma_i(x. y)\frac{\partial}{\partial x_i}$$

and $U(t)$ is a $\sigma_t(q)$-progressively measurable process for q (see Theorem 4).

Therefore using a routine method we have

$$\mathring{E}|\langle f, \Lambda(t)\rangle - \langle f, \nu\rangle|^2 \le K_3(t+1)\int_0^t \mathring{E}\langle 1, \Lambda(\theta)\rangle^2 d\theta \tag{7.13}$$

$$\mathring{E}\langle 1, \Lambda(t)\rangle = \langle 1, \nu\rangle = \|\nu\| \tag{7.14}$$

and

$$\mathring{E}\|\Lambda(t)\|^2 \le \|\nu\|^2 e^{K_4 t} \tag{7.15}$$

where K_3 and K_4 are independent of $\mathscr{A}$. Combining these evaluations, we have

$$\mathring{E}|\langle f, \Lambda(t)\rangle, - \langle f, \nu\rangle|^2 \le K_2\|\nu\|^2(t+1)te^{K_4 t}. \tag{7.16}$$

Let $\Phi \in \mathscr{D}$, say $\Phi(\nu) = F(\langle f_1, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)H_N(\langle 1, \nu\rangle)$. For simplicity we put $f_0 = 1$, and $\tilde{F}(z_0, z_1, \cdots, z_\ell) = F(z_1, \cdots, z_\ell)H_N(z_0)$. Appealing to Ito's formula, we see

$$
\begin{aligned}
d\Phi(\Lambda(t)) &= \sum_{i=0}^{\ell} \frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \Lambda(t)\rangle, \cdots, \langle f_\ell, \Lambda(t)\rangle)\langle A(U(t))f_i, \Lambda(t)\rangle dt \\
&\quad + \frac{1}{2}\sum_{ij=0}^{\ell} \frac{\partial^2 \tilde{F}}{\partial z_i \partial z_j}(\langle f_0, \Lambda(t)\rangle, \cdots, \langle f_\ell, \Lambda(t)\rangle)\langle hf_i, \Lambda(t)\rangle\langle hf_j, \Lambda(t)\rangle dt \qquad (7.17)\\
&\quad + \sum_{i=0}^{\ell} \frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \Lambda(t)\rangle, \cdots, \langle f_\ell, \Lambda(t)\rangle)\langle hf_i, \Lambda(t)\rangle dY(t)\,.
\end{aligned}
$$

Using (7.16) we have, for $s \leqq 1$,

$$
\begin{aligned}
\int_0^s \mathring{E}\Big| &\frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \Lambda(t)\rangle, \cdots, \langle f_\ell, \Lambda(t)\rangle)\langle A(U(t))f_i, \Lambda(t)\rangle \\
&- \frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle A(U(t))f_i, \nu\rangle\Big| dt \leq K_5\|\nu\|s^{3/2} \qquad (7.18)
\end{aligned}
$$

and

$$
\begin{aligned}
\int_0^s \mathring{E}\Big| &\frac{\partial^2 \tilde{F}}{\partial z_i \partial z_j}(\langle f_0, \Lambda(t)\rangle, \cdots, \langle f_\ell, \Lambda(t)\rangle)\langle hf_i, \Lambda(t)\rangle\langle hf_j, (t)\rangle \\
&- \frac{\partial^2 \tilde{F}}{\partial z_i \partial z_j}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle hf_i, \nu\rangle\langle hf_j, \nu\rangle\Big| dt \leq K_5\|\nu\|s^{3/2} \qquad (7.19)
\end{aligned}
$$

with K_5 which is independent of $\mathscr{A}$ and ν.

Define $G\Phi$ as follows

$$
\begin{aligned}
G\Phi(\nu) &= \sup_{u\in\Gamma}\left(\sum_{i=0}^{\ell} \frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle A(u)f_i, \nu\rangle\right) \\
&\quad + \frac{1}{2}\sum_{ij=0}^{\ell} \frac{\partial^2 \tilde{F}}{\partial z_i \partial z_j}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle hf_i, \nu\rangle\langle hf_j, \nu\rangle. \qquad (7.20)
\end{aligned}
$$

Since $\tilde{F}$ is smooth and $\langle A(u)f_i, \nu\rangle$ continuous in u and ν, $G\Phi(\nu)$ is continuous in ν. Moreover $G\Phi(\nu) = 0$ whenever $\|\nu\| \geq N + 1$. Therefore $G\Phi$ is bounded. This implies that $G\Phi \in C$ for $\Phi \in \mathscr{D}$.

We remark that

$$
\begin{aligned}
\sum_i \int_0^s &\mathring{E}\frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle A(U(t)f_i, \nu\rangle dt \\
&\leq \int_0^s \sup_{u\in\Gamma}\left(\sum_{i=0} \frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle A(u)f_i, \nu\rangle\right)dt \\
&= s, \sup_{u\in\Gamma}\left(\sum_{i=0}^{\ell} \frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle A(u)f_i, \nu\rangle\right) \qquad (7.21)\\
&= \sup_{u\in\Gamma}\sum_{i=0}^{\ell}\int_0^s \frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle A(u)f_i, \nu\rangle dt \\
&\leq \sup_{\substack{U;\ \text{usual control for} \\ \text{switch relaxed syst.}}} \sum_{i=0}^{\ell}\int_0^s \mathring{E}\frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle A(U(t))f_i, \nu\rangle dt.
\end{aligned}
$$

Taking the supremum of left side of (7.21) with respect to $U(t)$ for switching admissible systems, we have

$$\text{(7.22)} \quad \frac{1}{s} \sup_{U;\ \text{switch syst.}} \sum_{i=0}^{\ell} \int_0^s \mathring{E} \frac{\partial \tilde{F}}{\partial z_i}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle A(U(t))f_i, \nu\rangle dt = G\Phi(\nu).$$

On the other hand (7.17) ~ (7.19) tell us that

$$\text{(7.23)} \quad \begin{aligned} &\Big| \mathring{E}\Phi(\Lambda(s, Y_{\mathscr{A}}, q_{\mathscr{A}}, \nu)) - \Phi(\nu) - \int_0^s \sum_{i=0}^{\ell} \mathring{E} \frac{\partial \tilde{F}}{\partial z_j}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle A(U(t))f_i, \nu\rangle \\ &\qquad + \frac{1}{2} \sum_{ij=0}^{\ell} \frac{\partial^2 \tilde{F}}{\partial z_i \partial z_j}(\langle f_0, \nu\rangle, \cdots, \langle f_\ell, \nu\rangle)\langle hf_i, \nu\rangle\langle hf_j, \nu\rangle dt \Big| \\ &\quad \leq 2K_5(n+1)^2 \|\nu\| s^{3/2}. \end{aligned}$$

Appealing to (7.22) we have

$$\text{(7.24)} \quad \left| \frac{1}{s}(S(s)\Phi(\nu) - \Phi(\nu)) - G\Phi(\nu) \right| \leq 2K_5(n+1)^2 \|\nu\| s^{1/2}$$

Recalling (5.16) we have

$$\text{(7.25)} \quad \|\Lambda(s, Y_{\mathscr{A}}, q_{\mathscr{A}}, \nu)\| \geq \|\nu\| e^{-K_1(a+1)^2(s+1)} \qquad \text{whenever } \sup_{t \leq s} |Y_{\mathscr{A}}(t)| \leq a.$$

Since $Y_{\mathscr{A}}$ is a Brownian motion, a martingale inequality implies

$$\text{(7.26)} \quad \mathring{P}(\sup_{t \leq s} |Y_{\mathscr{A}}(t)| \geq a) \leq \frac{s}{a^2}.$$

Putting $a = 1/\sqrt{\varepsilon}$ and $\tilde{N}(\varepsilon) = (N+1)e^{K_1(a+1)^2(s+1)}$, we see, from (7.25) and (7.26)

$$\text{(7.27)} \quad \mathring{P}(\|\Lambda(s, Y_{\mathscr{A}}, q_{\mathscr{A}}, \nu)\| \leq N+1) \leq \mathring{P}\left(\sup_{t \leq s} |Y_{\mathscr{A}}(t)| > \frac{1}{\sqrt{\varepsilon}}\right) < \varepsilon s \quad \text{whenever } \|\nu\| \geq \tilde{N}(\varepsilon).$$

Therefore, if $\|\nu\| \geq \tilde{N}(\varepsilon)$, then

$$\text{(7.28)} \quad E|\Phi(\Lambda(s, Y_{\mathscr{A}}, q_{\mathscr{A}}, \nu))| \leq \|\Phi\| \varepsilon s.$$

This implies, by virtue of "$\Phi(\nu) = G\Phi(\nu) = 0$ for $\|\nu\| \geq \tilde{N}(\varepsilon)$",

$$\text{(7.29)} \quad \frac{1}{s}(S(s)\Phi(\nu) - \Phi(\nu)) - G\Phi(\nu)| \leq \|\Phi\| \varepsilon, \qquad \text{whenever } \|\nu\| \geq \tilde{N}(\varepsilon).$$

Appealing to (7.24), we have

(7.30) $$\left|\frac{1}{s}(S(s)\Phi(\nu)-\Phi(\nu))-G\Phi(\nu)\right|\leq(1+\|\Phi\|)\varepsilon,$$
$$\text{whenever } s\leq\varepsilon/(2K_5(n+1)^2\tilde{N}(\varepsilon))^2.$$

This implies that $1/s(S(s)\Phi(\nu)-\Phi(\nu))$ converges to $G\Phi(\nu)$ uniformly on $M(R^n)$, as $s\downarrow 0$.

THEOREM 8. *$\mathscr{D}(\mathfrak{G})\supset\mathscr{D}$ and*

(7.31) $$\mathfrak{G}\Phi=G\Phi \quad \text{on} \quad \mathscr{D}$$

that is, the generator is an extension of G.

§ 8. Time discrete approximation

First we recall an approximation theorem of Proposition 4.1 and Theorem 4, namely, for $\mathscr{A}=(\Omega, F, \mathring{P}, \xi, B, T, q)$ there exists an approximate usual control U_n, such that

(8.1) U_n is an $\sigma_t(q)$-progressively measurable Γ-valued process

(8.2) $q_n(t, A)=\int_0^t \delta_{U_n(s)}(A)ds$ is a switching relaxed control of $\mathscr{A}_n=(\Omega, F, \mathring{P}, \xi, B, Y, q_n)$

and

(8.3) $J(t, \mathscr{A}, \nu, \phi)=\lim_{n\to\infty} J(t, \mathscr{A}_n, \nu, \phi)$ for $^\forall\phi\in C$.

Now we define a usual admissible system $\tilde{\mathscr{A}}=(\Omega, F, \mathring{P}, \xi, B, Y, U)$ as follows: $(\Omega, F, \mathring{P}, \xi, B, Y)$ satisfies the same conditions as an admissible (relaxed) system, U is a Γ-valued process, ξ, B and (Y, U) are independent and Y is a $\sigma_t(Y, U)$-Brownian motion.

$\mathfrak{A}$ denotes the totality of usual admissible systems, and we apply similar notations as for the relaxed case. Putting $\tilde{q}_U(t, A)=\int_0^t \delta_{U(s,\omega)}(A)ds$, we see $\mathscr{A}_U=(\Omega, F, P, \xi, B, Y, \tilde{q}_U)\in\mathfrak{A}$. Thus a usual admissible system can be regarded as an admissible (relaxed) system. Moreover a unique solution $\tilde{X}$ of the SDE,

(8.4) $$\begin{cases} d\tilde{X}(t)=\alpha(\tilde{X}(t))dB(t)+\gamma((t), U(t))dt \\ \tilde{X}(0)=\xi \end{cases}$$

gives a unique solution $X_{\mathscr{A}_U}(=\hat{X})$.

Since (ξ, B) and (Y, U) are independent, we can calculate

$$\mathring{E}(f(\hat{X}(\theta))L(\theta, \tilde{\mathscr{A}})/\sigma_t(Y, U))$$

in the same way as (5.8), and get

$$\begin{aligned}
&\mathring{E}(f(\hat{X}(\theta))L(\theta, \tilde{\mathscr{A}})/Y = y,\ U(t) = v(t) \text{ a.a. } t) \\
(8.5) \qquad &= \mathring{E}f(\eta_v(\theta)\mathscr{L}(\theta, \xi, B, y, q_v) \\
&= \mathring{E}(f(X_{\mathscr{A}_U}(\theta))L(\theta, \mathscr{A}_U)/Y = y,\ q_U = \lambda_v)
\end{aligned}$$

where $\lambda_v(t, A) = \int_0^t \delta_{v(s)}(A)ds$. We remark that, if $v(t) = v^*(t)$, a.a. t, then $\lambda_v = \lambda_{v^*}$. The unnormalized conditional distribution $\tilde{\Lambda}(t, \tilde{\mathscr{A}}, v)$ is defined by

$$(8.6) \qquad \langle f, \tilde{\Lambda}(\theta, \tilde{\mathscr{A}}, v)\rangle = \mathring{E}_\nu\langle f(\hat{X}(\theta))L(\theta, \tilde{\mathscr{A}})/\sigma_t(Y, U)) \qquad \text{for } f \in C_b(R^n).$$

Hence (8.5) implies

(8.7) $\tilde{\Lambda}(\theta, \tilde{\mathscr{A}}, v) = \Lambda(\theta, Y, q_U, \nu)$, $\mathring{P}$-almost surely. We sometimes put $\tilde{\Lambda}(\theta, \tilde{\mathscr{A}}, \nu) = \tilde{\Lambda}(\theta, Y, U, \nu)$ and $\tilde{J}(\theta, \tilde{\mathscr{A}}, \nu, \phi) = \mathring{E}\phi(\tilde{\Lambda}(\theta, \tilde{\mathscr{A}}, \nu)) = \tilde{J}(\theta, Y, U, \nu, \phi)$.

We approximate U_n of (8.1) (say W for simplicity) by a switching usual control, by a routine method, i.e. $\tilde{U}_k$ and $\tilde{U}_{k,p}$ are defined as follows

$$(8.7) \qquad \tilde{U}_k(t) = 2^k \int_{t-2^{-k}}^t W(s)ds \quad \text{and} \quad \tilde{U}_{k,p}(t) = \tilde{U}_k\left(\frac{[2^p t]}{2^p}\right).$$

Then $\lim_{k\to\infty} \lim_{p\to\infty} \tilde{U}_{k,p}(t) = W$ in $L^2([0\ T] \times \Omega)$. This fact implies that there exists an approximate switching usual control W_k, which is $\sigma_t(U)$-progressively measurable and satisfies

$$(8.8) \qquad \mathring{E}\int_0^T | W_k(s) - W(s)|^2 ds \to 0, \quad \text{as} \quad k \to \infty$$

and

$$(8.9) \qquad \tilde{J}(t, Y, W, \nu, \phi) = \lim_{k\to\infty} \tilde{J}(t, Y, W_k, \nu, \phi).$$

By (8.8) some subsequence of W_k converges to W a.e. in $[0\ T] \times \Omega$, we assume "$W_k \to W$ a.e." for simplicity. Therefore, for a.a. $\omega(\mathring{P})$,

$$\begin{aligned}
&\int_\Gamma g(u)\tilde{q}_{W_k}(t, du) = \int_0^t \int_\Gamma g(u)\delta_{W_k}(du)ds \\
(8.10) \qquad &= \int^t g(W_k(s))ds \xrightarrow[k\to\infty]{} \int_0^t g(W(s))ds = \int_\Gamma g(u)\tilde{q}_W(t, du) \\
&\qquad \text{for any } t \text{ and } g \in C_b(\Gamma).
\end{aligned}$$

This implies, as $k \to \infty$.

(8.11) $$\hat{d}_T(\tilde{q}_{W_k}, \tilde{q}_W) \to 0\,, \qquad \mathring{P}\text{-almost surely}.$$

Hence by Propositions 2.2 and the Corollary of Proposition 3.1 we see

(8.12) $$\mathscr{A}_{W_k} \to \mathscr{A}_W \quad \text{and} \quad (X_{W_k}, \mathscr{A}_{W_k}) \to (X_W, \mathscr{A}_W) \text{ in } D_T\,.$$

Therefore Theorem 6 implies that, for $^\forall\theta$,

(8.13) $$\tilde{\Lambda}(\theta, Y, W_k, \nu) \to \tilde{\Lambda}(\theta, Y, W, \nu) \qquad \text{in metric } \delta\,.$$

So we have

PROPOSITION 8.1. *For $\mathscr{A} = (\Omega, F, P, \xi, B, Y, q)$ there exists a switching usual control W_k, $k = 1, 2, \cdots$, which is $\sigma_t(q)$-progressively measurable and for any θ*

$$\tilde{\Lambda}(\theta, Y, W_k, \nu) \to \Lambda(\theta, Y, q, \nu) \qquad \textit{in metric } \delta, \textit{ as } k \to \infty\,.$$

This means that switching usual controls are rich enough in the class of relaxed controls.

Put $\mathfrak{A}_N$ = totality of usual admissible systems whose usual controls are switching with time interval 2^{-N}, i.e. $\tilde{\mathscr{A}} = (\Omega, F, \mathring{P}, \xi, B, Y, U) \in \mathfrak{A}_N$, iff $U(t) = U([2^N t]/2^N)$. We denote $\tilde{\mathscr{A}} \in \mathfrak{A}_0$, if U is constant control. So $\mathfrak{A}_0 \subset \mathfrak{A}_N$. When $P_\xi = \nu/\|\nu\|$, we say $\tilde{\mathscr{A}} \in \mathfrak{A}_N(\nu/\|\nu\|)$. Put $\mathfrak{A} = \bigcup_{N=0}^\infty \mathfrak{A}_N$.

From Proposition 8.1 we see

(8.14) $$\begin{aligned} S(t)\phi(\nu) &= \sup_{\tilde{\mathscr{A}} \in \mathfrak{A}(\nu/\|\nu\|)} \tilde{J}(t, \tilde{\mathscr{A}}, \nu, \phi) \\ &= \lim_{N \to \infty} \sup_{\tilde{\mathscr{A}} \in \mathfrak{A}_N(\nu/\|\nu\|)} \tilde{J}(t, \tilde{\mathscr{A}}, \nu, \phi)\,. \end{aligned}$$

Define $Q = Q_N$ by

(8.15) $$Q\phi(\nu) = \sup_{\tilde{\mathscr{A}} \in \mathfrak{A}_N(\nu/\|\nu\|)} J(2^{-N}, \tilde{\mathscr{A}}, \nu, \phi)\,.$$

We remark that

(8.16) $$\begin{aligned} J(2^{-N}, \tilde{\mathscr{A}}, \nu, \phi) &= \mathring{E}\phi(\tilde{\Lambda}(2^{-N}, Y, U(0), \nu)) \\ &= \int_\Gamma \mathring{E}(\phi(\tilde{\Lambda}(2^{-N}, Y, U(0), \nu))/U(0) = u) P_{U_0}(du)\,. \end{aligned}$$

Since $(Y(\theta), \theta \geq 0)$ is independent of $U(0)$,

(8.17) $$\mathring{E}(\phi(\tilde{\Lambda}(2^{-N}, Y, U(0), \nu/U(0) = u))) = \tilde{J}(2^{-N}, Y, u, \nu, \phi)\,.$$

Moreover the value of the left side depends only on N, u, ν, ϕ, since Y is a Brownian motion with respect to $\mathring{P}(\cdot/U(0) = u)$. We put

$$\tilde{J}(2^{-N}, Y, u, \nu, \phi) = \Phi_N(u, \nu; \phi) \tag{8.18}$$

and

$$\tilde{Q}\phi(\nu) = \sup_{u \in \Gamma} \Phi_N(u, \nu, \phi). \tag{8.19}$$

Then $\Phi_N(u, \nu, \phi)$ is continuous in $(u, \nu) \in \Gamma \times (M(R^n), \Delta)$ and $\tilde{J}(2^{-N}, Y, u, \nu, \phi) \leq \tilde{Q}\phi(\nu)$. Combining with (8.12), we see

$$\begin{aligned} Q\phi(\nu) \leq \tilde{Q}\phi(\nu) &= \sup_{\tilde{\mathscr{A}} \in \tilde{\mathfrak{A}}_0(\nu/\|\nu\|)} \tilde{J}(2^{-N}, \tilde{\mathscr{A}}, \nu, \phi) \\ &\leq \sup_{\tilde{\mathscr{A}} \in \tilde{\mathfrak{A}}_N(\nu/\|\nu\|)} \tilde{J}(2^{-N}, \tilde{\mathscr{A}}, \nu, \phi) = Q\phi(\nu). \end{aligned} \tag{8.20}$$

Hence we have

$$Q_N = \tilde{Q}. \tag{8.21}$$

Since Γ is compact, using a measurable selection theorem, we can take a Borel measurable mapping, $v = v_\phi : M(R^n) \to \Gamma$ such that

$$\Phi_N(v_\phi(\nu), \nu, \phi)) = \sup_{u \in \Gamma} \Phi_N(u, \nu, \phi). \tag{8.22}$$

This gives

$$Q\phi(\nu) = \tilde{J}(2^{-N}, Y, v_\phi(\nu), \nu, \phi) \tag{8.23}$$

We have, for $\tilde{\mathscr{A}} \in \tilde{\mathfrak{A}}_N(\nu/\|\nu\|)$

$$\begin{aligned} J(2^{-N+1}, \tilde{\mathscr{A}}, \nu, \phi) &= \mathring{E}\phi(\tilde{\Lambda}(2^{-N+1}, Y, U, \nu)) \\ &= \mathring{E}\phi(\tilde{\Lambda}(2^{-N}, Y^+_{2^{-N}}, U^+_{2^{-N}}, \tilde{\Lambda}(2^{-N}, Y, U, \nu))) \\ &\leq (Q\phi)(\tilde{\Lambda}(2^{-N}, Y, U, \nu)) \leqq Q(Q\phi)(\nu) \\ &= Q^2\phi(\nu). \end{aligned} \tag{8.24}$$

Define $U_2 = U_{2,\phi,\nu}$ by

$$U_2(t) = \begin{cases} v_{Q\phi}(\nu), \ 0 \leq t \leq 2^{-N} \\ v_\phi(\tilde{\Lambda}(2^{-N}, Y, v_{Q\phi}(\nu), \nu)), \ 2^{-N} < t. \end{cases} \tag{8.25}$$

Then $\mathscr{A}^* = (\Omega, F, \mathring{P}, \xi, B, Y, U_2) \in \tilde{\mathfrak{A}}_N(\nu/\|\nu\|)$ and

$$\begin{aligned} J(2^{-N+1}, \mathscr{A}^*, \nu, \phi) &= \mathring{E}\phi(\tilde{\Lambda}(2^{-N+1}, Y, U_2, \nu)) \\ &= \mathring{E}\mathring{E}(\phi(\tilde{\Lambda}(2^{-N}, Y^+_{2^{-N}}, v_\phi(\tilde{\Lambda}(2^{-N}, Y, v_{Q\phi}(\nu), \nu)), \\ &\qquad \tilde{\Lambda}(2^{-N}, Y, v_{Q\phi}(\nu), \nu))/\sigma_{2^{-N}}(Y))) \\ &= \mathring{E}(Q\phi)(\tilde{\Lambda}(2^{-N}, Y, v_{Q\phi}(\nu), \nu)) \\ &= Q(Q\phi)(\nu) = Q^2\phi(\nu). \end{aligned} \tag{8.26}$$

Combining with (8.24), we have

$$
\begin{aligned}
Q^2\phi(\nu) &= \sup_{\mathscr{A} \in \mathfrak{A}_N(\nu/\|\nu\|)} \tilde{J}(2^{-N+1}, \tilde{\mathscr{A}}, \nu, \phi) \\
&= \tilde{J}(2^{-N+1}, \mathscr{A}^*, \nu, \phi) .
\end{aligned} \tag{8.27}
$$

Repeating a similar calculation we see

$$
Q^{k+1}\phi(\nu) = \sup_{\mathscr{A} \in \mathfrak{A}_N(\nu/\|\nu\|)} \tilde{J}((k+1)2^{-N}, \tilde{\mathscr{A}}, \nu, \phi) \tag{8.28}
$$

and an optimal one $U_{k+1} = U_{k+1,\nu,\phi}$ is given successively by

$$
U_{k+1}(t) = \begin{cases} U_{k,\nu,Q^k\phi}(t) , & 0 \le t \le k2^{-N} \\ v_\phi(\tilde{\Lambda}(k2^{-N}, Y, U_{k,\nu,Q^k\phi}, \nu) , & k2^{-N} < t . \end{cases}
$$

Recalling (8.14) we see, for binary t (say $j2^{-p}$)

$$
\lim_{N\to\infty} \sup_{\mathscr{A} \in \mathfrak{A}_N(\nu/\|\nu\|)} \tilde{J}(t, \tilde{\mathscr{A}}, \nu, \phi) = S(t)\phi(\nu) = \lim_{N\to\infty} Q_N^{[2^N t]}\phi(\nu)
$$

and an approximate optimal usual switching control is given by $U_{k,\nu,\phi}$.

References

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

[2] P. Billingsley, Convergence of Probability Measures, John Wiley 1968.

[3] J. M. Bismut, Partially observed diffusions and their control, SIAM. J. Control Optim., **20** (1982), 302–309.

[4] M. H. A. Davis and M. Kohlmann, On the nonlinear semigroup of stochastic control under partial observations, preprint.

[5] W. H. Fleming, Generalized solutions in optimal stochastic control, Differential games and control theory II, Lect. Notes in pure and Appl. math., **30**, Dekker.

[6] ——, Nonlinear semigroup for controlled partially observed diffusions, SIAM. J. Control Optim., **20** (1982), 286–301.

[7] —— and M. Nisio, On the existence of optimal stochastic controls, J. Math. Mech., **15** (1966), 777–794.

[8] —— and E. Pardoux, Optimal control for partially observed diffusions, SIAM. J. Control Optim., **20** (1982), 258–285.

[9] U. G. Haussman, On the existence of optimal controls for partially observed diffusions, SIAM. J. Control Optim., **20** (1982), 385–407.

[10] M. Nisio, Stochastic Control Theory, ISI Lect. Note **9**, McMillan India, 1981.

[11] K. R. Parthasarathy, Probability Measures on Metric Spaces, Acad. Press, 1967.

[12] A. V. Skorobod, Studies in the Theory of Random Processes, Addism-Wesly, 1965 (orig. Russian).

W. H. Fleming
Division of Applied Mathematics
Brown University
Prividence R. I.
USA

M. Nisio
Department of Mathematics
Kobe University
Rokko, Kobe 657
Japan