Hamilton-Jacobi-Bellman equations in infinite dimensions

Marco Fuhrman

Politecnico di Milano

Fausto Gozzi

LUISS University, Rome

Spring School "Stochastic Control in Finance"
March 7-18 2010, Roscoff

Preliminary results and general framework.

1. A summary of stochastic integration in Hilbert spaces
2. Deterministic and stochastic evolution equations
3. Basic examples: the heat equation, delay equations
4. The optimal control problem
5. The Hamilton-Jacobi-Bellman equation and a verification theorem

Spaces and operators

H, K denote Hilbert spaces. All Hilbert spaces are assumed to be real separable. Scalar product is denoted $\langle \cdot, \cdot \rangle$.

$L(K, H)$ is the space of linear bounded operators $T : K \to H$. $L(H) := L(H, H)$.

$L_2(K, H)$ is the subspace of Hilbert-Schmidt operators, i.e. of all $T \in L(K, H)$ such that

$$\|T\|^2_{L_2(K,H)} = \sum_{i=1}^{\infty} \|Te_i\|^2_H < \infty,$$

where (e_i) is an arbitrary basis of K (i.e. a complete orthonormal system). $L_2(K, H)$ is a separable Hilbert space with scalar product

$$\langle T, S \rangle_{L_2(K,H)} = \sum_{i=1}^{\infty} \langle Te_i, Se_i \rangle_H = \sum_{i=1}^{\infty} \langle S^* Te_i, e_i \rangle_H = \text{Trace}[S^* T].$$

Other notation:

- $L^2(\Omega; K)$ is the Hilbert space of random variables $X : \Omega \to K$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$\|X\|^2_{L^2(\Omega;K)} = \mathbb{E} \, \|X\|^2_K < \infty.$$

- $C(I; K)$, for $I \subset \mathbb{R}$, is the space of continuous functions $f : I \to K$.
- etc.

Wiener process in Hilbert spaces

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ take a sequence of independent standard brownian motions

$$\beta^i = (\beta^i_t)_{t \geq 0}, \qquad i \in \mathbb{N}.$$

Given a basis of K, a cylindrical Wiener process $(W_t)_{t\geq 0}$ in K is defined as

$$W_t = \sum_{i=1}^{\infty} \beta^i_t \, e_i, \qquad t \geq 0.$$

The series is convergent (in $L^2(\Omega; K_1)$ and $\mathbb{P}$-a.s.) in an arbitrary Hilbert space K_1 such that $K \subset K_1$ with Hilbert-Schmidt embedding.

For $G \in L(K)$ define

$$GW_t = \sum_{i=1}^{\infty} \beta^i_t \, Ge_i, \qquad t \geq 0.$$

Then the series converges (in $L^2(\Omega; K)$ and $\mathbb{P}$-a.s.) if and only if $G \in L_2(K)$.

Suppose $Ge_i = \sqrt{\lambda_i} e_i$ for a basis (e_i) and numbers $0 \leq \lambda_i \leq \sup_i \lambda_i < \infty$. Then

$$GW_t = \sum_{i=1}^{\infty} \beta^i_t \sqrt{\lambda_i} \, e_i, \qquad t \geq 0$$

and the series converges as above if and only if $\sum_i \lambda_i < \infty \iff G \in L_2(K)$. This happens in particular if $\lambda_i = 0$ for all i large (finite-dimensional Wiener process).

Stochastic integrals in Hilbert spaces

Let (W_t) be a cylindrical Wiener process in K. One can define the stochastic integral

$$I_t = \int_0^t \Phi_s \, dW_s, \qquad t \in [0, T]$$

as a process in H, under the following conditions:

1) $(\Phi_t)_{t\in[0,T]}$ is a stochastic process in $L_2(K, H)$ $(\Phi_t(\omega) \in L_2(K, H))$.

2) (Φ_t) is progressive with respect to $(\mathcal{F}_t)_{t\geq 0}$, the natural completed filtration of W: denoting by $\mathcal{N}$ the $\mathbb{P}$-null sets,

$$\mathcal{F}_t^0 = \sigma(\beta_s^i \, : \, s \in [0, t], i \in \mathbb{N}), \qquad \mathcal{F}_t = \sigma(\mathcal{F}_t^0, \mathcal{N}).$$

3) $\int_0^T \|\Phi_s\|^2_{L_2(K,H)} \, ds < \infty$, $\mathbb{P}$-a.s.

Then $(I_t)_{t\in[0,T]}$ is a stochastic process in H with continuous paths and it is a local martingale: there exist $(\mathcal{F}_t)$ stopping times $\tau_n \uparrow \infty$ such that the stopped processes

$$\Phi_{t\wedge\tau_n}, \qquad t \in [0, T]$$

are martingales in H.

If (Φ_t) is an $(\mathcal{F}_t)$-progressive process in $L_2(K,H)$ satisfying the stronger condition

$$\mathbb{E}\int_0^T \|\Phi_s\|^2_{L_2(K,H)}\,ds < \infty, \tag{1}$$

then $(I_t)_{t\in[0,T]}$ is a mean-zero, continuous martingale in H and the Ito isometry holds:

$$\mathbb{E}\left\|\int_0^T \Phi_s\,dW_s\right\|^2_H = \mathbb{E}\int_0^T \|\Phi_s\|^2_{L_2(K,H)}\,ds.$$

Three basic tools: representation theorem, Ito's formula, Girsanov's theorem.

• Representation theorem: if (Ψ_t) is an $(\mathcal{F}_t)$-martingale in H and $\Psi_T \in L^2(\Omega;H)$ then

$$\Psi_t = \Psi_0 + \int_0^t \Phi_s\,dW_s, \qquad t\in[0,T],$$

for an appropriate $(\mathcal{F}_t)$-progressive process (Φ_t) with values in $L_2(K,H)$ satisfying (1).

• The Ito formula will be recalled later, when needed, in appropriate form.

• The Girsanov theorem: let $(u_t)_{t\in[0,T]}$ be an $(\mathcal{F}_t)$-progressive process in K satisfying $\int_0^T \|u_s\|_K^2\,ds < \infty$, $\mathbb{P}$-a.s. Define

$$\rho_t = \exp\left(\int_0^t u_s^*\,dW_s - \frac{1}{2}\int_0^t \|u_s\|_K^2\,ds\right), \qquad t\in[0,T],$$

where $u_s^*(\omega)$ denotes $k \mapsto \langle u_s(\omega), k\rangle_U$, belonging to $L_2(K,\mathbb{R})$.

If $\mathbb{E}\,\rho_T = 1$ then $(\rho_t)_{t\in[0,T]}$ is a martingale and the process

$$\bar{W}_t = W_t - \int_0^t u_s\,ds, \qquad t\in[0,T],$$

is a cylindrical Wiener process in K with respect to the probability $\mathbb{Q}$ defined on $(\Omega,\mathcal{F})$ by the formula $\mathbb{Q}(d\omega) = \rho_T(\omega)\,\mathbb{P}(d\omega)$.

Evolution equations

The "obvious" generalization for an SDE in H would be:

$$dX_t = F(X_t)\,dt + G(X_t)\,dW_t, \qquad X_0 = x \in H,$$

for an unknown process $(X_t)_{t\in[0,T]}$ in H.
W is a cylindrical Wiener process in K, the equation is understood as

$$X_t = x + \int_0^t F(X_s)\,ds + \int_0^t G(X_s)\,dW_s, \qquad t \in [0,T],$$

and

$$F : H \to H, \qquad G : H \to L_2(K,H),$$

are appropriate coefficients.

However, the useful form of the equation is

$$dX_t = AX_t\,dt + F(X_t)\,dt + G(X_t)\,dW_t, \qquad X_0 = x \in H,$$

where A is a linear unbounded operator in H:

$$A : D(A) \to H, \qquad D(A) \subset H.$$

So we will first address the simpler equation

$$dX_t = AX_t, \qquad X_0 = x \in H,$$

i.e. the deterministic abstract Cauchy problem

$$\frac{d}{dt}y(t) = Ay(t), \qquad y(0) = x \in H,$$

with unknown $y : [0,T] \to H$.

Semigroups of operators

$$\frac{d}{dt}y(t) = Ay(t), \qquad y(0) = x \in H,\ t \in [0,T].$$

If $A \in L(H)$ then the solution is given by the power series formula

$$y(t) = e^{tA}x = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n x.$$

Setting $S(t) = e^{tA}$ one has, for $t, s \geq 0$, $x \in H$,

$$S(0) = I, \qquad S(t+s) = S(t)S(s), \qquad S(t)x \to x \text{ in } H \text{ as } t \to 0. \tag{2}$$

Note that, for all $x \in H$,

$$Ax = \lim_{t \to 0} \frac{S(t)x - x}{t} \quad \text{in } H. \tag{3}$$

Definition. $(S(t))_{t \geq 0} \subset L(H)$ is called a strongly continuous semigroup of linear bounded operators on H if (2) holds.
Its infinitesimal generator is the operator A given by (3) and defined on

$$D(A) = \{x \in H \ :\ \text{the limit (3) exists in } H\}.$$

Note that in general $D(A) \subsetneq H$ and A is not H-continuous on $D(A)$.

Standing notation: from now on, A denotes the generator of a semigroup S and we use the exponential notation e^{tA} instead of $S(t)$ (even if the power series formula does not hold).

Deterministic evolution equations

Given a generator A, we ask whether the function

$$y(t) = e^{tA}x$$

is a solution to the homogeneous abstract Cauchy problem

$$\frac{d}{dt}y(t) = Ay(t), \qquad y(0) = x \in H,\ t \in [0, T].$$

One can prove that if $x \in D(A)$ then y is a strict solution, i.e. $y \in C^1([0, T]; H)$, $y(t) \in D(A)$ and the equation holds. The strict solution is unique.
If only $x \in H$ then $y \in C([0, T]; H)$ and it is called mild solution.

Given a generator A and $f : [0, T] \to H$, we consider the nonhomogeneous abstract Cauchy problem

$$\frac{d}{dt}y(t) = Ay(t) + f(t), \qquad y(0) = x \in H,\ t \in [0, T].$$

One can prove that if

$$x \in D(A), \qquad f \in C^1([0, T]; H),$$

then there exists a unique strict solution given by the variation of costants formula

$$y(t) = e^{tA}x + \int_0^t e^{(t-s)A} f(s)\, ds,$$

i.e. $y \in C^1([0, T]; H)$, $y(t) \in D(A)$ and the equation holds.
If only $x \in H$ and $f \in C([0, T]; H)$ (or even $f \in L^1([0, T]; H)$) then $y \in C([0, T]; H)$ is called mild solution.

Example 1: the heat semigroup

We take $\mathcal{O} \subset \mathbb{R}^d$ open bounded with smooth boundary and set

$$H = L^2(\mathcal{O}).$$

Thus $x \in H$ is a real function $x(\xi)$, $\xi \in \mathcal{O}$ in $L^2(\mathcal{O})$, and $y : [0,T] \to H$ is a real function $y(t,\xi)$, $t \in [0,T], \xi \in \mathcal{O}$ such that $y(t,\cdot) \in L^2(\mathcal{O})$.

We define an unbounded linear operator in $H = L^2(\mathcal{O})$ setting

$$A = \Delta_\xi, \qquad D(A) = H^2(\mathcal{O}) \cap H^1_0(\mathcal{O}),$$

where Δ_ξ is the Laplace operator with respect to the space variable $\xi \in \mathcal{O}$. The equation

$$\frac{d}{dt} y(t) = Ay(t), \qquad y(0) = x \in H,\ t \in [0,T]$$

is an abstract form of the heat equation with homogeneous Dirichlet boundary conditions:

$$\begin{cases} \partial_t y(t,\xi) = \Delta_\xi y(t,\xi), & \xi \in \mathcal{O},\ t \in [0,T], \\ y(0,\xi) = x(\xi), & \xi \in \mathcal{O}, \\ y(t,\xi) = 0, & \xi \in \partial\mathcal{O},\ t \in [0,T]. \end{cases}$$

A is a positive self-adjoint operator in H and generate a semigroup. There exists a basis (e_i) of H and numbers $0 < \alpha_i \uparrow \infty$ such that

$$Ae_i = -\alpha_i e_i,$$

and we have $e^{tA}x = \sum_i e^{-\alpha_i t} \langle x, e_i \rangle_H \, e_i$ for all $x \in H,\ t \geq 0$.

Stochastic evolution equations: a special case

Consider the linear equation with additive noise (Langevin equation)

$$dX_t = AX_t\,dt + G\,dW_t, \qquad X_0 = x \in H,\ t \in [0,T],$$

where A is a generator in H, W is a cylindrical Wiener process in K, $G \in L(K,H)$. We define the mild solution as

$$X_t = e^{tA}x + \int_0^t e^{(t-s)A}G\,dW_s,$$

provided the (deterministic) integrand $\Phi_s = e^{(t-s)A}G$ satisfies ($\mathbb{P}$-a.s.)

$$\int_0^T \|\Phi_s\|^2_{L_2(K,H)}\,ds = \int_0^T \|e^{sA}G\|^2_{L_2(K,H)}\,ds < \infty.$$

X is called the Ornstein-Uhlenbeck process in H.
One proves that if there exists $\gamma \in [0, 1/2)$ and $K > 0$ such that

$$\|e^{tA}G\|_{L_2(K,H)} \le Kt^{-\gamma}, \qquad t \in (0,T].$$

then X has continuous paths in H, $\mathbb{P}$-a.s.
This condition always holds (with $\gamma = 0$) if $G \in L_2(K,H)$.

Stochastic evolution equations: general case

Given $F : H \to H$, $G : H \to L(K,H)$ and a generator A consider

$$dX_t = AX_t\,dt + F(X_t)\,dt + G(X_t)\,dW_t, \qquad X_0 = x \in H,\ t \in [0,T].$$

We call X a mild solution if it is an $(\mathcal{F}_t)$-adapted process in H with continuous paths, satisfying $\mathbb{P}$-a.s.

$$X_t = e^{tA}x + \int_0^t e^{(t-s)A}F(X_s)\,ds + \int_0^t e^{(t-s)A}G(X_s)\,dW_s,\ t \in [0,T].$$

Standing assumptions:

- A is a generator in H, W is a cylindrical Wiener process in K.
- $F : H \to H$ and $G : H \to L(K,H)$ satisfy, for all $t \in (0,T]$, $x, y \in H$,

$$\|e^{tA}(G(x) - G(y))\|_{L_2(K,H)} \le Lt^{-\gamma}\|x - y\|_H,\ \|e^{tA}G(x)\|_{L_2(K,H)} \le Kt^{-\gamma}(1 + \|x\|_H),$$

$$\|F(x) - F(y)\|_H \le L\|x - y\|_H, \qquad \|F(x)\|_H \le K(1 + \|x\|_H),$$

for some $\gamma \in [0, 1/2)$ and $K, L > 0$.

One can prove that there exists a unique mild solution satisfying, for every $p \in [1,\infty)$,

$$\mathbb{E}\sup_{t\in[0,T]} \|X_t\|_H^p \le C(1 + \|x\|)^p,$$

with $C = C(p, L, K, \gamma, A, T)$.

Generalizations to time-dependent or stochastic coefficients are possible.

Example 1 revisited: the stochastic heat equation

We take $\mathcal{O} \subset \mathbb{R}^d$ as before and $H = L^2(\mathcal{O})$. Thus a process (X_t) in H is in particular a real process $\{X_t(\omega, \xi)\,;\, \omega \in \Omega, t \in [0, T], \xi \in \mathcal{O}\}$ depending on a spatial parameter ξ.

Define the Laplace operator $A = \Delta_\xi$, $D(A) = H^2(\mathcal{O}) \cap H_0^1(\mathcal{O})$ and let W be a cylindrical Wiener process in $K = L^2(\mathcal{O})$. Then the equation

$$dX_t = AX_t\,dt + G\,dW_t, \qquad X_0 = x \in H,\ t \in [0, T],$$

is an abstract form of the stochastic heat equation

$$\begin{cases} dX_t(\xi) = \Delta_\xi X_t(\xi)\,dt + G\,dW_t(\xi), & \xi \in \mathcal{O},\ t \in [0, T], \\ X_0(\xi) = x(\xi), & \xi \in \mathcal{O}, \\ X_t(\xi) = 0, & \xi \in \partial\mathcal{O},\ t \in [0, T]. \end{cases}$$

Warning: $W_t(\omega, \cdot) \notin L^2(\mathcal{O})$, but $W_t(\omega, \cdot) \in H^{-m}(\mathcal{O})$ for suitable $m \geq 0$.

It remains to define $G \in L(K, H) = L(L^2(\mathcal{O}))$ satisfying

$$\|e^{tA} G\|_{L_2(K,H)} \leq K t^{-\gamma}, \qquad t \in (0, T],$$

for some $\gamma \in [0, 1/2)$ and $K > 0$, and the mild solution given by

$$X_t = e^{tA} x + \int_0^t e^{(t-s)A} G\,dW_s,$$

will be well-defined, and $\mathbb{P}$-a.s. continuous in H.

Recall that $$Ae_i = -\alpha_i e_i,$$
for a basis (e_i) of H and numbers $0 < \alpha_i \uparrow \infty$. Define $G \in L(L^2(\mathcal{O}))$ setting

$$Ge_i = \sqrt{q_i} e_i, \qquad \text{i.e.} \qquad Gx = \sum_i q_i \langle x, e_i \rangle_H \, e_i, \qquad x \in H,$$

where (q_i) is a numerical sequence, $0 \le q_i \le C$ for some $C > 0$.
We finally require

$$\|e^{tA} G\|_{L_2(K,H)} = (\textstyle\sum_i q_i e^{-2\alpha_i t})^{1/2} \le K t^{-\gamma}, \qquad t \in (0,T],$$

for some $\gamma \in [0, 1/2)$ and $K > 0$.
Special case: if $d = 1$ ($\mathcal{O} = (a,b) \subset \mathbb{R}$) then $\alpha_i = Ci^2$ and we can take $q_i \equiv 1$, i.e. $G = I$, obtaining $\gamma = 1/4$ and solving the equation

$$dX_t(\xi) = \partial_{\xi\xi} X_t(\xi)\, dt + dW_t(\xi), \qquad \xi \in (a,b),\ t \in [0,T].$$

driven by the space-time white noise on (a,b).

In the general case $\mathcal{O} \subset \mathbb{R}^d$ we can consider the nonlinear heat equation

$$\begin{cases} dX_t(\xi) = \Delta_\xi X_t(\xi)\, dt + f(X_t(\xi))\, dt + G\, dW_t(\xi), & \xi \in \mathcal{O},\ t \in [0,T], \\ X_0(\xi) = x(\xi), & \xi \in \mathcal{O}, \\ X_t(\xi) = 0, & \xi \in \partial\mathcal{O},\ t \in [0,T]. \end{cases}$$

with $f : \mathbb{R} \to \mathbb{R}$ Lipschitz. We define $F : H \to H$ setting

$$F(x)(\xi) = f(x(\xi)), \qquad \xi \in \mathcal{O},\ x \in H = L^2(\mathcal{O}),$$

and we obtain a unique mild solution of the equation

$$dX_t = AX_t\, dt + F(X_t)\, dt + G\, dW_t, \qquad X_0 = x \in H,\ t \in [0,T].$$

Example 2: delay equations (hereditary systems)

Given a delay $r > 0$, the prototype of a (scalar) delay equation is

$$\frac{d}{dt}z(t) = z(t-r), \qquad t > 0.$$

Note that the initial condition $z(0) = x_0 \in \mathbb{R}$ is not enough to have a well-posed problem. We need to specify a function $x_1(\theta)$, $\theta \in [-r, 0]$, and require

$$z(\theta) = x_1(\theta), \qquad \theta \in [-r, 0].$$

We will assume $x_1 \in C([-r, 0])$ for the moment.
Note that the equation can be written in the form $z'(t) = \int_{-r}^{0} z(t+\theta)\, \delta_{-r}(d\theta)$.

More generally we will consider the problem

$$\begin{cases} \frac{d}{dt}z(t) = \int_{-r}^{0} z(t+\theta)\, a(d\theta), & t > 0, \\ z(0) = x_0, & \\ z(\theta) = x_1(\theta), & \theta \in [-r, 0], \end{cases}$$

where $a(d\theta)$ is a signed finite measure on $[-r, 0]$.

By direct methods one proves that there exists a unique classical solution $z : [-r, \infty) \to \mathbb{R}$ provided $x_1(0) = x_0$. We look for a Hilbert space setting for the equation, useful for dealing with control problems and for a more general initial datum $x_1 \in L^2(-r, 0)$.

$$\begin{cases} \frac{d}{dt}z(t) = \int_{-r}^{0} z(t+\theta)\, a(d\theta), & t > 0, \\ z(0) = x_0, & \\ z(\theta) = x_1(\theta), & \theta \in [-r, 0], \end{cases} \tag{4}$$

Suppose that z is a classical solution and define

$$y(t) := \begin{pmatrix} z(t) \\ z(t+\cdot) \end{pmatrix} \in \mathbb{R} \times L^2(-r, 0) =: H.$$

Then $$\frac{d}{dt}y(t) = \begin{pmatrix} \frac{d}{dt}z(t) \\ \frac{d}{dt}z(t+\cdot) \end{pmatrix} = \begin{pmatrix} \int_{-r}^{0} z(t+\theta)\, a(d\theta) \\ \frac{d}{d\theta}z(t+\cdot) \end{pmatrix} =: Ay(t)$$

provided we define an operator A in $H = \mathbb{R} \times L^2(-r, 0)$ by

$$A \begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix} = \begin{pmatrix} \int_{-r}^{0} x_1(t+\theta)\, a(d\theta) \\ \frac{d}{d\theta}x_1(\cdot) \end{pmatrix},$$

with domain $D(A) = \{(x_0, x_1(\cdot)) \in H \,:\, x_1(\cdot) \in H^1(-r, 0), x_1(0) = x_0\}$.

It can be proved that A is a generator of a semigroup (e^{tA}) in H.
If $x_1 \in C([-r, 0])$ and $x_1(0) = x_0$ then the first component of

$$y(t) := e^{tA} \begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix}$$

is a classical solution to (4).

In the general case $x_0 \in \mathbb{R}$, $x_1 \in L^2(-r, 0)$, y is called mild solution to (4).
Note that x_0 is not determined by $x_1(\cdot)$.

Next we consider the general case of a stochastic nonlinear delay equation

$$\begin{cases} dz(t) = \left(\int_{-r}^{0} z(t+\theta)\, a(d\theta)\right) dt + f(z(t))\, dt + dW_t, & t \in [0,T], \\ z(0) = x_0, & \\ z(\theta) = x_1(\theta), & \theta \in [-r,0], \end{cases} \tag{5}$$

where W is a Wiener process in $K = \mathbb{R}$, $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz.
We recall that $x_0 \in \mathbb{R}, x_1 \in L^2(-r,0)$, $a(d\theta)$ is a signed finite measure on $[-r,0]$.

We will write this equation as an abstract evolution equation of the form

$$dX_t = AX_t\, dt + F(X_t)\, dt + G\, dW_t, \qquad X_0 = x \in H,$$

in the space $H = \mathbb{R} \times L^2(-r,0)$.
We define A as before, $F : H \to H$ and $G : \mathbb{R} \to H$ as

$$F \begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix} = \begin{pmatrix} f(x_0) \\ 0 \end{pmatrix}, \qquad Gu = \begin{pmatrix} u \\ 0 \end{pmatrix},$$

and the initial datum $x = \begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix}$.

The standard assumptions are satified and the mild solution X (or its first component) will be called a mild solution to (5).

Controlled stochastic evolution equations

$$dX_t = AX_t\,dt + F(X_t, u_t)\,dt + G(X_t, u_t)\,dW_t, \qquad X_0 = x_0 \in H,\ t \in [0,T].$$

The solution depends on a process $u(\cdot)$ with values in another (measurable) space U, called the action space. We assume that $u(\cdot)$ belongs to the set of admissible controls $\mathcal{U}$:

$$\mathcal{U} = \{(\mathcal{F}_t)-\text{progressive processes with values in } U\}.$$

Standing assumptions:

- A is a generator in H, W is a cylindrical Wiener process in K.
- $F : H \times U \to H$ and $G : H \times U \to L(K, H)$ satisfy, for all $t \in (0, T]$, $x, y \in H$, $u \in U$ and for some $\gamma \in [0, 1/2)$ and $K, L > 0$,

$$\|e^{tA}(G(x,u) - G(y,u))\|_{L_2(K,H)} \le Lt^{-\gamma}\|x - y\|_H,$$

$$\|e^{tA}G(x,u)\|_{L_2(K,H)} \le Kt^{-\gamma}(1 + \|x\|_H),$$

$$\|F(x,u) - F(y,u)\|_H \le L\|x - y\|_H, \qquad \|F(x,u)\|_H \le K(1 + \|x\|_H).$$

For $u(\cdot) \in \mathcal{U}$ we call trajectory the corresponding mild solution X^u i.e. an $(\mathcal{F}_t)$-adapted continuous process in H satisfying $\mathbb{P}$-a.s.

$$X_t^u = e^{tA}x + \int_0^t e^{(t-s)A}F(X_s^u, u_s)\,ds + \int_0^t e^{(t-s)A}G(X_s^u, u_s)\,dW_s,\ t \in [0,T].$$

X^u is unique satisfying, for every $p \in [1, \infty)$, and some $C = C(p, L, K, \gamma, A, T)$,

$$\mathbb{E} \sup_{t\in[0,T]} \|X_t\|_H^p \le C(1 + \|x\|)^p.$$

Optimal control problem

We introduce the cost functional

$$J(u(\cdot)) = \mathbb{E}\int_0^T l(X_s^u, u_s)\,ds + \mathbb{E}\phi(X_T^u),$$

where the running cost $l : H \times U \to \mathbb{R}$ and the final cost $\phi : H \to \mathbb{R}$ are assumed to safisfy (as part of the Standing Assumptions), for some $K > 0$ and $m \geq 0$,

$$|l(x,u)| + |\phi(x)| \leq K(1 + \|x\|_H)^m, \qquad x \in H,\, u \in U.$$

The optimal control problem consists in giving conditions for the existence and to characterize an optimal control, i.e. an element $u^*(\cdot) \in \mathcal{U}$ such that

$$J(u^*(\cdot)) \leq J(u(\cdot)), \qquad u(\cdot) \in \mathcal{U}.$$

Note that given data of the optimal control problem are:

$$H, K, A, F, G, l, \phi, x_0, T,$$

and also

$$\Omega, \mathcal{F}, \mathbb{P}, W.$$

We then minimize $J(u(\cdot))$ for $u(\cdot) \in \mathcal{U}$.

Remark. Other formulations of the control problem are also possible, where the set-up $(\Omega, \mathcal{F}, \mathbb{P}, W)$ is not fixed in advance. Note that the cost depends only on the law of (X, u) and not on the particular set-up (provided uniqueness in law holds for the state equation).

Dynamic programming and the HJB equation

The optimal control problem can be embedded in a family of optimal control problems parameterized by the starting point $t \in [0, T]$ and the initial condition $x \in H$:

$$dX_s = AX_s\, ds + F(X_s, u_s)\, ds + G(X_s, u_s)\, dW_s, \qquad X_t = x \in H,\ s \in [t, T] \subset [0, T],$$

$$J^{t,x}(u(\cdot)) = \mathbb{E} \int_t^T l(X_s, u_s)\, ds + \mathbb{E}\phi(X_T).$$

We then define the value function:

$$V(t, x) = \inf_{u(\cdot) \in \mathcal{U}} J^{t,x}(u(\cdot)), \qquad t \in [0, T],\ x \in H.$$

The function V is often a solution to a deterministic partial differential equation on $[0, T] \times H$, called the Hamilton-Jacobi-Bellman equation (HJB).

If uniqueness holds, HJB determines the value function. In many cases, the study of HJB a preliminary step to prove that an optimal control actually exists, and a certain relation between optimal controls and the value function can be established.

To present HJB we need to introduce some differential operators on H.

Spaces of Gâteaux differentiable functions

$g : H \to B$ (Banach space) is said to be G-differentiable if there exists $\nabla g : H \to L(H, B)$ such that

$$\left\| \frac{g(x + \epsilon h) - g(x)}{\epsilon} - \nabla g(x)h \right\|_B \to 0, \text{ as } \epsilon \downarrow 0, \quad x, h \in H.$$

g is said to be twice G-differentiable if $\nabla g : H \to L(H, K)$ is G-differentiable. In particular $\nabla^2 g : H \to L(H, L(H, B)) \simeq$ continuous bilinear forms $H \times H \to B$.

For $f : H \to \mathbb{R}$ we have $\nabla f(x) \in L(H, \mathbb{R}) \simeq H^*$, $\nabla^2 f(x) \in L(H, H^*) \simeq L(H)$ (via the Riesz isometry).
If $H = \mathbb{R}^n$ (column vectors) then $\nabla f(x)$ is a row vector and $\nabla^2 f(x)$ is the hessian matrix.

We write $g \in C^1(H; B)$ if g is continuous, G-differentiable, and $x \mapsto \nabla g(x)h$ is continuous $H \to B$ for every fixed $h \in H$.
Example: $H = B = L^2(a, b)$, $\phi : \mathbb{R} \to \mathbb{R}$ Lipschitz of class C^1, $g(x) = \phi \circ x$.
We set $C^1(H) := C^1(H; \mathbb{R})$.

We say that $v = v(t, x) : [0, T] \times H \to \mathbb{R}$ is in $C^{0,1}([0, T] \times H)$ if v is continuous, G-differentiable with respect to x, and $x \mapsto \nabla_x v(t, x)h$ is continuous $[0, T] \times H$ for every fixed $h \in H$.

Recall the state equation

$$dX_s = AX_s\,ds + F(X_s,u_s)\,ds + G(X_s,u_s)\,dW_s, \qquad X_t = x \in H,\ s \in [t,T] \subset [0,T],$$

For a "test function" $f : H \to \mathbb{R}$ we define

$$\mathcal{L}^u f(x) = \frac{1}{2}\,\mathrm{Trace}\,[\nabla^2 f(x) G(x,u) G(x,u)^*] + \nabla f(x)\,Ax + \nabla f(x)\,F(x,u).$$

Here we assume that the indicated derivatives of f exist, that the trace is well-defined, and that the functional $x \mapsto \nabla f(x)\,Ax$ is continuous for the norm of H and can thus be uniquely extended from $D(A)$ to the whole H.
If no control occurs in the equation for X, $\mathcal{L}$ is called the Kolmogorov operator of X.

$\mathcal{L}^u$ plays a role in the following version of the Ito formula: if $u(\cdot) \in \mathcal{U}$ and X is the corresponding trajectory then

$$df(X_s) = \mathcal{L}^{u(s)} f(X_s)\,ds + \nabla f(X_s) G(X_s,u_s)\,dW_s.$$

If $v(t,x)$ is likewise regular then

$$dv(s,X_s) = \partial_s v(s,X_s)\,ds + \mathcal{L}^{u(s)} v(s,X_s)\,ds + \nabla_x v(s,X_s) G(X_s,u_s)\,dW_s.$$

For an unknown real function $v(t,x)$ $(t \in [0,T], x \in H)$ the HJB equation is

$$\begin{cases} \partial_t v(t,x) + \inf_{u\in U} \{\mathcal{L}^u v(t,x) + l(x,u)\} = 0, & t \in [0,T],\ x \in H, \\ v(T,x) = \phi(x) & x \in H. \end{cases}$$

Verification theorem

Theorem. *Assume that v is a regular solution of HJB. Then $v \le V$. Moreover let $\underline{u} : [0,T] \times H \to U$ be a measurable function such that for all $t \in [0,T], x \in H$*

$$\inf_{u \in U} \{\mathcal{L}^u v(t,x) + l(x,u)\} \quad \text{is attained at} \quad u = \underline{u}(t,x).$$

Also assume that some $u^*(\cdot) \in \mathcal{U}$ and its trajectory X^{u^*} satisfy

$$u_t^* = \underline{u}(t, X_t^{u^*}).$$

Then $u^*(\cdot)$ is optimal and $v \equiv V$.

$\underline{u}$ is called optimal feedback law.
We remark that to find a pair $(u^*(\cdot), X^{u^*})$ related by the optimal feedback one tries to solve the closed-loop equation:

$$dX_s = AX_s\, ds + F(X_s, \underline{u}(s,X_s))\, ds + G(X_s, \underline{u}(s,X_s))\, dW_s, \qquad X_0 = x \in H,\ s \in [t,T],$$

obtained formally replacing u_s by $\underline{u}(s,X_s)$ in the state equation. If there exists a mild solution set

$$u^*(s) := \underline{u}(s, X_s)$$

By the closed-loop equation, X is the trajectory corresponding to $u^*(\cdot)$:

$$X_s = X_s^{u^*},$$

so that we have $u^*(s) = \underline{u}(s, X_s^{u^*})$ by construction.

For later use note that from HJB $\partial_t v + \inf_{u \in U} \{\mathcal{L}^u v + l\} = 0$ we have: $\forall t, x, u$

$$\partial_t v(t,x) + \mathcal{L}^u v(t,x) + l(x,u) \ge 0, \quad \partial_t v(t,x) + \mathcal{L}^{\underline{u}(t,x)} v(t,x) + l(x, \underline{u}(t,x)) = 0.$$

Proof of the theorem. Let $u(\cdot) \in \mathcal{U}$ be arbitrary, $X_s = X_s^u$ starting at (t, x).

$$dv(s, X_s) = \partial_s v(s, X_s)\, ds + \mathcal{L}^{u(s)} v(s, X_s)\, ds + \nabla_x v(s, X_s) G(X_s, u_s)\, dW_s.$$

$$v(T, X_T) - v(t, x) = \phi(X_T) - v(t, x) = \int_t^T \{\partial_s v + \mathcal{L}^{u(s)} v\}\, ds + \int_t^T \nabla v G\, dW_s.$$

Taking expectation and rearranging

$$v(t, x) = \mathbb{E}\phi(X_T) - \mathbb{E} \int_t^T \{\partial_s v + \mathcal{L}^{u(s)} v\}\, ds.$$

Since $J^{t,x}(u(\cdot)) = \mathbb{E}\phi(X_T) + \mathbb{E} \int_t^T l(X_s, u_s)\, ds$, we arrive at the so-called fundamental relation:

$$v(t, x) = J^{t,x}(u(\cdot)) - \mathbb{E} \int_t^T \{\partial_s v(s, X_s)\, ds + \mathcal{L}^{u(s)} v(s, X_s) + l(X_s, u_s)\}\, ds.$$

1) By HJB, $\partial_t v(t, x) + \mathcal{L}^u v(t, x) + l(x, u) \geq 0$, $\forall t, x, u$, so

$$v(t, x) \leq J^{t,x}(u(\cdot)) \qquad \Longrightarrow \qquad v(t, x) \leq \inf_{u(\cdot) \in \mathcal{U}} J^{t,x}(u(\cdot)) = V(t, x).$$

2) For $\underline{u}$ as above we have, as noted earlier,

$$\partial_t v(t, x) + \mathcal{L}^{\underline{u}(t,x)} v(t, x) + l(x, \underline{u}(t, x)) = 0, \quad \forall t, x.$$

If $u^*(\cdot) \in \mathcal{U}$ and $u_t^* = \underline{u}(t, X_t^{u^*})$, replacing t by s and x by $X_s^{u^*}$,

$$\partial_s v(s, X_s^{u^*}) + \mathcal{L}^{u_s^*} v(s, X_s^{u^*} x) + l(x, u_s^*) = 0,$$

and we obtain $v(t, x) = J^{t,x}(u^*(\cdot))$.

Various forms of the HJB equation

Cost functional: $J = \mathbb{E} \int_t^T l(X_s, u_s)\, ds + \mathbb{E}\, \phi(X_T)$

1) General case: $dX_s = AX_s\, ds + F(X_s, u_s)\, dt + G(X_s, u_s)\, dW_s$
Setting

$$\mathcal{L}^u f(x) = \frac{1}{2} \text{Trace}\, [\nabla^2 f(x) G(x,u) G(x,u)^*] + \nabla f(x)\, (Ax + F(X_t, u_t)),$$

the HJB equation is fully non-linear:

$$\partial_t v(t,x) + \inf_{u \in U} \{\mathcal{L}^u v(t,x) + l(x,u)\} = 0.$$

In general no regular (i.e. classical) solution exists and the verification theorem does not apply.

A viscosity solution approach is useful here. It allows weak assumptions on the coefficients. Existence results follow from the so-called Bellman's optimality principle. Uniqueness results are analytic in character, and difficult.

In general, one obtains a solution $v(t,x)$ which is only continuous and coincides with the value function. Existence of an optimal control does not follow immediately. Optimal controls are not generally characterized by feedback laws.

The verification theorem can be applied in the special case of linear state equation and quadratic cost: it leads to the so-called Riccati equation.

In special cases other approaches are possible: we will first consider the case

$$dX_s = AX_s\, ds + F(X_s)\, dt + B(X_s, u_s)\, ds + G(X_s)\, dW_s$$

2) First special case: $dX_s = AX_s\,ds + F(X_s)\,dt + B(X_s,u_s)\,ds + G(X_s)\,dW_s$
Here we have

$$\mathcal{L}^u f(x) = \underbrace{\frac{1}{2}\,\text{Trace}\,[\nabla^2 f(x) G(x) G(x)^*] + \nabla f(x)\,(Ax + F(x))}_{=\ \mathcal{L}f(x)} + \nabla f(x)\,B(x,u),$$

where $\mathcal{L}$ is the Kolmogorov operator in the case $B \equiv 0$. The HJB equation is

$$\partial_t v(t,x) + \mathcal{L}v(t,x) + \inf_{u \in U} \{\nabla_x v(t,x)\,B(x,u) + l(x,u)\} = 0,$$

$$\partial_t v(t,x) + \mathcal{L}v(t,x) = h(x, \nabla_x v(t,x)),$$

where the Hamiltonian h is defined by

$$h(x,p) = -\inf_{u \in U} \{p\,B(x,u) + l(x,u)\}\,, \qquad x \in H,\ p \in H^*.$$

Here the HJB equation is semilinear.

Another approach is possible, based on the concept of mild solutions to HJB. A mild solution $v(t,x)$ is (generally) continuous together with its space derivative $\nabla_x v(t,x)$. Under some conditions, direct analytic proofs of existence and uniqueness results are possible.

They generally allow to show that $v(t,x)$ coincides with the value function and that an optimal control exists in feedback form.

This approach requires some non-degeneracy assumptions on the diffusion coefficients G. These assumptions can be removed in a more particular case:

$$dX_s = AX_s\,ds + F(X_s)\,dt + G(X_s)\,R(X_s,u_s)\,ds + G(X_s)\,dW_s.$$

3) Second special case: $dX_s = AX_s\,ds + F(X_s)\,dt + G(X_s)\,(R(X_s,u_s)\,ds + dW_s)$
We have

$$\mathcal{L}^u f(x) = \mathcal{L} f(x) + \nabla f(x)\, G(x)\, R(x,u),$$

where $\mathcal{L}$ is the Kolmogorov operator in the case $R \equiv 0$:

$$\mathcal{L} f(x) = \frac{1}{2}\,\mathrm{Trace}\,[\nabla^2 f(x) G(x) G(x)^*] + \nabla f(x)\,(Ax + F(x)),$$

The HJB equation $\partial_t v + \inf_{u\in U}\{\mathcal{L}^u v + l\} = 0$ becomes

$$\partial_t v(t,x) + \mathcal{L} v(t,x) + \inf_{u\in U}\{\nabla_x v(t,x) G(x) R(x,u) + l(x,u)\} = 0,$$

$$\partial_t v(t,x) + \mathcal{L} v(t,x) = \psi(x, \nabla_x v(t,x) G(x)),$$

where Hamiltonian ψ is now defined by

$$\psi(x,z) = -\inf_{u\in U}\{z\, R(x,u) + l(x,u)\}, \qquad x \in H,\ z \in K^*.$$

The HJB equation is semilinear of special form, i.e. the nonlinear term depends on $\nabla_x v G$, not simply $\nabla_x v$.

The concept of mild solution $v(t,x)$ will be introduced and used (v and $\nabla_x v$ continuous functions).

A probabilistic approach based on backward stochastic differential equations (BSDEs) allows to prove existence and uniqueness, to show that v coincides with the value function and that an optimal control exists in feedback form.

Non-degeneracy assumptions on G are removed, at the expense of a more particular form of the state equation.

Some references.

Stochastic calculus in infinite-dimensional spaces.

G. Da Prato and J. Zabczyk, Stochastic equations in infinite dimensions. Cambridge University Press, 1992.

C. Prévôt and M. Röckner. A Concise Course on Stochastic Partial Differential Equations. Springer, 2007.

T. Zhang. Stochastic evolution equations. Lecture notes, ITN School September 2009, Manchester.

Semigroup theory.

A. Pazy. Semigroups of linear operators and applications to partial differential equations. Springer, 1983.

K-J. Engel, R. Nagel. One-parameter semigroups for linear evolution equations. Springer, 2000.

Deterministic delay equations.

J. Hale, Theory of functional differential equations. Springer, 1977.

J.K. Hale, S.M. Verduyn Lunel. Introduction to functional differential equations. Springer 1993.

A. Bátkai, S. Piazzera, Semigroups for delay equations. Springer, 2005.

Stochastic delay equations.

S.E.A. Mohammed, Stochastic functional differential equations. Pitman, 1984.

S.E.A. Mohammed, Stochastic differential systems with memory: theory, examples and applications, in Stochastic analysis and related topics. VI. Proceedings of the 6th Oslo-Silivri Workshop held in Geilo, July 29–August 6, 1996. Progress in Probability, 42. Birkhäuser Boston, 1998.

Backward stochastic differential equations.

E. Pardoux, S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems and Control Lett. **14**, 1990, 55-61.

N. El Karoui, S. Peng, M. C. Quenez, Backward stochastic differential equations in finance. Mathematical Finance 7 (1997), no. 1, 1-71.

E. Pardoux, Backward Stochastic Differential Equations and Viscosity Solutions of Systems of Semilinear Parabolic and Elliptic PDEs of Second Order. In: Stochastic Analysis and Related Topics VI. (Eds.: L. Decreusefond, J. Gjerde, B. Øksendal, A.S. Üstünel) The Geilo Workshop, 1996. Progress in Probability 42, Birkhäuser, Boston 1998,

E. Pardoux, BSDE's, weak convergence and homogeneization of semilinear PDE's. In: Nonlinear analysis, differential equations and control, eds. F.H. Clarke, R.J. Stern, 503-549, Kluwer, 1999.

Classical solutions of HJB in infinite-dimensional spaces.

Barbu-Da Prato (book, 1983), Da Prato-Zabczyk (3 books).

Viscosity solutions of HJB in infinite-dimensional spaces.

P.L. Lions (1988, 1989a, 1989b), Crandall-Kocan-Święch (1993/94), Święch (1994), Kocan-Święch (1995), Gozzi-Rouy-Święch (2000), Gozzi-Święch (2000).

See the lectures by Fausto Gozzi.

Mild solutions of HJB in infinite-dimensional spaces: analytic approach

Da Prato-Cannarsa (1991, 1992), Gozzi (1995, 1996), Cerrai (LNM 1762), Masiero (2005).

Mild solutions of HJB in infinite-dimensional spaces: the approach via BSDEs.

M. Fuhrman, G. Tessitore, (2002a, 2002b, 2004, 2005). Ph. Briand, F. Confortola, (2008a, 2008b, 2008c). Masiero (2008).

G. Tessitore, Lecture notes, ITN School Roscoff (previous edition).

Approach to HJB via backward stochastic differential equations.

1. Parabolic partial differential equations on Hilbert spaces
2. The associated backward stochastic differential equation (BSDE)
3. Existence and uniqueness of mild solutions
4. The optimal control problem and the HJB equation
5. Examples

Parabolic PDEs on Hilbert spaces

We consider again the general evolution equation of the form

$$dX_t = AX_t\,dt + F(X_t)\,dt + G(X_t)\,dW_t, \qquad X_0 = x \in H,\ t \in [0,T].$$

Standing assumptions:

- A is a generator in H, W is a cylindrical Wiener process in K.
- $F : H \to H$ and $G : H \to L(K,H)$ satisfy, for all $t \in (0,T]$, $x, y \in H$,

$$\|e^{tA}(G(x) - G(y))\|_{L_2(K,H)} \le Lt^{-\gamma}\|x-y\|_H,\ \|e^{tA}G(x)\|_{L_2(K,H)} \le Kt^{-\gamma}(1+\|x\|_H),$$

$$\|F(x) - F(y)\|_H \le L\|x-y\|_H, \qquad \|F(x)\|_H \le K(1+\|x\|_H),$$

for some $\gamma \in [0,1/2)$ and $K, L > 0$.

We look for a solution X in the space $\mathcal{S}_p$ ($1 \le p < \infty$), i.e. an $(\mathcal{F}_t)$-adapted process in H with continuous paths, satisfying

$$\|X\|_{\mathcal{S}_p}^p := \mathbb{E} \sup_{t\in[0,T]} \|X_t\|_H^p < \infty.$$

There exists a unique mild solution X, i.e. a process in $\mathcal{S}_p$ for every $p \in [1,\infty)$ satisfying $\mathbb{P}$-a.s.

$$X_t = e^{tA}x + \int_0^t e^{(t-s)A}F(X_s)\,ds + \int_0^t e^{(t-s)A}G(X_s)\,dW_s,\ t \in [0,T].$$

For every starting point $t \in [0, T]$ and initial condition $x \in H$ we solve

$$\begin{cases} dX_s = AX_s\, ds + F(X_s)\, ds + G(X_s)\, dW_s, & s \in [t, T] \subset [0, T], \\ X_t = x \in H. \end{cases}$$

The mild solution $X = \{X_s^{t,x},\ 0 \le t \le s \le T,\ x \in H\}$ defines a Markov process in H, whose transition semigroup we denote $\{P_{t,s} : 0 \le t \le s \le T\}$.

This means the following: each $P_{t,s}$ is an operator acting on bounded measurable functions $\phi : H \to \mathbb{R}$ by the formula:

$$P_{t,s}[\phi](x) = \mathbb{E}\, \phi(X_s^{t,x}), \qquad x \in H.$$

and the Markov property is: for $x \in H$, $u \le t \le s$,

$$\mathbb{E}\,[\phi(X_s^{u,x}) \,|\, \mathcal{F}_t] = P_{t,s}[\phi](X_t^{u,x}), \qquad \mathbb{P} - a.s.$$

We denote $\mathcal{L}$ the Kolmogorov operator of X: for regular f,

$$\mathcal{L}f(x) = \frac{1}{2}\,\mathrm{Trace}\,[\nabla^2 f(x) G(x) G(x)^*] + \nabla f(x)\,(Ax + F(x)),$$

The (linear) Kolmogorov equation is

$$\begin{cases} \partial_t u(t,x) + \mathcal{L}u(t,x) = 0, & t \in [0, T],\ x \in H, \\ u(T,x) = \phi(x), & x \in H. \end{cases}$$

It is known that the function $u(t,x) = P_{t,T}[\phi](x)$, is a candidate solution of the Kolmogorov equation. (see precise statements in [Da Prato-Zabczyk 2002, book, 2nd order PDEs on Hilbert spaces]).

More generally, the solution of the linear nonhomogeneous equation

$$\begin{cases} \partial_t u(t,x) + \mathcal{L}u(t,x) = f(t,x), & t \in [0,T],\ x \in H, \\ u(T,x) = \phi(x), & x \in H. \end{cases}$$

has the candidate solution given by the variation of constants formula:

$$u(t,x) = P_{t,T}[\phi](x) - \int_t^T P_{t,s}\left[f(s,\cdot)\right](x)\, ds.$$

Now let us consider a semilinear parabolic PDE on H of the form

$$\begin{cases} \partial_t v(t,x) + \mathcal{L}v(t,x) = \psi(x, v(t,x), \nabla v(t,x) G(x)), & t \in [0,T],\ x \in H, \\ v(T,x) = \phi(x), & x \in H, \end{cases}$$

for some $\psi : H \times \mathbb{R} \times K^* \to \mathbb{R}$.
The HJB equation was a special case, with ψ given by the Hamiltonian of a control problem (and independent of v).

In analogy with the variation of constants formula, we call v a mild solution if it satisfies, for $t \in [0,T]$, $x \in H$,

$$v(t,x) = P_{t,T}[\phi](x) - \int_t^T P_{t,s}\left[\psi(\cdot, v(s,\cdot), \nabla v(s,\cdot) G(s,\cdot))\right](x)\, ds.$$

We will require the existence of $\nabla v(t,x)$ and appropriate continuity and growth conditions for v and ψ.

The backward stochastic differential equation

With $\{X_s^{t,x},\ s \in [t,T]\}$ we consider the backward differential equation for the unknown process $\{(Y_s, Z_s),\ s \in [t,T]\}$, in the sense of Pardoux and Peng 1990:

$$\begin{cases} dY_s = \psi(X_s^{t,x}, Y_s, Z_s)\, ds + Z_s\, dW_s, & s \in [t,T], \\ Y_T = \phi(X_T^{t,x}). \end{cases}$$

Y is real and Z takes values in K^*.

Standing assumptions on ϕ and ψ:

• $\phi : H \to \mathbb{R}$, $\psi : H \times \mathbb{R} \times K^* \to \mathbb{R}$ satisfy, for some $L > 0$, $m \geq 0$,

$$\begin{aligned} &|\phi(x_1) - \phi(x_2)| \leq L\, |x_1 - x_2|, \\ &|\psi(x, y_1, z_1) - \psi(x, y_2, z_2)| \leq L\, (|y_1 - y_2| + |z_1 - z_2|), \\ &|\psi(x_1, y, z) - \psi(x_2, y, z)| \leq L\, |x_2 - x_1|(1 + |z|)(1 + |x_1| + |x_2| + |y|)^m, \end{aligned}$$

for all $x, x_1, x_2 \in H$, $y, y_1, y_2 \in \mathbb{R}$, $z, z_1, z_2 \in K^*$.

We recall some facts from the standard theory of BSDEs.

• There exists a unique solution $(Y, Z) \in \mathcal{S}_p \times \mathcal{H}_p$ for all $1 \leq p < \infty$.
This means Y and Z are $(\mathcal{F}_t)$-adapted processes in $\mathbb{R}$ and K^* satisfying

$$\|Y\|^p_{\mathcal{S}_p} := \mathbb{E} \sup_{s \in [t,T]} |Y_s|^p < \infty, \qquad \|Z\|^p_{\mathcal{H}_p} := \mathbb{E} \left(\int_t^T \|Z_s\|^2_{K^*} ds \right)^p < \infty.$$

• Y_t is deterministic.

$$\begin{cases} dY_s = \psi(X_s^{t,x}, Y_s, Z_s)\, ds + Z_s\, dW_s, & s \in [t,T], \\ Y_T = \phi(X_T^{t,x}). \end{cases}$$

We denote $Y_s = Y_s^{t,x}$, $Z_s = Z_s^{t,x}$ and we set

$$v(t,x) = Y_t^{t,x}, \qquad t \in [0,T],\ x \in H.$$

• There exists a Borel function $\zeta(t,x)$ such that for $0 \le t \le s \le T$, $x \in H$,

$$Y_s^{t,x} = v(s, X_s^{t,x}), \qquad Z_s^{t,x} = \zeta(s, X_s^{t,x}).$$

In Peng 1992, Pardoux-Peng 1992, in the case $H = \mathbb{R}^n$ it was proved that if the coefficients are sufficiently regular then the function $v(t,x)$ is a classical solution to our semilinear PDE:

$$\begin{cases} \partial_t v(t,x) + \mathcal{L}v(t,x) = \psi(x, v(t,x), \nabla_x v(t,x) G(x)), & t \in [0,T],\ x \in H, \\ v(T,x) = \phi(x), & x \in H, \end{cases}$$

It was also proved that if the coefficients are not smooth then $v(t,x)$ is the unique viscosity solution of the PDE.

In the case of a Hilbert space we will prove that v is the unique mild solution, under appropriate conditions. In particular, $v(t,\cdot) \in C^1$.

Regularity of v and identification of Z

Given $(t,x) \in [0,T] \times H$ we set $v(t,x) = Y_t^{t,x}$, where

$$\begin{cases} dX_s^{t,x} = AX_s^{t,x}\, ds + F(X_s^{t,x})\, ds + G(X_s^{t,x})\, dW_s, & s \in [t,T] \subset [0,T], \\ X_t^{t,x} = x, \\ dY_s^{t,x} = \psi(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\, ds + Z_s^{t,x}\, dW_s, \\ Y_T^{t,x} = \phi(X_T^{t,x}). \end{cases}$$

Regularity assumptions:

$$F(\cdot) \in C^1(H;H), \qquad e^{tA}G(\cdot) \in C^1(H; L_2(K,H)),$$

$$\phi(\cdot) \in C^1(H;\mathbb{R}), \qquad \psi(\cdot,\cdot,\cdot) \in C^1(H \times \mathbb{R} \times K^*;\mathbb{R}).$$

Theorem (regularity of v) *Under the standard and the regularity assumptions we have*

$$v \in C^{0,1}([0,T] \times H),$$

and v, $\nabla_x v$ have polynomial growth in x uniformly with respect to t.

The last assertion means that for some $C > 0$, $r \geq 0$ we have

$$|v(t,x)| + \|\nabla_x v(t,x)\|_{H^*} \leq C(1 + \|x\|_H)^r \qquad t \in [0,T],\ x \in H.$$

Idea of the proof. One proves the differentiability of the maps

$$x \mapsto (X_s^{t,x})_s, \quad x \mapsto (Y_s^{t,x})_s, \quad x \mapsto (Z_s^{t,x})_s,$$

from H to $\mathcal{S}_p$, $\mathcal{S}_p$, $\mathcal{H}_p$ respectively. Differentiability of $x \mapsto Y_t^{t,x} = v(t,x)$ then follows immediately.
[Note: this does not imply that the maps

$$x \mapsto X_s^{t,x}, \quad x \mapsto Y_s^{t,x}, \quad x \mapsto Z_s^{t,x},$$

are differentiable $\mathbb{P}$-a.s. from H to H, $\mathbb{R}$, K^* respectively, for any fixed t, s: no stochastic flow exists in general in infinite dimensions!]

One formally differentiates the forward-backward system

$$\begin{cases} dX_s^{t,x} = AX_s^{t,x}\, ds + F(X_s^{t,x})\, ds + G(X_s^{t,x})\, dW_s, \quad s \in [t,T], \\ X_t^{t,x} = x, \\ dY_s^{t,x} = \psi(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\, ds + Z_s^{t,x}\, dW_s, \\ Y_T^{t,x} = \phi(X_T^{t,x}). \end{cases}$$

and proves that $(\nabla X, \nabla Y, \nabla Z)$ solve the forward-backward system

$$\begin{cases} d\nabla X_s^{t,x} &=& A\nabla X_s^{t,x}\, ds + \nabla F(X_s^{t,x})\nabla X_s^{t,x}\, ds + \nabla G(X_s^{t,x})\nabla X_s^{t,x}\, dW_s, \\ \nabla X_t^{t,x} &=& I, \\ d\nabla Y_s^{t,x} &=& \nabla_x\psi(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\nabla X_s^{t,x}\, ds + \nabla_y\psi(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\nabla Y_s^{t,x}\, ds \\ && +\nabla_z\psi(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\nabla Z_s^{t,x}\, ds + \nabla Z_s^{t,x}\, dW_s, \\ \nabla Y_T^{t,x} &=& \nabla\phi(X_T^{t,x})\nabla X_T^{t,x}. \end{cases}$$

The other properties of $(\nabla X, \nabla Y, \nabla Z)$ also follow from the forward-backward system.

Theorem (identification of Z) *Under the standard and the regularity assumptions we have*

$$Z_s^{t,x} = \nabla_x v(s, X_s^{t,x})\, G(X_s^{t,x}), \qquad a.e.\ s \in [t,T],\ x \in H.$$

Justification. By the BSDE

$$v(s, X_s^{t,x}) = Y_s^{t,x} = Y_t^{t,x} + \int_t^s \psi(X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\, dr + \int_t^s Z_r^{t,x}\, dW_r,$$

so that, recalling that $W_t = \sum_i \beta_t^i e_i$,

$$\langle v(\cdot, X_\cdot^{t,x}), \beta^i \rangle_s = \int_t^s Z_r^{t,x} e_i\, dr.$$

If we could apply the Ito formula to v:

$$v(s, X_s^{t,x}) = v(t,x) + \int_t^s (\partial_t v(r, X_r^{t,x}) + \mathcal{L} v(r, X_r^{t,x})\, dr + \int_t^s \nabla_x v(r, X_r^{t,x}) G(X_r)\, dW_r,$$

we would obtain

$$\langle v(\cdot, X_\cdot^{t,x}), \beta^i \rangle_s = \int_t^s \nabla_x v(r, X_r^{t,x}) G(X_r) e_i\, dr, \tag{6}$$

and we would conclude $Z_s^{t,x} e_i = \nabla_x v(s, X_s^{t,x}) G(X_s) e_i$.
Even if the Ito formula does not apply, (6) still holds by the following:

Lemma. *(6) is true for every $v \in C^{0,1}([0,T] \times H)$ with v and $\nabla_x v$ having polynomial growth in x uniformly with respect to t.*

The proof of the lemma uses the Malliavin calculus.

Existence and uniqueness of the mild solution

We say that v is a mild solution of the PDE

$$\begin{cases} \partial_t v(t,x) + \mathcal{L}v(t,x) = \psi(x, v(t,x), \nabla_x v(t,x) G(x)), & t \in [0,T],\ x \in H, \\ v(T,x) = \phi(x), & x \in H, \end{cases}$$

if $v \in C^{0,1}([0,T] \times H)$, v and $\nabla_x v$ have polynomial growth in x uniformly with respect to t, and for $t \in [0,T]$, $x \in H$,

$$v(t,x) = P_{t,T}[\phi](x) - \int_t^T P_{t,s}\left[\psi(\cdot, v(s,\cdot), \nabla_x v(s,\cdot) G(s,\cdot))\right](x)\, ds.$$

Theorem *Assume the standard and the regularity conditions and for $x \in H$, $t \in [0,T]$ consider the system*

$$\begin{cases} dX_s^{t,x} = AX_s^{t,x}\, ds + F(X_s^{t,x})\, ds + G(X_s^{t,x})\, dW_s, & s \in [t,T] \subset [0,T], \\ X_t^{t,x} = x, \\ dY_s^{t,x} = \psi(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\, ds + Z_s^{t,x}\, dW_s, \\ Y_T^{t,x} = \phi(X_T^{t,x}). \end{cases}$$

Setting $v(t,x) = Y_t^{t,x}$, then v is the unique mild solution.

Proof. **Existence.** Mild solution:

$$v(t,x) = P_{t,T}[\phi](x) - \int_t^T P_{t,s}\left[\psi(\cdot, v(s,\cdot), \nabla_x v(s,\cdot) G(s,\cdot))\right](x)\, ds.$$

Fix t,x and denote $X_s = X_s^{t,x}$, $Y_s = Y_s^{t,x}$, $Z_s = Z_s^{t,x}$, $s \in [t,T]$.
We know $v \in C^{0,1}$ with appropriate growth and

$$Y_s = v(s, X_s), \qquad Z_s = \nabla_x v(s, X_s)\, G(s, X_s).$$

Then

$$P_{t,T}[\phi](x) = \mathbb{E}\, \phi(X_T)$$

and

$$\begin{aligned} & P_{t,s}[\psi(\cdot, v(s,\cdot), \nabla_x v(s,\cdot) G(s,\cdot))](x) \\ & \quad = \mathbb{E}\, \psi\big(X_s, v(s,X_s), \nabla v(s,X_s) G(s,X_s)\big) \\ & \quad = \mathbb{E}\, \psi\big(X_s, Y_s, Z_s\big). \end{aligned}$$

Next we recall the backward equation

$$Y_t + \int_t^T Z_s dW_s = \phi(X_T) - \int_t^T \psi(X_s, Y_s, Z_s) ds,$$

and we take expectation:

$$\begin{aligned} v(t,x) = Y_t &= \mathbb{E}\, \phi(X_T) - \int_t^T \mathbb{E}\, \psi(X_s, Y_s, Z_s) ds \\ &= P_{t,T}[\phi](x) - \int_t^T P_{t,s}[\psi(\cdot, v(s,\cdot), \nabla_x v(s,\cdot) G(s,\cdot))](x)\, ds. \end{aligned}$$

Uniqueness Let v be a mild solution: for $0 \leq s \leq T$, $y \in H$,

$$v(s,y) = P_{s,T}[\phi](y) - \int_s^T P_{s,r}\left[\psi(\cdot, v(r,\cdot), \nabla_x v(r,\cdot)G(\cdot))\right](y)\, dr.$$

Now take $0 \leq t \leq s$ and $x \in H$ and replace y with $X_s^{t,x} = X_s$ to get

$$v(s,X_s) = P_{s,T}[\phi](X_s) - \int_s^T P_{s,r}\left[\psi(\cdot, v(r,\cdot), \nabla_x v(r,\cdot)G(\cdot))\right](X_s)\, dr.$$

So by the Markov property $P_{s,T}[\phi](X_s) = \mathbb{E}^{\mathcal{F}_s}\left[\phi(X_T)\right]$ and

$$P_{s,r}\left[\psi(\cdot, v(r,\cdot), \nabla_x v(r,\cdot)G(\cdot))\right](X_s) = \mathbb{E}^{\mathcal{F}_s}\psi(X_r, v(r,X_r), \nabla_x v(r,X_r)G(X_r)).$$

So we obtain

$$\begin{aligned} v(s,X_s) &= \mathbb{E}^{\mathcal{F}_s}\left[\phi(X_T)\right] - \mathbb{E}^{\mathcal{F}_s}\left[\int_s^T \psi(X_r, v(r,X_r), \nabla_x v(r,X_r)G(X_r))\, dr\right] \\ &= \mathbb{E}^{\mathcal{F}_s}\left[\chi\right] + \int_t^s \psi(X_r, v(r,X_r), \nabla_x v(r,X_r)G(X_r))\, dr, \end{aligned}$$

where $\chi = \phi(X_T) - \int_t^T \psi(X_r, u(r,X_r), \nabla_x v(r,X_r)G(X_r))\, dr$.
By the representation theorem there exists $\widetilde{Z}$ for which

$$v(s,X_s) = \int_t^s \widetilde{Z}_r\, dW_r + v(t,x) + \int_t^s \psi(X_r, u(r,X_r), \nabla_x v(r,X_r)G(X_r))\, dr.$$

$$v(s,X_s)=\int_t^s \widetilde{Z}_r\,dW_r+v(t,x)+\int_t^s \psi(X_r,u(r,X_r),\nabla_x v(r,X_r)G(X_r))\,dr, \quad (7)$$

Computing the joint quadratic variation with β^i we get

$$\langle v(\cdot,X_\cdot),\beta^i\rangle_s=\int_t^s \nabla_x v(r,X_r)G(X_r)e_i\,dr, \quad \text{by a previous lemma, and}$$
$$\langle v(\cdot,X_\cdot),\beta^i\rangle_s=\int_t^s \widetilde{Z}_r e_i\,dr, \quad \text{by (7).}$$

Comparing we conclude that $\mathbb{P}$-a.s.

$$\nabla_x v(s,X_s)G(X_s)=\widetilde{Z}_s, \qquad \text{for a.e. } \; s\in[t,T].$$

Substituting into (7) we obtain

$$dv(s,X_s)=\nabla_x v(s,X_s)G(X_s)\,dW_s+\psi(X_s,v(s,X_s),\nabla_x v(s,X_s)G(X_s))\,ds,$$

Comparing with the backward equation

$$dY_s=Z_s\,dW_s+\psi(X_s,Y_s,Z_s)\,ds,$$

we see that the pairs

$$(v(s,X_s),\nabla_x v(s,X_s)G(X_s)), \qquad (Y_s,Z_s),$$

solve the same equation and, by uniqueness, they coincide. In particular,

$$v(s,X_s^{t,x})=Y_s^{t,x}, \qquad s\in[t,T],$$

and setting $s=t$ we finally obtain $v(t,x)=Y_t^{t,x}$.

The optimal control problem and the HJB equation

State equation: for $t \in [0,T]$, $x \in H$,

$$\begin{cases} dX_s = AX_s\,ds + F(X_s)\,dt + G(X_s)\,(R(X_s,u_s)\,ds + dW_s), & s \in [t,T] \subset [0,T], \\ X_t = x \in H. \end{cases}$$

Here $u(\cdot) \in \mathcal{U} = \{(\mathcal{F}_t)\text{-progressive } U\text{-valued processes}\}$. Then $X = X^{t,x,u}$.

Cost functional and value function: for $t \in [0,T]$, $x \in H$,

$$J^{t,x}(u(\cdot)) = \mathbb{E}\int_t^T l(X_s,u_s)\,ds + \mathbb{E}\phi(X_T), \qquad V(t,x) = \inf_{u(\cdot)\in\mathcal{U}} J^{t,x}(u(\cdot)).$$

We assume that $l : H \times U \to \mathbb{R}$ and $R : H \times U \to K$ are measurable and

$$|l(x,u)| \le C(1 + \|x\|_H)^m, \qquad \|R(x,u)\|_K \le C, \qquad x \in H, u \in U,$$

for some $C > 0$, $m \ge 0$.

The Hamiltonian ψ is

$$\psi(x,z) = -\inf_{u\in U}\{z\,R(x,u) + l(x,u)\}, \qquad x \in H,\ z \in K^*.$$

We assume that A, F, G, ϕ, ψ satisfy the standard and regularity assumptions.
[**Note:** it is not very satisfactory to impose assumptions directly on ψ.]

$$\psi(x,z) = -\inf_{u\in U} \{z\, R(x,u) + l(x,u)\}, \qquad x \in H,\ z \in K^*.$$

The HJB equation is

$$\partial_t v(t,x) + \mathcal{L} v(t,x) = \psi(x, \nabla_x v(t,x) G(x)),$$

where $\mathcal{L}$ is the Kolmogorov operator in the case $R \equiv 0$:

$$\mathcal{L} f(x) = \frac{1}{2} \mathrm{Trace}\,[\nabla^2 f(x) G(x) G(x)^*] + \nabla f(x)\,(Ax + F(x)).$$

We know that HJB has a unique mild solution v.

We will also assume that there exist nonempty $\Gamma(x,z) \subset U$ such that for $x \in H$, $z \in K^*$,

$$\inf_{u\in U} \{z\, R(x,u) + l(x,u)\} \quad \text{is attained at} \quad u \in \Gamma(x,z),$$

and that there exists a measurable function $\gamma : H \times K^* \to U$ such that

$$\gamma(x,z) \in \Gamma(x,z), \qquad x \in H, z \in K^*.$$

For later use note that $\psi(x,z) + z\, R(x,u) + l(x,u) \geq 0$ and

$$\psi(x,z) + z\, R(x,u) + l(x,u) = 0 \qquad \Longleftrightarrow \qquad u \in \Gamma(x,z),$$

and $\psi(x,z) + z\, R(x,\gamma(x,z)) + l(x,\gamma(x,z)) = 0$.

Theorem *For every* $t \in [0,T]$, $x \in H$, $u(\cdot) \in \mathcal{U}$ *we have*

$$v(t,x) \le J^{t,x}(u(\cdot)).$$

In particular, $v(t,x) \le V(t,x)$.

A control $u^*(\cdot)$ *attains the lower bound:* $v(t,x) = J^{t,x}(u^*(\cdot))$ *if and only if*

$$u^*(s) \in \Gamma(X_s^{t,x,u^*}, \nabla_x v(s, X_s^{t,x,u^*})), \qquad a.e.\ s,$$

and in this case it is optimal.

Let us define the feedback law

$$\underline{u}(t,x) = \gamma(x, \nabla_x v(t,x)), \qquad t \in [0,T],\ x \in H.$$

and the closed-loop equation:

$$dX_s = AX_s\,ds + F(X_s)\,ds + G(X_s)(R(X_s, \underline{u}(s,X_s))ds + dW_s),\ X_t = x \in H,\ s \in [t,T],$$

obtained formally replacing u_s by $\underline{u}(s,X_s)$ in the state equation.

Corollary *If the closed-loop equation has a solution then the control*

$$u^*(s) := \underline{u}(s, X_s)$$

is optimal.

Proof of the Corollary. By the closed-loop equation, X is the trajectory corresponding to $u^*(\cdot)$:

$$X_s = X_s^{u^*},$$

so we have $u^*(s) = \underline{u}(s, X_s^{u^*}) = \gamma(X_s^{u^*}, \nabla_x v(s, X_s^{u^*})) \in \Gamma(X_s^{t,x,u^*}, \nabla_x v(s, X_s^{t,x,u^*}))$.

Remark The closed loop equation always has a weak solution. Consequently, the optimal control problem has a solution in an appropriate weak formulation.

Proof of the Theorem. Fix $u(\cdot) \in \mathcal{U}$ and recall the state equation for $X = X^{t,x,u} = X^{t,x}$:

$$\begin{cases} dX_s = AX_s\, ds + F(X_s)\, dt + G(X_s)\,(R(X_s,u_s)\, ds + dW_s), & s \in [t,T] \subset [0,T], \\ X_t = x \in H. \end{cases}$$

By the Girsanov theorem the process $(W^u)_{s\in[t,T]}$ defined by

$$dW^u_s = R(X_s,u_s)\, ds + dW_s,$$

is a Wiener process under another probability $\mathbb{Q}^u$ on $(\Omega, \mathcal{F})$, equivalent to $\mathbb{P}$. Then

$$dX^{t,x}_s = AX^{t,x}_s\, ds + F(X^{t,x}_s)\, ds + G(X^{t,x}_s)\, dW^u_s, \qquad X_t = x,$$

and we consider the associated BSDE

$$dY^{t,x}_s = \psi(X^{t,x}_s, Z^{t,x}_s)\; ds + Z^{t,x}_s\; dW^u_s, \qquad Y^{t,x}_T = \phi(X^{t,x}_T).$$

here (X,Y,Z) depend on $u(\cdot)$ but its law does not, so

$$v(t,x) = Y^{t,x}_t = \mathbb{E}^u Y^{t,x}_t \quad \text{does not depend on } u(\cdot)$$

(it is determined by A, F, G, ϕ, ψ, T). By the BSDE

$$v(t,x) = Y_t = \phi(X^{t,x}_T) - \int_t^T \psi(X^{t,x}_s, Z^{t,x}_s)\; ds - \int_t^T Z^{t,x}_s\; (R(X^{t,x}_s, u_s)\, ds + dW_s).$$

Add and subtract $\int_t^T l(X_s,u_s)\, ds$ and take $\mathbb{E}$:

$$v(t,x) = J^{t,x}(u(\cdot)) - \int_t^T \{\psi(X^{t,x}_s, Z^{t,x}_s) + Z^{t,x}_s\, R(X^{t,x}_s, u_s) + l(X^{t,x}_s, u_s)\}\, ds.$$

This is called the fundamental relation.

$$v(t,x) = J(u(\cdot)) - \int_t^T \{\psi(X_s, Z_s) + Z_s\, R(X_s, u_s) + l(X_s, u_s)\}\, ds.$$

We noticed that $\psi(x,z) + z\, R(x,u) + l(x,u) \geq 0$ and

$$\psi(x,z) + z\, R(x,u) + l(x,u) = 0 \qquad \Longleftrightarrow \qquad u \in \Gamma(x,z).$$

First we have $\{\ldots\} \geq 0$ which implies

$$v(t,x) \leq J(u(\cdot)), \qquad u(\cdot) \in \mathcal{U},$$

and so $v(t,x) \leq V(t,x)$.

Then we have $v(t,x) = J(u^*(\cdot))$ for some $u^*(\cdot) \in \mathcal{U}$ if and only if $\{\ldots\} = 0$, i.e.

$$\psi(X_s, Z_s) + Z_s\, R(X_s, u^*(s)) + l(X_s, u^*(s)) = 0 \qquad \Longleftrightarrow \qquad u^*(s) \in \Gamma(X_s, Z_s).$$

Example 1: the controlled stochastic heat equation

We take $\mathcal{O} \subset \mathbb{R}^d$ open bounded regular, and the cylindrical Wiener process W in $K = L^2(\mathcal{O})$.

$$\begin{cases} dX_t(\xi) = \Delta_\xi X_t(\xi)\, dt + f(X_t(\xi))\, dt + G\,(u(t,\xi)\, dt + dW_t(\xi)), & \xi \in \mathcal{O},\ t \in [0,T], \\ X_0(\xi) = x(\xi), & \xi \in \mathcal{O}, \\ X_t(\xi) = 0, & \xi \in \partial\mathcal{O},\ t \in [0,T]. \end{cases}$$

We assume that $(u(t,\cdot))$ takes values in $U = \{u \in H\ :\ \|u\|_H \leq r\} = B_r(0) \subset H$.

Cost functional

$$J(u(\cdot)) = \mathbb{E}\int_0^T \int_{\mathcal{O}} [l_0(X_t(\xi)) + |u(t,\xi)|^2]\, d\xi\, dt + \mathbb{E}\int_{\mathcal{O}} \phi_0(X_T(\xi))\, d\xi.$$

The functions f, l_0, ϕ_0 are real-valued and belong to $Lip(\mathbb{R}) \cap C^1(\mathbb{R})$.

We define $H = L^2(\mathcal{O})$, $A = \Delta_\xi$, $D(A) = H^2(\mathcal{O}) \cap H^1_0(\mathcal{O})$ so that for a basis (e_i) of H we have $Ae_i = -\alpha_i e_i$, for $0 < \alpha_i \uparrow \infty$.

We take G satisfying $Ge_i = \sqrt{q_i} e_i$ with

$$0 \leq q_i \leq K, \qquad \Big(\sum_i q_i e^{-2\alpha_i t}\Big)^{1/2} \leq K t^{-\gamma}, \qquad t \in (0,T],$$

for some $\gamma \in [0,1/2)$ and $K > 0$.
[If $d = 1$ we can take $G = I$.]

State equation and cost:

$$dX_t(\xi) = \Delta_\xi X_t(\xi)\, dt + f(X_t(\xi))\, dt + G\,(u(t,\xi)\, dt + dW_t(\xi)),$$

$$J(u(\cdot)) = \mathbb{E}\int_0^T \int_{\mathcal{O}} [l_0(X_t(\xi)) + |u(t,\xi)|^2]\, d\xi\, dt + \mathbb{E}\int_{\mathcal{O}} \phi_0(X_T(\xi))\, d\xi.$$

We define $F : H \to H$, $l : H \times U \to H$, $\phi : H \to \mathbb{R}$ setting

$$F(x)(\xi) = f(x(\xi))\ (\xi \in \mathcal{O}), \qquad l(x,u) = \int_{\mathcal{O}} [l_0(x(\xi)) + |u(\xi)|^2]\, d\xi,$$

$$\phi(x) = \int_{\mathcal{O}} \phi_0(X_T(\xi))\, d\xi,$$

and we obtain

$$dX_t = AX_t\, dt + F(X_t)\, dt + G\,(u(t)\, dt + dW_t),$$

$$J(u(\cdot)) = \mathbb{E}\int_0^T l(X_t, u_t)\, dt + \mathbb{E}\phi(X_T).$$

All the required assumptions are satisfied and the previous results apply.

Example 2: controlled stochastic delay equations

$$\begin{cases} dz(t) = \left(\int_{-r}^{0} z(t+\theta)\, a(d\theta)\right) dt + f(z(t))\, dt + u(t)\, dt + dW_t, & t \in [0,T], \\ z(0) = x_0, & \\ z(\theta) = x_1(\theta), & \theta \in [-r,0], \end{cases}$$

where W is a Wiener process in $K = \mathbb{R}$, $x_0 \in \mathbb{R}, x_1 \in L^2(-r,0)$, $a(d\theta)$ is a signed finite measure on $[-r,0]$.

We assume that $u(\cdot)$ takes values in a compact interval $U \subset \mathbb{R}$.

Cost functional

$$J(u(\cdot)) = \mathbb{E} \int_0^T [l_0(z(t)) + |u(t)|^2]\, ds + \mathbb{E}\phi_0(z(T)).$$

The functions f, l_0, ϕ_0 are real-valued and belong to $Lip(\mathbb{R}) \cap C^1(\mathbb{R})$.

We introduced the state space $H = \mathbb{R} \times L^2(-r,0)$ and the generator

$$A \begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix} = \begin{pmatrix} \int_{-r}^{0} x_1(t+\theta)\, a(d\theta) \\ \frac{d}{d\theta} x_1(\cdot) \end{pmatrix},$$

with domain $D(A) = \{(x_0, x_1(\cdot)) \in H \,:\, x_1(\cdot) \in H^1(-r,0), x_1(0) = x_0\}$.

State equation and cost:

$$\begin{cases} dz(t) = \left(\int_{-r}^{0} z(t+\theta)\, a(d\theta)\right)\, dt + f(z(t))\, dt + u(t)\, dt + dW_t, & t \in [0,T], \\ z(0) = x_0, & \\ z(\theta) = x_1(\theta), & \theta \in [-r, 0], \end{cases}$$

$$J(u(\cdot)) = \mathbb{E}\int_0^T [l_0(z(t)) + |u(t)|^2]\, dt + \mathbb{E}\phi_0(z(T)).$$

In the space $H = \mathbb{R} \times L^2(-r, 0)$ we define

$$F : H \to H, \quad G : \mathbb{R} \to H, \quad x \in H, \quad l : H \times U \to H, \quad \phi : H \to \mathbb{R},$$

setting

$$F\begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix} = \begin{pmatrix} f(x_0) \\ 0 \end{pmatrix}, \qquad Gu = \begin{pmatrix} u \\ 0 \end{pmatrix}, \qquad x = \begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix},$$

$$l\left(\begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix}, u\right) = l_0(x_0) + |u|^2, \qquad \phi\begin{pmatrix} x_0 \\ x_1(\cdot) \end{pmatrix} = \phi_0(x_0),$$

and we obtain

$$dX_t = AX_t\, dt + F(X_t)\, dt + G\,(u(t)\, dt + dW_t), \qquad X_0 = x \in H,$$

$$J(u(\cdot)) = \mathbb{E}\int_0^T l(X_t, u_t)\, dt + \mathbb{E}\phi(X_T).$$

All the required assumptions are satisfied and the previous results apply.