Probab. Theory. Relat. Fields 116, 485–504 (2000)
Digital Object Identifier (DOI) 10.1007/s004409900038

Rainer Buckdahn · Marc Quincampoix · Aurel Răşcanu

Viability property for a backward stochastic differential equation and applications to partial differential equations

Received: 23 April 1998 / Published online: 14 February 2000

Abstract. In the present paper, we study conditions under which the solutions of a backward stochastic differential equation remains in a given set of constraints. This property is the so-called "viability property". In a separate section, this condition is translated to a class of partial differential equations.

1. Introduction

The aim of this paper is to state necessary and sufficient conditions under which the solution of a given backward stochastic differential equation (in short: BSDE)

$$Y_t = Y_T + \int_t^T F(s, Y_s, Z_s)ds - \int_t^T Z_s\, dW_s, \quad t \in [0, T] \ ,$$

remains in a given set K of constraints. We apply our results to a system of semilinear parabolic partial differential equations whose solution can be expressed through the solution of a system of suitable – Forward and backward – stochastic differential equations. This enables us to state an existence result for the above PDE's with constraints.

BSDE have been studied first by Pardoux and Peng [13] in 1990. They have turned out to describe the solution of systems of parabolic PDE by the related Feynman-Kac formula introduced in [15] and [14]. The thus described solutions of PDE's are viscosity solutions, a notion introduced by Crandall and P.L. Lions in the early 80's: we refer to [8] and its bibliography.

The strategy we adopt here to study nonsmooth solutions of PDE's by the mean of viability property – introduced by Aubin [2] – for differential equation has already been extensively used for some first order PDE's: the Hamilton Jacobi

R. Buckdahn, M. Quincampoix: Département de Mathématiques, Université de Bretagne Occidentale, 6 Avenue Victor Le Gorgeu, B.P. 809, 29285 Brest Cedex, France. e-mail: rainer.buckdahn@univ-brest.fr; marc.quincampoix@univ-brest.fr

A. Răşcanu*: Facultatea de Matematică, Universitatea "Alexandru Ioan Cuza", 6600 Iaşi, Romania. e-mail: rascanu@uaic.ro

* The work of this author was done during a visit in 1997 to the University of Brest, Department of Mathematics, whose generous support is gratefully acknowledged.

Mathematics Subject Classification (1991): 60H10, 60H30, 35K65, 35D05, 49L25

equations. For this approach using viability property for control systems to study solutions of Hamilton-Jacobi equations, the reader can be referred to [2] and its bibliography.

To our knowledge viability properties for stochastic differential equations and inclusions have been introduced and studied by Aubin and Da Prato in [3], [4] and [5]. The key point of their work consists in the introduction of a "stochastic tangent cone" which generalizes the well-known Bouligand's contingent cone used for deterministic systems. Our approach differs from their methods and bases on the convexity of the distance function of K (when K is a closed convex set). This enables us to deduce some condition in differential form on the distance function of K which is necessary as well as sufficient, and which generalizes the well-known Nagumo condition for first order differential equations with constraints (cf [10], [2]).

Let us now explain how this paper is organized. In the first section, we recall the basic statement of [13] on BSDE's and deduce some basic estimates for BSDE's. Then, in the second section, we state and prove the main result of the paper concerning the viability for BSDE. Finally, in the last section, we apply our main result to a class of systems of semilinear parabolic PDE's.

2. Backward stochastic viability in closed sets

Let $(\Omega, \mathscr{F}, P, \{\mathscr{F}_t, t \geq 0\})$ be a complete stochastic basis such that $\mathscr{F}_0$ contains all P-null elements of $\mathscr{F}$, $\mathscr{F}_{t^+} = \bigcap_{\varepsilon>0} \mathscr{F}_{t+\varepsilon}$, $t \geq 0$, and $\mathscr{F} = \mathscr{F}_T$, and suppose that the filtration is generated by a d-dimensional standard Wiener process $W = (W_t)_{0\leq t\leq T}$. By $T > 0$ we denote the finite real time horizon. We consider the following Backward stochastic differential equation – or shorter BSDE

$$Y_t = Y_T + \int_t^T F(s, Y_s, Z_s)\, ds - \int_t^T Z_s dW_s, \quad t \in [0, T] \ , \tag{1}$$

$$Y_T = \xi \ , \tag{2}$$

where $\xi \in L^2(\Omega, \mathscr{F}_T, P, \mathbb{R}^N)$ is a given random variable and $F : \Omega \times [0, T] \times \mathbb{R}^N \times \mathscr{L}(\mathbb{R}^d, \mathbb{R}^N) \mapsto \mathbb{R}^N$ a measurable function; the assumption on F will be specified below.

Throughout this paper, for any Euclidian space H, we denote by $L^2_{ad}(\Omega, C([0, T], H))$ the closed linear subspace of adapted processes of $L^2(\Omega, \mathscr{F}, P, C([0, T], H))$, and $L^2_{ad}(\Omega\times]0, T[, H))$ is the Hilbert space of adapted measurable stochastic processes $f : \Omega \times [0, T] \mapsto H$ such that

$$\|f\|_2 := (E \int_0^T \|f(t)\|^2\, dt)^{\frac{1}{2}} < \infty \ .$$

We suppose that there are some nonnegative real constants L, L_1, M such that

$$\begin{cases} i) \quad F(\cdot, \cdot, y, z) \text{ is progressively measurable,} \\ \qquad F(\omega, \cdot, y, z) \text{ is continuous} \\ ii) \quad \|F(t, y, z) - F(t, y', z')\| \leq L(\|y - y'\| + \|z - z'\|) \\ iii) \ \sup_{t\leq T} \|F(t, 0, 0)\| \in L^2(\Omega, \mathscr{F}_T, P) \end{cases} \tag{3}$$

for all (t, y, z, y', z'), P-almost everywhere on Ω.

Let us recall the existence and uniqueness result for BSDE (cf [14] or [6] for generalization to integral-partial differential equations):

Proposition 2.1. *Let (3) holds true.*

Then for any given $\xi \in L^2(\Omega, \mathscr{F}, P, \mathbb{R}^N)$, there exists an unique solution to (1) (2)

$$(Y, Z) \in L^2_{ad}(\Omega, C([0,T], \mathbb{R}^N)) \times L^2_{ad}(\Omega \times]0,T[, \mathscr{L}(\mathbb{R}^d, \mathbb{R}^N)) \ .$$

Remark that the notion of BSDE generalizes the well-known martingale representation property. Indeed, in the particular case of $F = 0$ we have the following Lemma (cf for instance [17])

Lemma 2.2. *For any $\xi \in L^2(\Omega, \mathscr{F}_T, P, \mathbb{R}^N)$, there is a unique $R(\xi)$ belonging to $L^2_{ad}(\Omega \times]0,T[, \mathscr{L}(\mathbb{R}^d, \mathbb{R}^N))$ such that*

$$\xi = E(\xi) + \int_0^T R_s(\xi) dW_s \ . \tag{4}$$

After this recall we give now the definition of stochastic viability.

Definition 2.3. *Let K be a nonempty closed subset of $\mathbb{R}^N$.*

a)- A stochastic process $\{Y_t,\ t \in [0,T]\}$ is viable *in K if and only if for P-almost $\omega \in \Omega$,*

$$Y_t(\omega) \in K, \quad \forall\, t \in [0,T] \ .$$

b)- The closed set K enjoys the Backward Stochastic Viability Property – denoted BSVP – for the equation (1) if and only if:

$\forall\, \tau \in [0,T], \forall\, \xi \in L^2(\Omega, \mathscr{F}_\tau, P, K)$, there exists a solution (Y, Z) to BSDE (1) (2) over the time interval $[0, \tau]$,

$$\begin{cases} Y_s = \xi + \int_s^\tau F(r, Y_r, Z_r) dr - \int_s^\tau Z_r\, dW_r,\ s \in [0,\tau], \\ (Y, Z) \in L^2_{ad}(\Omega, C([0,\tau], \mathbb{R}^N)) \times L^2_{ad}(\Omega \times]0,\tau[, \mathscr{L}(\mathbb{R}^d, \mathbb{R}^N)) \end{cases} \tag{5}$$

such that $\{Y_s,\ s \in [0,\tau]\}$ is viable in K.

Let us define for any closed set $K \subset \mathbb{R}^N$ the – multivalued – projection of a point a onto K:

$$\Pi_K(a) := \{ b \in K \mid \|a - b\| = \min_{c \in K} \|a - c\| = d_K(a) \} \ .$$

Recall that $\Pi_K(a)$ is a singleton[1] whenever d_K is differentiable at the point a. According to Motzkin's Theorem Π_K is single-valued if and only if K is convex.

[1] When $\Pi_K(a)$ is a singleton, we also denote by $\Pi_K(a)$ the unique element of $\Pi_K(a)$.

Let us recall that $d_K^2(\cdot)$ is convex when K is convex, and thus, due to Alexandrov's Theorem [1], $d_K^2(\cdot)$ is almost everywhere twice differentiable[2].

After this preparation, we can state our first main result.

Theorem 2.4. *Suppose that $F : \Omega \times [0, T] \times \mathbb{R}^N \times \mathbb{R}^{N\times d} \mapsto \mathbb{R}^N$ is a measurable function which satisfies condition (3). Let K be a nonempty closed set.*

If K enjoys the BSVP property for (1), then the set K is convex.

Proof of Theorem 2.4. Assume that K is not convex. We have to prove that some $\xi \in L^2(\Omega, \mathscr{F}_T, P, K)$ exists such that the solution (Y, Z) to BSDE (1) satisfies

$$P\{Y_t \notin K\} > 0, \text{ for some } t \in [0, T] \ . \tag{6}$$

If K is not convex, we can find a and b belonging to ∂K such that

$$a \neq b \quad \text{and} \quad K \cap]a, b[= \emptyset \ ,$$

where $]a, b[:= \{a + t(b - a) \,|\, 0 < t < 1\}$.

Fix $\varepsilon > 0$ such that

$$d_K\left(\frac{a+b}{2}\right) > 2\varepsilon \ .$$

Let us denote by W_t^1 the first coordinate of the d-dimensional Wiener process W_t and define

$$\xi := aI_{\{W_T^1 < \frac{1}{2}\}} + bI_{\{W_T^1 \geq \frac{1}{2}\}} \ .$$

Then

$$E[\xi|\mathscr{F}_t] = a + (b - a)\Phi\left(\frac{-\frac{1}{2} + W_t^1}{(T - t)^{\frac{1}{2}}}\right), \quad t < T \ , \tag{7}$$

where

$$\Phi(r) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{r} e^{-\frac{x^2}{2}} dx, \quad r \in \mathbb{R} \ .$$

Let (Y, Z) be the solution to BSDE (1).

Then, thanks to (3-i)-iii), $M := E[\int_0^T \|F(s, Y_s, Z_s)|^2\|ds] < +\infty$.

A standard argument yields

$$\begin{aligned} &E(\|Y_t - E[\xi|\mathscr{F}_t]\|^2) + E\int_t^T \|Z_r - R_r(\xi)\|^2 dr \\ &\leq (T - t)E\int_t^T \|F(s, Y_s, Z_s)|^2\|ds \leq (T - t)M \ . \end{aligned}$$

[2] By twice differentiable, we mean that the function admits a second order Taylor expansion. We underline that this may hold true even if the first derivative is not continuous.

Consequently,

$$
\begin{aligned}
&P\{\|Y_t - \frac{a+b}{2}\| \le 2\varepsilon\} \ge P\{\|E[\xi|\mathscr{F}_t] - \frac{a+b}{2}\| \le \varepsilon, \|Y_t - E[\xi|\mathscr{F}_t]\| \le \varepsilon\} \\
&= P\{\|E[\xi|\mathscr{F}_t] - \frac{a+b}{2}\| \le \varepsilon\} \\
&\quad - P\{\|E[\xi|\mathscr{F}_t] - \frac{a+b}{2}\| \le \varepsilon, \|Y_t - E[\xi|\mathscr{F}_t]\| > \varepsilon\} \\
&\ge P\{\|E[\xi|\mathscr{F}_t] - \frac{a+b}{2}\| \le \varepsilon\} - P\{\|Y_t - E[\xi|\mathscr{F}_t]\| > \varepsilon\} \\
&\ge P\{\|E[\xi|\mathscr{F}_t] - \frac{a+b}{2}\| \le \varepsilon\} - \frac{1}{\varepsilon^2} E(\|Y_t - E[\xi|\mathscr{F}_t]\|^2) \ .
\end{aligned}
$$

The last estimation comes from the Chebychev inequality. Thus

$$
P\{\|Y_t - \frac{a+b}{2}\| \le 2\varepsilon\} \ge P\{\|E[\xi|\mathscr{F}_t] - \frac{a+b}{2}\| \le \varepsilon\} - \frac{(T-t)M}{\varepsilon^2} \ .
$$

Note that by (7), $P\{\|E[\xi|\mathscr{F}_t] - \frac{a+b}{2}\| \le \varepsilon\} > 0,\ t \in \ (0, T)$.

Let us choose $t \in \ ((T - \frac{\varepsilon^2}{M} P\{\|E[\xi|\mathscr{F}_t] - \frac{a+b}{2}\| \le \varepsilon\})^+, T)$. Then

$$
P\{\|Y_t - \frac{a+b}{2}\| \le 2\varepsilon\} \ge P\{\|E[\xi|\mathscr{F}_t] - \frac{a+b}{2}\| \le \varepsilon\} - \frac{M(T-t)}{\varepsilon^2} > 0 \ .
$$

But this is a contradiction to the BSVP, and hence K must be convex.

Q.E.D.

Since the previous results means that only convex sets could have the BSVP, we restrict our attention to closed convex sets.

Theorem 2.5. *Suppose that $F : \Omega \times [0, T] \times \mathbb{R}^N \times \mathscr{L}(\mathbb{R}^d, \mathbb{R}^N) \mapsto \mathbb{R}^N$ is a measurable function which satisfies condition (3). Let K be a nonempty closed convex set.*

The set K enjoys the BSVP property for (1) if and only if

$$
\begin{cases}
\forall\ (t, z) \in \ [0, T] \times \mathscr{L}(\mathbb{R}^d, \mathbb{R}^N) \\
\text{and for all } y \in \mathbb{R}^N \text{ such that } d_K^2(\cdot) \text{ is twice differentiable at } y, \\
4\langle y - \Pi_K(y), F(t, y, z)\rangle \le \langle D^2 d_K^2(y) z, z\rangle + C d_K^2(y),\ P\text{-a.e.} \\
\text{where } C > 0 \text{ is a constant which does not depend on } (t, y, z)\ .
\end{cases}
\tag{8}
$$

Let us notice that, under assumption (3-ii), condition (8) takes form:

$$
4\langle y - \Pi_K(y), F(t, \Pi_K(y), z)\rangle - \langle D^2 d_K^2(y) z, z\rangle \le (C + 4L) d_K^2(y)
$$

On the other hand, for some $C' > 0$, the condition

$$4\langle y - \Pi_K(y), F(t, \Pi_K(y), z)\rangle - \langle D^2 d_K^2(y)z, z\rangle \le C' d_K^2(y) \tag{9}$$

implies (8) with constant $C' + 4L$ instead C.

This shows that (8) is a condition only on the values of $F(t, \cdot, z)$ on ∂K. Recall also that the behaviour of d_K on $\mathbb{R}^N$ is completely determined by that of ∂K.

Remark. Because $D^2[d_K^2]$ is almost everywhere positive semidefinite, a sufficient condition for the BSVP property of K is that there exists some constant $C \ge 0$ such that

$$\langle y - \Pi_K(y), F(t, y, z)\rangle \le C d_K^2(y), \text{ for all } (t, y, z) \ . \tag{10}$$

A very similar condition

$$\langle n(y), F(t, y, z)\rangle \le 0, \ \forall \ (t, y, z) \in [0, T] \times \partial K \times \mathscr{L}(\mathbb{R}^d, \mathbb{R}^N)$$

(where $n(x)$ is a normal vector – in the sense of convex analysis – at x to K) can be found in [11]. □

Example. Let K be a convex closed subset of $\mathbb{R}^N$ such that its boundary is a $(N-1)$ – submanifold of class C^2. Let us write the condition (8) at a point (t, y, z). Denote by $\kappa_1, \dots, \kappa_{N-1}$ the principal curvatures of ∂K at point $\Pi_K(y)$ along $\nabla d_K(y)$. Then (8) becomes

$$\begin{aligned} 4\langle y - \Pi_K(y), F(t, y, z)\rangle \le \frac{1}{2}\sum_{i=1}^{N-1} \frac{d_k(y)\kappa_i}{1 + \kappa_i d_K(y)} \|\tilde{z}_i\|^2 \\ + \frac{1}{2}\|\tilde{z}_N\|^2 + C d_K^2(y) \end{aligned}$$

where $(\tilde{z}_1, \dots, \tilde{z}_N)$ are the coordinates of z in the relative basis (tangential plane, normal) to K at $\Pi_K(y)$.

From this expression, one can easily check that the following example with $K = B(0, 1) \subset \mathbb{R}^2$ and $d = 1$ satisfies (8) but not (10):

$$F(s, y, z) = \frac{1}{2}\frac{\tilde{z}_1^2 \wedge \|\tilde{z}_1\|}{(1 + d_K(y))^2} y + \frac{\|\tilde{z}_2\|}{1 + d_K(y)}(y - \Pi_K(y))$$

where

$$\begin{cases} \tilde{z}_2 = \langle z, \frac{y}{\|y\|}\rangle \\ \tilde{z}_1 = z - \tilde{z}_2 \frac{y}{\|y\|} \ . \end{cases}$$ □

Before proving the preceeding theorem, we give an application to PDE's.

3. Viability for viscosity solutions to semilinear parabolic PDE's

In this section, we show that in the Markovian framework the BSVP provides new results for suitable systems of semilinear parabolic PDE's.

Let us consider the following PDE

$$\begin{cases} \frac{\partial}{\partial t}u_i(t,x) + Au_i(t,x) + f_i(x, u(t,x), (\nabla u_i \sigma)(t,x)) = 0 \\ u(T,x) = H(x) \\ \text{for } (t,x) \in [0,T] \times \mathbb{R}^d, \ 1 \le i \le N \ , \end{cases} \tag{11}$$

where $H \in C(\mathbb{R}^d, \mathbb{R}^N)$ is given and

$$A = \frac{1}{2} \sum_{i,j=1}^{N} (\sigma\sigma^\star)_{i,j}(s) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{N} b_i(x) \frac{\partial}{\partial x_i} \ .$$

We assume that $b\colon \mathbb{R}^d \mapsto \mathbb{R}^d$ and $\sigma\colon \mathbb{R}^d \mapsto \mathscr{L}(\mathbb{R}^d, \mathbb{R}^d)$ are Lipschitz and $f = (f_1, \ldots, f_N) \in C([0,T] \times \mathbb{R}^d \times \mathbb{R}^N \times \mathscr{L}(R^d, \mathbb{R}^N)$ is a bounded function with

$$f_i(t,x,y,z) := f_i(t,x,y,z_i), \ (t,x,y) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^N,$$
$$z = (z_1, \ldots, z_N) \in \mathscr{L}(\mathbb{R}^d, \mathbb{R}^N) \ .$$

Suppose that there exist some $p \ge 2$ and some $L > 0$ such that

$$\begin{cases} a) \ \|H(x)\| \le L(1 + \|x\|^p) \\ b) \ \|f(t,x,y,z)\| \le L(1 + \|x\|^p + \|y\| + \|z\|) \\ c) \ \|f(t,x,y,z) - f(t,x,\bar{y},\bar{z})\| \le L(\|y - \bar{y}\| + \|z - \bar{z}\|) \\ d) \ \|f_i(t,x,y,z_i) - f_i(t,x',y,z_i))\| \le m_R^i(\|x - x'\|(1 + \|z_i\|)) \\ \quad \text{for all } t \in [0,T], \ \|x\|, \|x'\|, \|y\| \le R, \\ \quad \text{where } \lim_{s \to 0^+} m_R^i(s) = 0, \ 1 \le i \le N \text{ and } R \ge 1 \ . \end{cases} \tag{12}$$

In [14], Pardoux and Peng have studied the correlation between PDE (11) and the BSDE.

$$Y_s^{t,x} = H(X_T^{t,x}) + \int_s^T f(X_r^{t,x}, Y_r^t, Z_r^{t,x})dr - \int_s^T Z_r^{t,x}\, dW_r \ , \tag{13}$$

where $X^{t,x} = \{X_s^{t,x}, \ t \le s \le T\}$ is the unique solution of the forward SDE

$$X_s^{t,x} = x + \int_t^s b(X_r^{t,x})dr + \int_t^s \sigma(X_r^{t,x})dW_r, \ t \le s \le T \ .$$

In [6], a generalization to integral-partial differential equations is provided.

In both papers it is shown that $u(t,x) = Y_t^{t,x}$ is a viscosity solution[3] to PDE (11) and uniqueness results are presented. We refer to the following one:

[3] The reader can refer to [8] and its bibliography, for a definition of viscosity solutions and a detailed study.

Theorem 3.1. *[6] Under the above assumptions, there exists at most one viscosity solution u such that*

$$\lim_{\|x\|\longrightarrow+\infty} |u(t,x)|e^{-\tilde{A}(Log\|x\|)^2} = 0 \ , \tag{14}$$

uniformely in $t \in [0,T]$, for some $\tilde{A} > 0$.

In particular, the function $u(t,x) = Y_t^{t,x}$ is the unique viscosity solution to (11) in the class of solutions which satisfy (14) for some real $\tilde{A} > 0$.

Let us now define the viability for PDE (11).

Definition 3.2. *The PDE (11) enjoys the viability property with respect to the closed set K if and only if for any $H \in C_p(\mathbb{R}^d, \mathbb{R}^N)$ taking its values in K the viscosity solution to (11) satisfies*

$$\forall (t,x) \in [0,T] \times \mathbb{R}^d, \quad u(t,x) \in K. \tag{15}$$

This enables us to state the main theorem of this section:

Theorem 3.3. *Suppose that b and σ are Lipschitz and $f \in C([0,T] \times \mathbb{R}^d \times \mathbb{R}^N \times \mathscr{L}(R^d, \mathbb{R}^N))$ is a bounded function. Let us assume furthermore that $K \subset \mathbb{R}^N$ is a closed convex nonempty set and that (12) holds true.*

Under these assumptions, we have

(a) *If the PDE (11) enjoys the viability property with respect to K, then there exists a constant $C > 0$ such that, for all $(t,x,y,z) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^N \times \mathscr{L}(R^d, \mathbb{R}^N)$ such that d_K^2 is twice differentiable at y,*

$$\langle y - \Pi_K(y), f(t,x,y,z\sigma(x))\rangle \leq \frac{1}{4}\langle D^2(d_K^2(y))(z\sigma(x)), z\sigma(x)\rangle + Cd_K^2(y) \ .$$

(b) *Conversely, the above necessary condition is sufficient when, moreover, σ is of class C^2.*

Proof (necessity). Let $(t,x,y,z) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^N \times \mathscr{L}(R^d, \mathbb{R}^N)$, $0 < \varepsilon < t$, and let

$$H(x') = y + z(x'-x), \ x' \in \mathbb{R}^d \ .$$

We set

$$F(\omega,s,y,z) = f(\omega, X_s^{t-\varepsilon,x}(\omega), y, z), \quad t-\varepsilon \leq s \leq t$$
$$\xi = H(X_t^{t-\varepsilon,x}) = y + z(X_t^{t-\varepsilon,x} - x) \ .$$

Then using the argument used for the proof of Theorem 2.5, we get the necessity of the above condition.

Proof (sufficiency). A standard approximation procedure applied to BSDE (13) gives

$$Z_s^{t,x} \in span\{z\sigma(X_s^{t,x}) \,|\, z \in \mathscr{L}(R^d, \mathbb{R}^N)\} \ ,$$
$$dsdP\text{-a.e. on } [t,T], \quad 0 \leq t \leq T \ .$$

This allows to proceed exactely as in the proof of the sufficiency for Theorem 2.5.

Q.E.D.

4. Appendix: Proofs of main theorem

For the proof of theorem 2.5, we need an auxiliary result on BSDEs. Given any $\xi \in L^2(\Omega, \mathscr{F}_T, P, \mathbb{R}^N)$ we denote by (Y^ξ, Z^ξ) the unique solution to BSDE (1) and by $R(\xi)$ the process associated to ξ by Lemma 2.2. With these notations we can state

Proposition 4.1. *Suppose that (3) holds true. Then there exists a real constant $C_0 > 0$ such that for all $\xi \in L^2(\Omega, \mathscr{F}_T, P, \mathbb{R}^N)$,*

$$E[\sup_{s\in[t,T]} \|Y_s^\xi\|^2|\mathscr{F}_t] + E[\int_t^T \|Z_s^\xi\|^2\, ds|\mathscr{F}_t]$$
$$\leq C_0[E[\|\xi\|^2|\mathscr{F}_t] + E[(\int_t^T \|F(s,0,0)\|ds)^2|\mathscr{F}_t]], \quad t \in [0,T] \ , \qquad (16)$$

and for all $\varepsilon \in [0, T]$,

$$\begin{cases} E(\sup_{s\in[T-\varepsilon,T]} \|Y_s^\xi - E(\xi|\mathscr{F}_s)\|^2|\mathscr{F}_{T-\varepsilon}) \\ + E[\int_{T-\varepsilon}^T \|Z_s^\xi - R_s(\xi)\|^2\, ds|\mathscr{F}_{T-\varepsilon}] \\ \leq C_0\varepsilon E[\int_{T-\varepsilon}^T \|F(s, E(\xi|\mathscr{F}_s), R_s(\xi))\|^2\, ds|\mathscr{F}_{T-\varepsilon}] \ . \end{cases} \qquad (17)$$

Proof. Let $\xi \in L^2(\Omega, \mathscr{F}_T, P, \mathbb{R}^N)$ and $Y = Y^\xi,\ Z = Z^\xi$. Itô's formula applied to $e^{\lambda t}\|Y_t\|^2$ yields for $s \in [0, T]$,

$$e^{\lambda s}\|Y_s\|^2 + \int_s^T e^{\lambda r}(\lambda\|Y_r\|^2 + \|Z_r\|^2)dr$$
$$= e^{\lambda T}\|\xi\|^2 + 2\int_s^T e^{\lambda r}(F(r, Y_r, Z_r), Y_r)dr - 2\int_s^T e^{\lambda r}(Y_r, Z_r\, dW_r) \ .$$

But,

$$2(F(r,y,z),y) = 2(F(r,y,z) - F(r,y,0),y) + 2(F(r,y,0) - F(r,0,0),y)$$
$$+2(F(r,0,0),y)$$
$$\leq \frac{1}{2}\|z\|^2 + 2(L^2 + L)\|y\|^2 + 2(y, F(r,0,0)) \ ,$$

and then

$$\begin{cases} e^{\lambda s}\|Y_s\|^2 + \int_s^T e^{\lambda r}[(\lambda - 2L^2 - 2L)\|Y_r\|^2 + \frac{1}{2}\|Z_r\|^2]dr \\ \leq e^{\lambda T}\|\xi\|^2 + 2\int_s^T e^{\lambda r}(Y_r, F(r,0,0))dr - 2\int_s^T e^{\lambda r}(Y_r, Z_r\, dW_r) \ , \end{cases} \qquad (18)$$

for $0 \le s \le T$. By replacing λ by $2L^2 + 2L$, inequality (18) yields

$$\begin{cases} \int_s^T e^{\lambda r}\|Z_r\|^2\,dr \le 2e^{\lambda T}\|\xi\|^2 + 4\int_s^T e^{\lambda r}(Y_r, F(r,0,0))dr \\ -4\int_s^T e^{\lambda r}(Y_r, Z_r\,dW_r),\ \text{for}\ \ 0 \le s \le T\ \ . \end{cases} \tag{19}$$

By the Burkholder-Davis-Gundy inequality, we have

$$\begin{aligned} &E[\sup_{s\in[t,T]}|\int_s^T e^{\lambda r}(Y_r, Z_r\,dW_r)|\,|\mathscr{F}_t] \\ &\le 2E[\sup_{s\in[t,T]}|\int_t^s e^{\lambda r}(Y_r, Z_r\,dW_r)|\,|\mathscr{F}_t] \\ &\le 6E([\int_t^T e^{2\lambda r}\|Y_r\|^2\|Z_r\|^2\,dr]^{\frac{1}{2}}|\mathscr{F}_t) \\ &\le 6E([\sup_{s\in[t,T]}(e^{\frac{1}{2}\lambda s}\|Y_s\|)(\int_t^T e^{\lambda r}\|Z_r\|^2\,dr)^{\frac{1}{2}}]|\mathscr{F}_t) \\ &\le \frac{1}{4}E[(\sup_{s\in[t,T]}e^{\lambda s}\|Y_s\|^2)|\mathscr{F}_t] + 36E[\int_t^T e^{\lambda r}\|Z_r\|^2\,dr|\mathscr{F}_t]\ \ . \end{aligned}$$

Let us denote by $C > 0$ a real constant depending only on L and T, and to which we allow to change from one formula to the other. From (18) and (19) we obtain

$$\begin{aligned} &\frac{1}{2}E[\sup_{s\in[t,T]}(e^{\lambda s}\|Y_s\|^2)|\mathscr{F}_t] \\ &\le Ce^{\lambda T}E[\|\xi\|^2|\mathscr{F}_t] + CE[\int_t^T e^{\lambda r}\|Y_r\|\|F(r,0,0)\|dr|\mathscr{F}_t] \\ &\le Ce^{\lambda T}E[\|\xi\|^2|\mathscr{F}_t] + CE[(\sup_{r\in[t,T]}e^{\frac{1}{2}\lambda r}\|Y_r\|\int_t^T e^{\frac{1}{2}\lambda r}\|F(r,0,0)\|dr)|\mathscr{F}_t]\ \ . \end{aligned}$$

Hence,

$$\begin{cases} E[\sup_{s\in[t,T]}(e^{\lambda s}\|Y_s\|^2)|\mathscr{F}_t] \le Ce^{\lambda T}E[\|Y_T\|^2|\mathscr{F}_t] \\ +CE[\{\int_t^T e^{\frac{1}{2}\lambda r}\|F(r,0,0)\|dr\}^2|\mathscr{F}_t]\ \ . \end{cases} \tag{20}$$

This, together with (19), gives

$$\begin{cases} E[\int_t^T e^{\lambda s}\|Z_s\|^2\,ds|\mathscr{F}_t] \le Ce^{\lambda T}E[\|\xi\|^2|\mathscr{F}_t] \\ + CE[(\int_t^T e^{\frac{\lambda r}{2}}\|F(r,0,0)\|dr)^2|\mathscr{F}_t]\ \ . \end{cases} \tag{21}$$

From (20) and (21) one can easily obtain (16).

In order to prove (17), note that (1) is equivalent to

$$\bar{Y}_t = \int_t^T G(s, \bar{Y}_s, \bar{Z}_s)ds - \int_t^T \bar{Z}_s dW_s, \quad t \in [0, T] \ , \tag{22}$$

where $\bar{Y}_t := Y_t - E(\xi|\mathscr{F}_t)$, $\bar{Z}_t := Z_t - R_t(\xi)$ and $G(s, \bar{y}, \bar{z}) := F(s, \bar{y} + E(\xi|\mathscr{F}_s), \bar{z} + R_s(\xi))$. Inequality (16) applied to $(\bar{Y}_t, \bar{Z}_t)$ implies (17).

Q.E.D.

Now we are ready to prove Theorem 2.5.

Proof of Theorem 2.5. This proof is splitted into several steps.

(a) Necessity. Let $t \in [0, T]$ and $\varepsilon > 0$ be such that, for some $t^\star \geq 0$, $t_\varepsilon := t - \varepsilon > t^\star \geq 0$. Fix $y \in \mathbb{R}^N$, $z \in \mathscr{L}(\mathbb{R}^N, \mathbb{R}^d)$ and $\xi := y + z(W_t - W_{t_\varepsilon})$.

Denote by (Y, Z) the unique solution to the BSDE

$$Y_s = \xi + \int_s^t F(r, Y_r, Z_r)dr - \int_s^t Z_r \, dW_r, \quad s \in [t_\varepsilon, t] \ .$$

Furthermore, we introduce the process $\widehat{Y} := \{\widehat{Y_s},\ s \in [t_\varepsilon, t]\}$ as follows:

$$\begin{aligned}\widehat{Y_s} &= \xi + (t - s)F(t_\varepsilon, y, z) - z(W_t - W_s) \\ &= y + (t - s)F(t_\varepsilon, y, z) + z(W_s - W_{t_\varepsilon}),\ t_\varepsilon \leq s \leq t \ .\end{aligned}$$

From Proposition 4.1, we deduce that there exists a constant $C > 0$ which depends only on y and z, and is such that

$$E[\sup_{t_\varepsilon \leq s \leq t} \|Y_s\|^2|\mathscr{F}_{t^\star}] + E[\int_{t_\varepsilon}^t \|Z_r\|^2 \, dr|\mathscr{F}_{t^\star}] \leq C \tag{23}$$

and

$$E[\sup_{t_\varepsilon \leq s \leq t} \|Y_s - E[\xi|\mathscr{F}_s\|^2|\mathscr{F}_{t^\star}] + E[\int_{t_\varepsilon}^t \|Z_r - z\|^2 \, dr|\mathscr{F}_{t^\star}] \leq C\varepsilon^2 \ . \tag{24}$$

Then

$$\begin{cases} E[\sup\limits_{t_\varepsilon \leq s \leq t} \|Y_s - y\|^2|\mathscr{F}_{t^\star}] \leq \\ 2E[\sup\limits_{t_\varepsilon \leq s \leq t} \|Y_s - E[\xi|\mathscr{F}_s\|^2|\mathscr{F}_{t^\star}] + 2E[\sup\limits_{t_\varepsilon \leq s \leq t} \|z(W_s - W_{t_\varepsilon})\|^2|\mathscr{F}_{t^\star}] \\ \leq C'(\varepsilon^2 + \|z\|^2\varepsilon) \leq C\varepsilon, \quad \text{for } \varepsilon \in]0, t - t^\star[\ . \end{cases} \tag{25}$$

Since for $t_\varepsilon \leq s \leq t$,

$$Y_s - \widehat{Y_s} = \int_s^t (F(r, Y_r, Z_r) - F(t_\varepsilon, y, z))dr - \int_s^t (Z_r - z)dW_r \ ,$$

we deduce from (24) and (25),

$$\begin{cases} E[\sup_{t_\varepsilon \le s \le t} \|Y_s - \widehat{Y}_s\|^2 | \mathscr{F}_{t^\star}] + E[\int_{t_\varepsilon}^t \|Z_r - z\|^2 \, dr | \mathscr{F}_{t^\star}] \\ \le C_1 \varepsilon E[\int_{t_\varepsilon}^t \|(F(r, Y_r, Z_r) - F(t_\varepsilon, y, z)\|^2 \, dr | \mathscr{F}_{t^\star}] \\ \le 3C_1 \varepsilon^2 (E[\sup_{t_\varepsilon \le r \le t} \|F(r, y, z) - F(t_\varepsilon, y, z)\|^2 | \mathscr{F}_{t^\star}] \\ + L^2 E[\sup_{t_\varepsilon \le r \le t} \|Y_r - y\|^2 | \mathscr{F}_{t^\star}]) \\ + 3C_1 \varepsilon E[\int_{t_\varepsilon}^t L^2 \|Z_r - z\|^2 \, dr | \mathscr{F}_{t^\star}] \\ \le C \varepsilon^2 \beta_\varepsilon^1 \ , \end{cases} \tag{26}$$

where

$$\beta_\varepsilon^1 = 6C_1 L^2 \varepsilon + 3\frac{C_1}{C} E[\sup_{t_\varepsilon \le s \le t} \|F(s, y, z) - F(t_\varepsilon, y, z)\|^2 | \mathscr{F}_{t^\star}] \ .$$

Observe that by (3):

$$\sup_{t_\varepsilon \le s \le t} \|F(s, y, z) - F(t_\varepsilon, y, z)\|^2 \to 0 \text{ when } \varepsilon \to 0 \ ,$$

and that

$$\begin{aligned} &\sup_{t_\varepsilon \le s \le t} \|F(s, y, z) - F(t_\varepsilon, y, z)\|^2 \\ &\quad \le 4 \sup_{0 \le t \le T} \|F(t, y, z)\|^2 \\ &\quad \le 8 \sup_{0 \le t \le T} \|F(t, 0, 0)\|^2 + 8L^2(\|y\| + \|z\|)^2 \in L^2(\Omega, \mathscr{F}_T, P) \ . \end{aligned}$$

Hence from Lebesgue's dominated convergence theorem, $\beta_\varepsilon^1 \to 0$, P-a.e., as $\varepsilon \to 0$.

Let us now establish two auxilliary results on the processes Y and $\widehat{Y}$ which will enable us to finish the proof of the necessity.

Lemma 4.2. *Under the assumptions made above, there is some real constant* $C = C(y, z) > 0$ *such that*

$$E[|d_K^2(Y_{t_\varepsilon}) - d_K^2(\widehat{Y_{t_\varepsilon}})| \, | \mathscr{F}_{t^\star}] \le C\varepsilon \sqrt{\beta_\varepsilon^1} \ ,$$

for any $\varepsilon > 0$ *with* $t - \varepsilon > t^\star$.

Proof of Lemma 4.2. Note that $\forall\, x, x' \in \mathbb{R}^d,\ \forall p \in \Pi_K(x),\ \forall\, p' \in \Pi_K(x')$:

$$d_K^2(x) - d_K^2(x') \le \|x - x'\|^2 + 2\langle x' - p', x - x'\rangle,$$
$$d_K^2(x) - d_K^2(x') \ge -\|x - x'\|^2 - 2\langle x - p, x' - x\rangle$$

Hence

$$|d_K^2(x) - d_K^2(x')| \le \|x - x'\|^2 + 2\|x - x'\|(\|x - p\| + \|x' - p'\|)$$
$$\le \|x - x'\|^2 + 2\|x - x'\|(d_K(x) + d_K(x'))\ \ .$$

Since $K \ne \emptyset$, there exists an element $a \in K$ such that

$$\forall\, b \in \mathbb{R}^N, \quad d_K(b) \le \|b\| + \|a\|\ \ .$$

Then, we have for some constant $C > 0$ and for any $x, x' \in \mathbb{R}^N$,

$$|d_K^2(x) - d_K^2(x')| \le C(1 + \|x\| + \|x'\|)\|x - x'\|\ \ .$$

Thus, from (23) and (26), for some constant again denoted by $C > 0$,

$$E[|d_K^2(Y_{t_\varepsilon}) - d_K^2(\widehat{Y_{t_\varepsilon}})|\ |\mathscr{F}_{t^\star}] \le C\varepsilon\sqrt{\beta_\varepsilon^1}\ \ .$$

This completes the proof of Lemma 4.2. Q.E.D.

Lemma 4.3. *We can find some $\mathscr{F}_{t^\star}$-measurable random variable $\gamma_\varepsilon = \gamma_\varepsilon(y, z)$ with $\liminf_{\varepsilon\to 0} \gamma_\varepsilon \ge 0$ such that*

$$\begin{cases} E[(d_K^2(\widehat{Y_{t_\varepsilon}}) - d_K^2(\xi))|\mathscr{F}_{t^\star}] \\ \ge \varepsilon\{E[\langle\nabla d_K^2(y), F(t, y, z)\rangle|\mathscr{F}_{t^\star}] - \frac{1}{2}\langle D^2[d_K^2(y)]z, z\rangle + \gamma_\varepsilon\} \end{cases} \tag{27}$$

Proof. We know that the function $d_K^2(\cdot)$ is twice differentiable almost everywhere. Let us denote by Λ_K the set of all points of $\mathbb{R}^N$ where d_K^2 is twice differentiable. This set is of full Lebesgue measure. Let us fix now $y \in \Lambda_K$ and define the following function $\alpha : \mathbb{R}^N \mapsto \mathbb{R}$:

$$\alpha(x) := d_K^2(x + y) - d_K^2(y) - \langle\nabla d_K^2(y), x\rangle - \frac{1}{2}\langle D^2[d_K^2(y)]x, x\rangle\ \ .$$

We shall prove two properties of $\alpha(\cdot)$ we shall need later on. The first one is an obvious consequence of the definition of Λ_K:

$$\lim_{\|x\|\longrightarrow 0} \frac{\alpha(x)}{\|x\|^2} = 0\ \ . \tag{28}$$

The second property is the following

$$\forall\, x \in \mathbb{R}^N, \quad \alpha(x) \le \|x^2\|(1 + \|D^2 d_K^2(y)\|)\ \ . \tag{29}$$

In fact, $\Pi_K(y)$ is a singleton because $\nabla d_K^2(y)$ exists, and

$$2\Pi_K(y) = \nabla[\|y\|^2 - d_K^2(y)] \ .$$

Then

$$\begin{aligned} d_K^2(x+y) - d_K^2(y) - \|x\|^2 &\leq \|(x+y) - \Pi_K(y)\|^2 - \|y - \Pi_K(y)\|^2 - \|x\|^2 \\ &= 2\langle y - \Pi_K(y), x\rangle = \langle \nabla d_K^2(y), x\rangle \ . \end{aligned}$$

Hence

$$\alpha(x) \leq \langle (Id - \frac{1}{2}D^2 d_K^2(y))x, x\rangle \leq \|x\|^2(1 + \|D^2 d_K^2(y)\|) \ .$$

First we substitute $x = \widehat{Y_{t_\varepsilon}} - y = \varepsilon F(t_\varepsilon, y, z)$ in the definition of $\alpha(\cdot)$. This provides us

$$E[d_K^2(\widehat{Y_{t_\varepsilon}})|\mathscr{F}_{t^\star}] = d_K^2(y) + \varepsilon E[\langle \nabla d_K^2(y), F(t_\varepsilon, y, z)\rangle|\mathscr{F}_{t^\star}] + \varepsilon\gamma_\varepsilon^1 \ ,$$

where

$$\gamma_\varepsilon^1 := \frac{\varepsilon}{2}E[\langle D^2 d_K^2(y)F(t_\varepsilon, y, z), F(t_\varepsilon, y, z)\rangle + \frac{1}{\varepsilon}\alpha(\varepsilon F(t_\varepsilon, y, z))|\mathscr{F}_{t^\star}] \ .$$

Since F is bounded in t, there is some constant $C > 0$ which depends only on y, z such that

$$|\gamma_\varepsilon^1| \leq C\varepsilon + \frac{1}{\varepsilon}\sup_{\|x\|\leq C}\alpha(\varepsilon x), \ \ \varepsilon \in \ (0, t - t^\star) \ .$$

Then in particular $\gamma_\varepsilon^1 \longrightarrow 0$, P-almost everywhere, as ε tends to 0.

We now substitute $x = \xi - y = z(W_t - W_{t_\varepsilon})$ in the definition of $\alpha(\cdot)$. This gives

$$\begin{aligned} E[d_K^2(\xi) - d_K^2(y)|\mathscr{F}_{t^\star}] \ &= \frac{1}{2}E[\langle D^2 d_K^2(y)z(W_t - W_{t_\varepsilon}), z(W_t - W_{t_\varepsilon})\rangle|\mathscr{F}_{t^\star}] \\ &\quad + E[\alpha(z(Wt - W_{t_\varepsilon}))|\mathscr{F}_{t^\star}] \\ &= \frac{1}{2}\varepsilon E[\langle D^2 d_K^2(y)z, z\rangle|\mathscr{F}_{t^\star}] + \varepsilon\gamma_\varepsilon^2 \end{aligned}$$

where

$$\gamma_\varepsilon^2 = \frac{1}{\varepsilon}E(\alpha(z(W_t - W_{t_\varepsilon})|\mathscr{F}_{t^\star}) = \frac{1}{\varepsilon}E[\alpha(\sqrt{\varepsilon}zW_1)]$$

is such that

$$\limsup_{\varepsilon\longrightarrow 0}\gamma_\varepsilon^2 \leq 0 \ .$$

In fact, on one hand, $\frac{1}{\varepsilon}\alpha(\sqrt{\varepsilon}zW_1) \longrightarrow 0$, P-almost everywhere, as $\varepsilon \longrightarrow 0$, and on the other hand, for $\varepsilon > 0$,

$$\frac{1}{\varepsilon}\alpha(\sqrt{\varepsilon}zW_1) \leq (1+\|D^2d_K^2(y)\|)\|z\|^2\|W_1\|^2 \in L^1(P) \ .$$

Finally, we get from Fatou's Lemma

$$\limsup_{\varepsilon\longrightarrow 0}\gamma_\varepsilon^2 \leq E[\limsup_{\varepsilon\longrightarrow 0}\frac{1}{\varepsilon}\alpha(\sqrt{\varepsilon}zW_1)|\mathscr{F}_{t^\star}] = 0, \ P\text{-a.e.}$$

Therefore

$$\begin{cases} E[d_K^2(\widehat{Y_{t_\varepsilon}}) - d_K^2(\xi)|\mathscr{F}_{t^\star}] \geq \\ \varepsilon\{E[\langle\nabla d_K^2(y), F(t,y,z)\rangle|\mathscr{F}_{t^\star}] - \frac{1}{2}\langle D^2[d_K^2(y)]z,z\rangle + \gamma_\varepsilon^1 - \gamma_\varepsilon^2\} \ , \end{cases}$$

and the proof of Lemma 4.3 is complete by setting $\gamma_\varepsilon := \gamma_\varepsilon^1 - \gamma_\varepsilon^2$.

Q.E.D.

Note that due to the Lemmata 4.2 and 4.3:

$$\begin{aligned} &E[d_K^2(Y_{t_\varepsilon}) - d_K^2(\xi)|\mathscr{F}_{t^\star}] \\ &\quad \geq \varepsilon\{\langle\nabla d_K^2(y), F(t,y,z)\rangle - \frac{1}{2}\langle D^2[d_K^2(y)]z,z\rangle + \gamma_\varepsilon\} \end{aligned}$$

for some $\gamma_\varepsilon = \gamma_\varepsilon(y,z) \in \mathbb{R}$ such that $\liminf_{\varepsilon\to 0}\gamma_\varepsilon \geq 0$, P-almost everywhere.

Let us now return to the proof of the necessity. For this denote by $(\widetilde{Y},\widetilde{Z})$ the unique solution to the following BSDE

$$\widetilde{Y}_s = \eta + \int_s^t F(r,\widetilde{Y}_r,\widetilde{Z}_r)dr - \int_s^t \widetilde{Z}_r\, dW_r, \quad t_\varepsilon \leq s \leq t \ ,$$

where $\eta \in L^2(\Omega,\mathscr{F},P)$ is a measurable selection of the set

$$\{(\omega,x) \in \Omega\times\mathbb{R}^N \mid x \in \Pi_K(\xi(\omega))\} \in \mathscr{F}_t\otimes\mathscr{B}(\mathbb{R}^N) \ .$$

We assume that K enjoys the BSVP. Hence $\widetilde{Y}_s \in K$, for $t_\varepsilon \leq s \leq t$, P-almost everywhere. This implies

$$\begin{aligned} 0 &\geq d_K^2(Y_{t_\varepsilon}) - \|Y_{t_\varepsilon} - \widetilde{Y_{t_\varepsilon}}\|^2 \\ &= E[d_K^2(Y_{t_\varepsilon}) - d_K^2(\xi)|\mathscr{F}_{t^\star}] - E[\|Y_{t_\varepsilon} - \widetilde{Y_{t_\varepsilon}}\|^2 - \|\xi-\eta\|^2|\mathscr{F}_{t^\star}] \ . \end{aligned}$$

From Itô's formula,

$$\begin{aligned} &E[\|Y_{t_\varepsilon} - \widetilde{Y_{t_\varepsilon}}\|^2 - \|\xi-\eta\|^2|\mathscr{F}_{t^\star}] \\ &\quad = 2E[\int_{t_\varepsilon}^t \langle Y_s - \widetilde{Y}_s, F(s,Y_s,Z_s) - F(s,\widetilde{Y}_s,\widetilde{Z}_s)\rangle ds|\mathscr{F}_{t^\star}] \\ &\qquad - E[\int_{t_\varepsilon}^t \|Z_s - \widetilde{Z}_s\|^2 ds|\mathscr{F}_{t^\star}] \leq L\int_{t_\varepsilon}^t E[\|Y_s - \widetilde{Y}_s\|^2|\mathscr{F}_{t^\star}]ds \ , \end{aligned}$$

where $L > 0$ is some real constant which depends only on F. Consequently

$$
\begin{aligned}
0 &\geq \frac{1}{\varepsilon}(d_K^2(Y_{t_\varepsilon}) - \|Y_{t_\varepsilon} - \widetilde{Y_{t_\varepsilon}}\|^2) \\
&\geq \frac{1}{\varepsilon}E[d_K^2(Y_{t_\varepsilon}) - d_K^2(\xi)|\mathscr{F}_{t^\star}] - \frac{L}{\varepsilon}\int_{t_\varepsilon}^{t} E[\|Y_r - \widetilde{Y}_r\|^2|\mathscr{F}_{t^\star}]dr \\
&\geq E[\langle \nabla d_K^2(y), F(t,y,z)\rangle|\mathscr{F}_{t^\star}] - \frac{1}{2}\langle D^2 d_K^2(y)z, z\rangle \\
&\quad + \gamma_\varepsilon - L\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t} E[\|Y_r - \widetilde{Y}_r\|^2|\mathscr{F}_{t^\star}]dr \ .
\end{aligned}
$$

In order to estimate the integral term in the last estimate, note that

$$
\begin{aligned}
E[\|Y_s - \widetilde{Y}_s\|^2|\mathscr{F}_{t^\star}] &= E[\|(Y_s-\xi) - (\widetilde{Y}_s - \eta) + (\xi - \eta)\|^2|\mathscr{F}_{t^\star}] \\
&\leq E[d_K^2(\xi)|\mathscr{F}_{t^\star}] + 2(E[\|Y_s - \xi\|^2|\mathscr{F}_{t^\star}] + E[\|\widetilde{Y}_s - \eta\|^2|\mathscr{F}_{t^\star}] \\
&\quad + 2(E[d_K^2(\xi)|\mathscr{F}_{t^\star}])^{\frac{1}{2}}((E[\|Y_s - \xi\|^2|\mathscr{F}_{t^\star}])^{\frac{1}{2}} + (E[\|\widetilde{Y}_s - \eta\|^2|\mathscr{F}_{t^\star}])^{\frac{1}{2}}) \ .
\end{aligned}
$$

From (17) we conclude

$$
\sup_{t_\varepsilon \leq s \leq t} E[\|Y_s - \xi\|^2 + \|\widetilde{Y}_s - \eta\|^2|\mathscr{F}_{t^\star}] \longrightarrow 0
$$

as ε tends to 0. On the other hand, the function d_K^2 is continuous at y and d_K^2 is of at most quadratic growth. Hence, according the Lebesgue theorem of dominated convergence,

$$
E[d_K^2(\xi)|\mathscr{F}_{t^\star}] \longrightarrow d_K^2(y),\ \ P\text{-almost everywhere} \ ,
$$

as $\varepsilon \longrightarrow 0$.

Consequently,

$$
E[\|Y_s - \widetilde{Y}_s\|^2|\mathscr{F}_{t^\star}] \leq d_K^2(y) + \beta_\varepsilon^2,\ \ t_\varepsilon \leq s \leq t,\ \ \varepsilon > 0 \ ,
$$

where β_ε^2 converges to 0, P-almost everywhere, as ε tends to 0.

Therefore

$$
L\frac{1}{\varepsilon}\int_{t_\varepsilon}^{t} E[\|Y_r - \widetilde{Y}_r\|^2|\mathscr{F}_{t^\star}]dr \leq L(d_K^2(y) + \beta_\varepsilon^2)
$$

and

$$
\begin{aligned}
&E[\langle \nabla d_K^2(y), F(t,y,z)\rangle|\mathscr{F}_{t^\star}] - \frac{1}{2}\langle D^2 d_K^2(y)z, z\rangle \\
&\quad - L d_K^2(y) + (\gamma_\varepsilon - L\beta_\varepsilon^2) \leq 0,\ \ P\text{-a.e.}
\end{aligned}
$$

Finally, since $\liminf_{\varepsilon \longrightarrow 0}(\gamma_\varepsilon + L\beta_\varepsilon^2) \geq 0$, P-a.e., we get

$$
E[\langle \nabla d_K^2(y), F(t,y,z)\rangle|\mathscr{F}_{t^\star}] \leq \frac{1}{2}\langle D^2 d_K^2(y)z, z\rangle + L d_K^2(y) \ ,
$$

P-almost everywhere, for $t^\star \in [0, T[$. Passing to the limit $t^\star \longrightarrow t$, we obtain the wished result.

(b) Sufficiency. Let K be a nonempty convex closed subset of R^N. Suppose that (8) holds true. Note that (cf [9]) if $d_K^2 \in C^{1,1}$, then Π_K is single-valued and

$$\begin{cases} \nabla d_K^2(y) = 2(y - \Pi_K(y)), \ \forall\, y \in \mathbb{R}^N \\ \|\Pi_K(y) - \Pi_K(x+y)\| \le \|x\|, \forall (x,y) \in \mathbb{R}^N \times \mathbb{R}^N \ . \end{cases}$$

Recall that the measurable mapping $D^2(d_K^2) : \Lambda_K \mapsto S(\mathbb{R}^N)$ is defined by the second order development of d_K^2 in $y \in \Lambda_K$:

$$d_K^2(x+y) = d_K^2(y) + \langle \nabla d_K^2(y), x\rangle + \frac{1}{2}\langle D^2(d_K^2(y))x, x\rangle + \alpha(y,x)$$

with

$$\frac{1}{\|x\|^2}\alpha(y,x) \longrightarrow 0, \ \text{as } x \longrightarrow 0 \ .$$

We claim that

$$\begin{cases} i) \ \ 0 \le \frac{1}{2} D^2[d_K^2(y)] \le I, \ \forall\, y \in \Lambda_K, \\ ii) \ |\alpha(y,x)| \le \|x\|^2, \ \forall (x,y) \in \mathbb{R}^N \times \Lambda_K \ . \end{cases} \tag{30}$$

For proving this, we fix $y \in \Lambda_K$. On one hand, since d_K^2 is convex we have

$$d_K^2(x+y) - d_K^2(y) - \langle \nabla d_K^2(y), x\rangle \ge 0, \ \forall\, x \in \mathbb{R}^N \ ,$$

on the other hand

$$\begin{aligned} d_K^2(x+y) - d_K^2(y) - \langle \nabla d_K^2(y), x\rangle \le \ & \|(y+x) - \Pi_K(y)\|^2 - \|y - \Pi_K(y)\|^2 \\ & -2\langle x, y - \Pi_K(y)\rangle = \|x\|^2 \ . \end{aligned}$$

Thus, by substituting $x = te$, $t > 0$, where e denotes an arbitrary unit vector of $\mathbb{R}^N$, we obtain from the previous inequalities and the definition of α, that

$$\begin{aligned} -\frac{1}{t^2}\alpha(y,te) &\le \frac{1}{t^2}(d_K^2(y+te) - d_K^2(y) - t\langle \nabla d_K^2(y), e\rangle) - \frac{\alpha(y,te)}{t^2} \\ &= \frac{1}{2}\langle D^2[d_K^2(y)]e, e\rangle \le 1 - \frac{1}{t^2}\alpha(y,te) \end{aligned}$$

Passing to the limit $t \longrightarrow 0^+$, this gives (30-i). From (30-i), one can easily deduce (30-ii).

Let $\eta \in C^\infty(\mathbb{R}^N)$ be a nonnegative function with support in the unit ball and such that $\int_{R^N} \eta(x)dx = 1$.

For $\delta > 0$, we put

$$\begin{aligned} \eta_\delta(x) &:= \frac{1}{\delta^N}\eta(\frac{1}{\delta}x) \\ \phi_\delta(x) &:= d_K^2 \star \eta_\delta(x) := \int_{R^N} d_K^2(x - x')\eta_\delta(x')dx', \quad x \in \mathbb{R}^N \ . \end{aligned}$$

Obviously, $\phi_\delta \in C^\infty(\mathbb{R}^N)$. The function ϕ_δ satisfies the following properties

$$\begin{cases} i) & 0 \le \phi_\delta(x) \le (d_K(x)+\delta)^2, \\ ii) & \nabla\phi_\delta(x) = \int_{R^N}(\nabla d_K^2)(x')\eta_\delta(x-x')dx', \\ & \|\nabla\phi_\delta(x)\| \le 2(d_K(x)+\delta), \\ iii) & D^2\phi_\delta(x) = \int_{R^N} D^2[d_K^2](x')\eta_\delta(x-x')dx', \\ & 0 \le D^2\phi_\delta(x) \le 2I \ , \end{cases} \tag{31}$$

for all $x \in \mathbb{R}^N$. The property (31-i) is clear, let us focus on (31-ii)-iii). Since Λ_K is of full measure, we have for any x, x' belonging to $\mathbb{R}^N$,

$$\begin{aligned} \phi_\delta(x'+x) - \phi_\delta(x') &= \int_{R^N}\{d_K^2(y+x) - d_K^2(y)\}\eta_\delta(x'-y)dy \\ &= \langle \int_{R^N} \nabla d_K^2(y)\eta_\delta(x'-y)dy, x\rangle \\ &\quad + \frac{1}{2}\langle(\int_{R^N} D^2[d_K^2(y)]\eta_\delta(x'-y)dy)x, x\rangle + \varepsilon(x', x) \end{aligned}$$

where $\varepsilon(x', x) := \int_{R^N} \alpha(y, x)\eta_\delta(x'-y)dy$. Obviously, from Lebesgue's dominated convergence Theorem (cf (30-ii)),

$$\frac{\varepsilon(x', x)}{\|x\|^2} = \int_{R^N}(\frac{\alpha(y, x)}{\|x\|^2})\eta_\delta(x'-y)dy \longrightarrow 0, \quad \text{as } \|x\| \longrightarrow 0, \ \forall\, x' \in \mathbb{R}^N \ .$$

Consider $\xi \in L^2(\Omega, \mathscr{F}_T, P, K)$ and let (Y, Z) be the unique solution to the following BSDE

$$Y_t = \xi + \int_t^T F(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad t \in [0, T] \ .$$

Relation (31) enables us to apply Itô's formula to $\phi_\delta(Y_t)$ and to deduce that, for $0 \le t \le T, \delta > 0$,

$$\begin{aligned} E\phi_\delta(Y_t) = E\phi_\delta(\xi) &+ E\int_t^T \langle(\nabla\phi_\delta)(Y_s), F(s, Y_s, Z_s)\rangle ds \\ &- \frac{1}{2}E\int_t^T \langle(D^2\phi_\delta)(Y_s)Z_s, Z_s\rangle ds \\ \le \delta^2 + E\int_t^T &\int_{R^N}\{\langle\nabla d_K^2(y), F(s, y, Z_s)\rangle \\ &- \frac{1}{2}\langle D^2(d_K^2(y))Z_s, Z_s\rangle\}\eta_\delta(Y_s - y)dy\, ds \\ &- E\int_t^T \int_{R^N}\{\langle\nabla d_K^2(y), F(s, y, Z_s) - F(s, Y_s, Z_s)\rangle\eta_\delta(Y_s - y)\}dy\, ds \end{aligned}$$

Then from (8) and (31), for $\delta \in [0,1]$,

$$\begin{aligned} &E\phi_\delta(Y_t) \\ &\quad \le \delta^2 + CE\int_t^T \int_{R^N} d_K^2(y)\eta_\delta(Y_s - y)dy\, ds \\ &\qquad + E\int_t^T\!\!\int_{R^N} 2d_K(y) \max_{y:\|y-Y_s\|\le\delta} \|F(s,y,Z_s) - F(s,Y_s,Z_s\|\eta_\delta(Y_s-y)dy\, ds \\ &\quad \le \delta^2 + C\int_t^T E[\phi_\delta(Y_s)]ds \\ &\qquad + E\int_t^T (1+\phi_\delta(Y_s)) \max_{y:\|y-Y_s\|\le\delta} \|F(s,y,Z_s) - F(s,Y_s,Z_s\|ds \ . \end{aligned}$$

Taking into account that the function F is uniformely continuous in its second variable, uniformely with respect to the other ones, we obtain that for some continuous increasing function $g : \mathbb{R}_+ \mapsto \mathbb{R}_+$ with $g(0) = 0$,

$$\begin{cases} E\phi_\delta(Y_t) \le \delta^2 + g(\delta) \\ +(C+1)\int_t^T E[\phi_\delta(Y_s)]ds \ , \end{cases} \tag{32}$$

for $0 \le t \le T$, $\delta > 0$ small enough. On the other hand, from (31-ii), we deduce

$$E[\phi_\delta(Y_t)] < +\infty \ .$$

This allows to apply Gronwall's inequality to (32). So there is a real $C > 0$ which does not depend on $t \in [0,T]$, $\delta \in]0,1[$, such that

$$E\phi_\delta(Y_t) \le C(\delta^2 + g(\delta)), \quad 0 \le t \le T \ .$$

Finally, since F is bounded, from Fatou's Lemmma and the dominated convergence Theorem, we conclude that

$$Ed_K^2(Y_t) \le \liminf_{\delta \longrightarrow 0} E\phi_\delta(Y_t) = 0 \quad \text{for } 0 \le t \le T \ ,$$

i.e., $Y_t \in K$, for any $t \in [0,T]$, P-almost everywhere. Q.E.D.

References

1. Alexandrov, A.D.: The existence almost everywhere of the second differential of a convex function and some associated properties of convex surfaces. (in Russian), Ucenye Zapiski Leningrad. Gos. Univ. Ser. Math. **37**(6), 3–35 (1939)
2. Aubin, J.-P.: Viability Theory, Birkhäuser (1992)
3. Aubin, J.-P., Da Prato, G.: Stochastic Viability and Invariance, Annali Scuola Normale di Pisa, N. 27, pp. 595–694 (1990)
4. Aubin, J.-P., De Prato, G.: Stochastic Nagumo's Viability Theorem, Stochastic Analysis and Applications, N. 13, pp. 1–11 (1995)
5. Aubin, J.-P., Da Prato, G.: The viability theorem for stochastic differential inclusions, Stochastic Anal. Appl. **16**(1), 1–15 (1998)

6. Barles, G., Buckdahn, R., Pardoux, E.: Backward stochastic differential equations and integral-partial differential equations, Stochastics Stochastics Rep. **60** (1–2), 57–83 (1997)
7. Crandall, M.G., Evans, L.C., Lions, P.L.: Some properties of viscosity solutions of Hamilton-Jacobi Equations, Transactions of A.M.S. **282**, 487–502 (1984)
8. Crandall, M., Ishii, H., Lions, P.L.: User's Guide to the Viscosity Solutions of Second Order Partial Differential Equations, Bull. AMS, **27**, 1–67 (1992)
9. Delfour, M., Zolezio, J.P.: Oriented distance functions in shape analysis and optimization. Lagnese, John E. (ed.) et al., Control and optimal design of distributed parameter systems. Proceedings of a 1992 workshop, held at IMA, Minneapolis, MN, USA. New York, NY: Springer-Verlag, IMA Vol. Math. Appl. **70**, 39–72 (1995)
10. Nagumo, M.: Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen, Prc. Phys. Math. Soc. Japan **24**, 551–559 (1942)
11. Gegout-Petit, A., Pardoux, E.: Equations différentielles stochastiques rétrogrades réfléchies dans un convexe, Stochastics **57**, 11–128 (1995)
12. Pardoux, E.: Backward Stochastic Differential Equations and Semilinear Partial Differential Equations, Summer School, Geilo (1996)
13. Pardoux, E., Peng, S.: Adapted solution of a backward stochastic differential equation, Syst., Cont. Letters, N. 14, pp. 55–61 (1990)
14. Pardoux, E., Peng, S.: Backward Stochastic differential equations and quasilinear parabolic partial differential equations. In: Stochastic partial differential equations and their applications, (B.L. Rotovskii, R.B. Sowers, Eds.), Lecture Notes Control Inf. Sci. N. 176, pp. 200–217, Springer, Berlin, Heidelberg, New York (1992)
15. Peng, S.: Probabilistic Interpretation for systems of Quasilinear Parabolic Partial Differential Equations, Stochastics and Stochastics Report **37**, 61–74 (1991)
16. Peng, S.: A generalized dynamic programming principle and Hamilton-Jacobi-Bellman Equation, Stochastics and Stochastics Report **38**(2), 119–134 (1992)
17. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. Springer-Verlag (1991)