Optimality problem for infinite dimensional bilinear systems

A. Berrabah, N. Bensalem, F. Pelletier *

Université de Savoie, Département de Mathématiques, Technolac, Campus Scientifique,
73376 Le Bourget du Lac cedex, France

Received 21 January 2006
Available online 18 April 2006

Abstract

For any finite dimensional control system with arbitrary cost, Pontryagin’s Maximum Principle (PMP) [N. Bensalem, Localisation des courbes anormales et problème d’accessibilité sur un groupe de Lie hilber- tien nilpotent de degré 2, Thèse de doctorat, Université de Savoie, 1998. [6]] gives necessary conditions for optimality of trajectories. In the infinite dimensional case, it is well known that these conditions are no more true in general. The purpose of this paper is to establish an “approached” version of PMP for infinite dimensional bilinear systems, with fixed final time and without constraints on the final state. Moreover, if the set of control is contained in a closed bounded convex subset with operators defining its dynamics are compact, or if it is contained in a finite dimensional space, we get an “exact” version of PMP. We also give two applications of these results. The first one deals with sub-Riemannian geometry on nilpotent Hilbertian Lie groups for which we can define a sub-Riemannian distance. The second one deals with heat equation for which we analyse the necessary conditions to give the optimal controls.

© 2006 Elsevier SAS. All rights reserved.

MSC: 22; 34; 47; 53; 93

Keywords: Pontryagin’s Maximum Principle; Infinite dimensional systems; Sub-Riemannian geometry; Heat equation

1. Introduction

The first study about optimal control for infinite dimensional systems has been made by Butkovsky and Lerner [8]. In this context, some generalizations of Pontryagin’s Maximum Prin-
443

ciple (PMP) have been given by Butkovsky for systems governed by integral equations [4], by
Kharatishvili for systems governed by ordinary differential equations [23], and by A.I. Egorov
for nonlinear evolution equations [14,15].

In 1960, Yu.V. Egorov constructs an example showing that PMP is not always true for an
arbitrary controlled system in the infinite dimensional case [16,17]; in this same work, under
some additional conditions, the author set up a PMP by using a generalization of some classical
proof of PMP in finite dimension. This counter-example gives rise to a great number of papers
about optimal time control for infinite dimensional systems (see [1–3,19,20,24,29–33], etc.).

The purpose of this paper is to establish an “approached” version of PMP for infinite dimen-
sional bilinear systems with fixed final time and without constraints on the final state, and when
the set of control is contained in a bounded subset $K$ of an infinite dimensional Hilbert space.
Moreover, if the set of control is contained in a closed bounded convex subset with operators
defining its dynamics are compact, or if it is contained in a finite dimensional space, we get an
“exact” version of PMP. We then give some applications of these results. The first one deals
with sub-Riemannian geometry on nilpotent Hilbertian Lie groups for which we can define a
sub-Riemannian distance. The second one deals with heat equation for which we analyse the
necessary conditions to give the optimal controls.

2. “Approached” Maximum Principle in infinite dimension

Let $E$ and $F$ be two Hilbert spaces, $D$ a dense subspace in $E$ and $[0, T]$ be a fixed time
interval. We denote by $L(D; E)$ (resp. $L(F; E)$) the space of bounded linear operators from $D$
(resp. $F$) to $E$.

Consider the following system:

$$
\begin{cases}
\dot{x}(t) = \Delta x(t) + A(t)u(t) + B(u(t), x(t)), \\
x(0) = x_0,
\end{cases}
$$

where: $\Delta \in L(D; E), A : \mathbb{R}^+ \to L(F; E)$ are a continuous map, $B$ is an element of the space
$L(F \times E; E)$ of bilinear operators from $F \times E$ to $E$ and $x(t) \in E$ is the trajectory starting from
a fixed point $x_0 \in D$ associated to the control $u(t) \in F$.

We assume that $\Delta$ is a generator of strongly continuous semigroup $G(t)$, of course we then
have:

$$
G(0) = \text{Id} \quad \text{and} \quad \|G(t)\| \leq M,
$$

for all $t \in [0, T]$, where $T > 0$ is fixed. Consider $u \in L^2([0, T]; K)$, where $K$ is a bounded subset
of $F$.

With the previous assumptions and [26, pp. 182–183], for all $x_0 \in D$ the system (1) has a unique “mild solution” $x \in C([0, T]; E)$ satisfying the integral equation:

$$
x(t) = G(t)x_0 + \int_0^t G(t-s)(A(s)u(s) + B(u(s), x(s)))\, ds,
$$

according to the Gronwall’s lemma, for all $t \in [0, T]$ we get

$$
\|x(t)\| \leq (M\|x_0\| + M_0 T)e^{M_1 T},
$$

for some positive constants $M_0$ and $M_1$. 
Give us a function $L: L^2([0, T]; F) \times E \to \mathbb{R}$ which is continuous for the weak topology on $L^2([0, T], F)$ and the usual topology on $E$ and a function $\phi: E \to \mathbb{R}$ which is $C^1$, and assume that these functions are bounded on all bounded subset. In a natural way, the function $L$ gives rise to a function, again denoted by $L$, from $F \times E$ as follows: $L(v, x)$ is the value of $L$ on the pair $(u(t), x)$ where $u(t)$ is the constant control with value $v$. This function is continuous with respect to the usual topology on $F \times E$.

We consider the following optimal problem: for a fixed point $x_0 \in D$, find a control $u$ which minimizes the functional

$$\Psi(u) = \int_0^T L(u(t), x_u(t)) \, dt + \phi(x_u(T)), \tag{3}$$

when the control $u$ ranges in $L^2([0, T]; K)$, where $x_u(t)$ still denotes the trajectory starting from $x_0$ and associated with the control $u$.

If we denote by $\inf \Psi$ the infimum of $\Psi$ on $L^2([0, T]; K)$, we will prove the following “approached” version of the maximum principle:

**Theorem 2.1.** For all $\epsilon > 0$, there exists a control $u_\epsilon \in L^2([0, T]; K)$ such that the associated trajectory $x_\epsilon$ satisfies the following properties:

$$\inf \Psi \leq \Psi(u_\epsilon) \leq \inf \Psi + \epsilon \tag{4}$$

and

$$\left\{ A(t)u_\epsilon(t) + B(u_\epsilon(t), x_\epsilon(t)), p_\epsilon(t) \right\} + L(u_\epsilon(t), x_\epsilon(t)) \leq \min_{v \in K} \left\{ A(t)v + B(v, x_\epsilon(t)), p_\epsilon(t) \right\} + L(v, x_\epsilon(t)) + \epsilon, \tag{5}$$

for almost all $t \in [0, T]$ and where $p(t)$ is a mild solution of the adjoint system:

$$\begin{cases}
\dot{p}_\epsilon(t) = -\Delta^* p_\epsilon(t) - (B_{u_\epsilon})^* p_\epsilon(t), \\
p_\epsilon(T) = d\phi(x_\epsilon(T)),
\end{cases}$$

and $B_u$ is the bounded linear operator associated with the bilinear operator $B$ for all $u$. $\Delta^*$ and $(B_u)^*$ are the adjoint of $\Delta$ and $B_u$ respectively.

Now we define

$$y_u(t) = \int_0^t L(u(s), x(s)) \, ds,$$

and consider the new system:

$$\begin{cases}
\dot{x}(t) = \Delta x(t) + A(t)u(t) + B(u(t), x(t)), \\
\dot{y}_u(t) = L(u(t), x(t)), \\
x(0) = x_0, \\
y_u(0) = 0.
\end{cases}$$

Let $g_u: E \times \mathbb{R} \to \mathbb{R}$ be the function defined by

$$g_u(x, y) = \phi(x) + y,$$

for simplicity, we note $y$ (resp. $x$) instead of $y_u$ (resp. $x_u$). Theorem 2.1 can then be expressed in the following way:
Theorem 2.2. For all $\epsilon > 0$, there exists a control $u_\epsilon \in L^2([0, T]; K)$ such that the associated trajectory $(x_\epsilon, y_\epsilon)$ satisfies
\[ g_u(x_\epsilon(T), y_\epsilon(T)) \leq \inf_{u \in K} g_u(x(T), y(T)) + \epsilon, \]
\[ \left\{ A(t)u_\epsilon(t) + B(u_\epsilon(t), x_\epsilon(t)), p_\epsilon(t) \right\} + L(u_\epsilon(t), x_\epsilon(t)) \leq \min_{u \in K} \left\{ A(t)u + B(u, x_\epsilon(t)), p_\epsilon(t) \right\} + L(u, x_\epsilon(t)) + \epsilon, \]
for almost all $t \in [0, T]$ and where $p_\epsilon(t)$ is a mild solution of the adjoint system:
\[ \dot{p}_\epsilon(t) = -\Delta^*p_\epsilon(t) - (Bu_\epsilon)^*p_\epsilon(t), \]
\[ p_\epsilon(T) = d\phi(x_\epsilon(T)). \]

Theorem 2.3. Let be $V$ a complete metric space and $\Phi : V \to \mathbb{R} \cup \{ + \infty \}$ a lower semicontinuous function, with finite values and bounded from below. For all $u \in V$ which satisfies
\[ \inf \Phi \leq \Phi(u) \leq \inf \Phi + \epsilon \text{ and for all } \lambda > 0, \text{ there exists a point } v_\lambda \in V \text{ such that:} \]
\[ \Phi(v_\lambda) \leq \Phi(u), \]
\[ d(v_\lambda, u) \leq \lambda, \]
\[ \Phi(w) > \Phi(v_\lambda) - \frac{\epsilon}{\lambda}d(v_\lambda, w) \text{ for all } w \neq v_\lambda. \]

Before proving Theorem 2.2, we are going to establish some preliminary lemmas. Denote by $U$ the set of controls $u$ which belong to $L^2([0, T]; K)$. On $U$ we define the distance $\delta$ by
\[ \delta(u_1, u_2) = \text{mes}\{ t \in [0, T] : u_1(t) \neq u_2(t) \}. \]
Note that we always have
\[ \|u_1 - u_2\|_{L^2} \leq C\delta(u_1, u_2), \]
for some suitable positive constant $C$.

**Lemma 2.1.** [18] $(U, \delta)$ is a complete metric space.

**Proof.** Let $(u_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in $U$. With formula (8), $u_n$ is a Cauchy in $L^2([0, T]; K)$. There exists $u \in L^2([0, T]; K)$ such that $\|u_n - u\|_{L^2}$ converges to zero. We can extract a subsequence $(u_{n_k}(t))_{k \in \mathbb{N}}$ which converges to $u(t)$ for almost all $t \in [0, T]$. Hence there exists $N \subset [0, T]$ such that $\text{mes}(N) = 0$ and for all $t \in N^c$, $\lim_{k \to \infty} u_{n_k}(t) = u(t)$ in $[0, T]$.

$t \in N^c$ is equivalent to: $\forall \epsilon > 0, \exists k_\epsilon \in \mathbb{N}$, such that $\forall k \geq k_\epsilon, t \in A_{k, \epsilon} = \{ t \in [0, T] : \|u_{n_k}(t) - u(t)\| < \epsilon \}$. We deduce that,
\[ t \in \bigcap_{\epsilon > 0} \bigcap_{m \in \mathbb{N}} A_{k, \epsilon}, \]
therefore
\[ N^c = \bigcap_{\epsilon > 0} \bigcup_{m \in \mathbb{N}} \bigcap_{k \geq m} A_{k, \epsilon}. \]
Which yields

\[ N = \bigcup_{\delta > 0} \bigcap_{m \in \mathbb{N}} \bigcup_{k \geq m} \{ t \in [0, T]: \| u_{nk}(t) - u(t) \| \geq \delta \}. \]

By hypothesis \( \text{mes}(N) = 0 \), consequently \( \text{mes} \left( \bigcap_{m \in \mathbb{N}} \bigcup_{k \geq m} \{ t \in [0, T]: \| u_{nk}(t) - u(t) \| \geq \delta \} \right) = 0 \), for all \( \delta > 0 \).

Now, set \( S_m = \bigcup_{k \geq m} \{ t \in [0, T]: \| u_{nk}(t) - u(t) \| \geq \delta \} \). Since \( S_{m+1} \subset S_m \) for all \( m \in \mathbb{N} \) and \( \text{mes}([0, T]) < +\infty \), then for all \( m \), \( \text{mes}(S_m) < +\infty \) and we have

\[ \text{mes} \left( \bigcap_{m \in \mathbb{N}} S_m \right) = \lim_{m \to \infty} S_m = 0, \]

but \( \{ t \in [0, T]: \| u_{nm}(t) - u(t) \| \geq \delta \} \subset S_m \), hence

\[ \lim_{m \to \infty} \text{mes} \left( \{ t \in [0, T]: \| u_{nm}(t) - u(t) \| \geq \delta \} \right) \leq \lim_{m \to \infty} \text{mes}(S_m) = 0. \]

Finally for all \( \delta > 0 \), we have

\[ \lim_{m \to \infty} \delta(u_{nm}, u) = \lim_{m \to \infty} \text{mes} \left( \{ t \in [0, T]: \| u_{nm}(t) - u(t) \| \geq \delta \} \right) = 0. \]

By definition, the subsequence \( (u_{nk})_{k \in \mathbb{N}} \) converges to \( u \). As the sequence \( (u_n) \) is Cauchy, it converges to \( u \) as a whole with respect to the metric \( \delta \). Which ends the proof. \( \square \)

**Lemma 2.2.** The map \( \Phi : \mathcal{U} \to \mathbb{R} \) defined by

\[ u \mapsto g_u(x(T), y(T)) = \phi(x(T)) + y(T), \]

where \((x(t), y(t))\) is the trajectory associated to \( u(t) \), is continuous.

**Proof.** Let \( (u_n)_{n \in \mathbb{N}} \) be a sequence in \( \mathcal{U} \) which converges to an element \( \bar{u} \) with respect to the metric \( \delta \). Denote by \( ((x_n, y_n))_{n \in \mathbb{N}} \) the sequence of trajectories associated with \( (u_n)_{n \in \mathbb{N}} \) and by \((\bar{x}, \bar{y})\) the trajectory associated with \( \bar{u} \).

For all \( n \in \mathbb{N} \), let us define

\[ A_n = \{ t \in [0, T]: u_n(t) \neq \bar{u}(t) \}. \]

On one hand, for \( t \in [0, T] \) fixed, we have

\[
\| x_n(t) - \bar{x}(t) \|_E \leq \left\| \int_0^t G(t-s)(A(s)u_n(s) - A(s)\bar{u}(s)) \, ds \right\|
+ \left\| \int_0^t G(t-s)(B(u_n(s), \bar{x}(s)) - B(\bar{u}(s), \bar{x}(s))) \, ds \right\|
+ \left\| \int_0^t G(t-s)(B(u_n(s), x_n(s)) - B(u_n(s), \bar{x}(s))) \, ds \right\|.
\] (9)
From which we get

\[ \| x_n(t) - \bar{x}(t) \|_E \leq M \left( C_0 \int_0^t \| u_n(s) - \bar{u}(s) \| \, ds + C_1 \int_0^t \| x_n(s) - \bar{x}(s) \| \, ds \right) \]

\[ \leq M \left( C_0 \int_{[0,T]} \| u_n(s) - \bar{u}(s) \| \, ds + \int_0^t C_1 \| x_n(s) - \bar{x}(s) \| \, ds \right) \]

\[ \leq MC_0 \left( \int_{[0,T]} \| u_n(s) - \bar{u}(s) \| \, ds \right) e^{C_1 T} \quad \text{(Gronwall's lemma)} \]

\[ \leq MC_0 e^{C_1 T} \int_{A_n} \| u_n(s) - \bar{u}(s) \| \, ds \]

\[ \leq 2MC_0 \text{Diam } K e^{C_1 T} \int_{A_n} ds \]

\[ \leq 2MC_0 \text{Diam } K e^{C_1 T} \delta(u_n, \bar{u}). \quad (10) \]

On the other hand,

\[ \left| y_n(t) - \bar{y}(t) \right| \leq \int_0^t \left| L \left( u_n(s), x_n(s) \right) - L \left( \bar{u}(s), \bar{x}(s) \right) \right| \, ds \]

\[ \leq \int_0^T \left| L \left( u_n(s), x_n(s) \right) - L \left( u_n(s), \bar{x}(s) \right) \right| \, ds \]

\[ + \int_0^T \left| L \left( u_n(s), \bar{x}(s) \right) - L \left( \bar{u}(s), \bar{x}(s) \right) \right| \, ds \]

\[ \leq 2 \sup L\delta(u_n, \bar{u}) + \int_0^T \left| L \left( u_n(s), x_n(s) \right) - L \left( u_n(s), \bar{x}(s) \right) \right| \, ds. \quad (11) \]

From Lebesgue’s theorem, we then deduce:

\[ \lim_{n \to \infty} \int_0^T \left| L \left( u_n(s), x_n(s) \right) - L \left( u_n(s), \bar{x}(s) \right) \right| \, ds = 0. \quad (12) \]

Taking into account (10), (11), (12) and the continuity of \( \Phi \), the lemma is proved. \( \square \)

**Proof of Theorem 2.1.** Now, we can apply Theorem 2.3. In consequence, there exists a control \( u_\varepsilon \in \mathcal{U} \) such that:

\[ \Phi(u_\varepsilon) \leq \inf \Phi + \varepsilon^2, \]

\[ \Phi(u) \geq \Phi(u_\varepsilon) - \varepsilon \delta(u, u_\varepsilon), \quad \forall u \in \mathcal{U}. \quad (13) \]
The associated trajectory \((x_\epsilon, y_\epsilon)\) is given by
\[
\begin{align*}
\dot{x}_\epsilon(t) &= \Delta x_\epsilon(t) + A(t)u_\epsilon(t) + B(u_\epsilon(t), x_\epsilon(t)), \\
\dot{y}_\epsilon(t) &= L(x_\epsilon(t), u_\epsilon(t)), \\
x_\epsilon(0) &= x_0, \\
y_\epsilon(0) &= 0.
\end{align*}
\]
(14)

Give us a time \(t_0\) in \([0, T]\) for which we have
\[
\dot{x}_\epsilon(t_0) = \Delta x_\epsilon(t_0) + A(t_0)u_\epsilon(t_0) + B(u_\epsilon(t_0), x_\epsilon(t_0)).
\]

For \(u_0\) in \(K\), we define \(u_\tau \in U\) for all \(\tau \geq 0\) almost everywhere in the following way:
\[
u_\tau(t) = \begin{cases} 
 u_0 & \text{if } t \in [0, T] \cap [t_0 - \tau, t_0[, \\
 u_\epsilon(t) & \text{if } t \not\in [0, T] \cap [t_0 - \tau, t_0[,
\end{cases}
\]
(15)
clearly we have
\[
\delta(u_\epsilon, u_\tau) \leq \tau.
\]

Denote by \(x_\tau\) the trajectory associated to \(u_\tau\) with initial condition \(x_\tau(0) = x_0\).

For all \(t \in [0, T]\) and \(t_0 \in [0, T]\), \(x_\tau\) satisfies the integral equation:
\[
x_\tau(t) = G(t - t_0)x_\tau(t_0) + \int_{t_0}^{t} G(t - s)(A(s)u_\tau(s) + B(u_\tau(s), x_\tau(s))) \, ds.
\]
(16)

Now (13) gives us:
\[
g(x_\tau(T), y_\tau(T)) - g(x_\epsilon(T), y_\epsilon(T)) \geq -\epsilon_\tau,
\]
so we obtain
\[
\frac{d}{d\tau} g(x_\tau(T), y_\tau(T)) \bigg|_{\tau=0} = \frac{d}{d\tau} \phi(x_\tau(T)) \bigg|_{\tau=0} + \frac{d}{d\tau} y_\tau(T) \bigg|_{\tau=0} \geq -\epsilon.
\]

It remains to show the existence of \(\frac{d}{d\tau} \phi(x_\tau(T)) \big|_{\tau=0}\) and also to compute \(\frac{d}{d\tau} \phi(x_\tau(T)) \big|_{\tau=0}\) and \(\frac{d}{d\tau} y_\tau(T) \big|_{\tau=0}\). Since we must have
\[
\frac{d}{d\tau} \phi(x_\tau(t)) \bigg|_{\tau=0} = \left(\frac{d}{d\tau} x_\tau(t) \bigg|_{\tau=0}, \frac{d}{d\tau} \phi(x_\tau(t)) \bigg|_{\tau=0}\right),
\]
(18)

for all \(t \in [0, T]\), we will now look for the existence and computation of \(\frac{d}{d\tau} x_\tau(t) \big|_{\tau=0}\).

We consider three situations:

1st case, \(t < t_0\): in this case, we choose \(\tau\) enough small so that \(t \not\in [t_0 - \tau, t_0[\). The formula (16) gives:
\[
\frac{x_\tau(t) - x_\epsilon(t)}{\tau} = 0,
\]
for all \(t < t_0\), and then \(\frac{d}{d\tau} (x_\tau(t)) \big|_{\tau=0} = 0\).

2nd case, \(t = t_0\): in this case, we compute \(\frac{d}{d\tau} (x_\tau(t_0)) \big|_{\tau=0}\).
\[ x_\tau(t_0) - x_\epsilon(t_0) \]
\[ = \int_0^{t_0} G(t_0 - s) \left[ A(s)u_\tau(s) + B(u_\tau(s), x_\tau(s)) - A(s)u_\epsilon(s) - B(u_\epsilon(s), x_\epsilon(s)) \right] ds \]
\[ = \int_{t_0-\tau}^{t_0} G(t_0 - s) \left[ A(s)u_0 + B(u_0, x_\tau(s)) - A(s)u_\epsilon(s) - B(u_\epsilon(s), x_\epsilon(s)) \right] ds, \]

from continuity property, we have
\[
\lim_{\tau \to 0} \frac{1}{\tau} \int_{t_0-\tau}^{t_0} G(t_0 - s) \left[ A(s)u_0 + B(u_0, x_\tau(s)) \right] ds = A(t_0)u_0 + B(u_0, x_\epsilon(t_0)),
\]
on the other hand
\[
\left\| \frac{1}{\tau} \int_{t_0-\tau}^{t_0} G(t_0 - s) B(u_0, x_\tau(s) - x_\epsilon(s)) \right\| ds \leq K \int_{t_0-\tau}^{t_0} \| x_\tau(s) - x_\epsilon(s) \| ds
\]
\[ = K \| x_\tau(\xi) - x_\epsilon(\xi) \|,
\]
with \( \xi \in [t_0 - \tau, t_0] \), therefore we obtain
\[
\lim_{\tau \to 0} \frac{1}{\tau} \int_{t_0-\tau}^{t_0} G(t_0 - s) \left[ A(s)u_0 + B(u_0, x_\tau(s)) \right] ds = A(t_0)u_0 + B(u_0, x_\epsilon(t_0)).
\]

Now let us show that
\[
\lim_{\tau \to 0} \frac{1}{\tau} \int_{t_0-\tau}^{t_0} G(t_0 - s) \left[ A(s)u_\epsilon(s) + B(u_\epsilon(s), x_\epsilon(s)) \right] ds = A(t_0)u_\epsilon(t_0) + B(u_\epsilon(t_0), x_\epsilon(t_0)),
\]
as, on \([t_0 - \tau, t_0]\) almost everywhere we have
\[
G(t_0 - s) \left[ A(s)u_\epsilon(s) + B(u_\epsilon(s), x_\epsilon(s)) \right] = G(t_0 - s) \left[ \dot{x}_\epsilon(s) - \Delta x_\epsilon(s) \right]
\]
\[ = \frac{d}{ds} \left\{ G(t_0 - s)x_\epsilon(s) \right\},
\]
with formula (16), we obtain
\[
G(t_0 - t)x_\tau(t) = x_\tau(t_0) + \int_{t_0}^{t} G(t_0 - s) \left( A(s)u_\tau(s) + B(u_\tau(s), x_\tau(s)) \right) ds.
\]
From what \( t \to G(t_0 - t)x_\tau(t) \) is absolutely continuous, and then,
\[
\frac{1}{\tau} \int_{t_0-\tau}^{t_0} G(t_0 - s) \left[ A(s)u_\epsilon(s) + B(u_\epsilon(s), x_\epsilon(s)) \right] ds = \frac{1}{\tau} \left[ G(t_0 - s)x_\epsilon(s) \right]_{s=t_0}^{s=t_0-\tau},
\]
As from our assumption we have, \( x_\epsilon \) is derivable at \( t_0 \), it follows that:
\[
\lim_{\tau \to 0} \frac{1}{\tau} \int_{t_0 - \tau}^{t_0} G(t_0 - s) \left[ A(s)u_\epsilon(s) + B(u_\epsilon(s), x_\epsilon(s)) \right] ds
\]
\[
= \left. \frac{d}{ds} \left\{ G(t_0 - s)x_\epsilon(s) \right\} \right|_{s=t_0}
\]
\[
= A(t_0)u_\epsilon(t_0) + B(u_\epsilon(t_0), x_\epsilon(t_0)),
\]
so we have
\[
\frac{d}{d\tau} x_{\tau}(t_0) \bigg|_{\tau=0} = \left( A(t_0)u_0 + B(u_0, x_\epsilon(t_0)) \right) - \left( A(t_0)u_\epsilon(t_0) + B(u_\epsilon(t_0), x_\epsilon(t_0)) \right).
\]

3rd case, \( t > t_0 \): we set
\[
N^\epsilon_{\tau}(t) = \frac{x_{\tau}(t) - x_\epsilon(t)}{\tau}
\]
and
\[
M^\epsilon_0(t) = \lim_{\tau \to 0} N^\epsilon_{\tau}(t),
\]
let us show that this limit exists for all \( t > t_0 \).

From the formula (16), and if we denote by \( B_u \) the bounded linear operator associated with the bilinear operator \( B \) for all \( u \) [7, p. 14], we have for all \( t > t_0 \):
\[
N^\epsilon_{\tau}(t) = G(t - t_0)N^\epsilon_{\tau}(t_0) + \int_{t_0}^{t} G(t - s)B_{u_\epsilon(s)}N^\epsilon_{\tau}(s) ds,
\]
(19)
since \( G(t) \) is a strongly continuous, by passing to the limit and from Lebesgue’s theorem we deduce:
\[
\lim_{\tau \to 0} N^\epsilon_{\tau}(t) = G(t - t_0) \lim_{\tau \to 0} N^\epsilon_{\tau}(t_0) + \int_{t_0}^{t} G(t - s)B_{u_\epsilon(s)} \lim_{\tau \to 0} N^\epsilon_{\tau}(s) ds,
\]
so we have
\[
M^\epsilon_0(t) = G(t - t_0)M^\epsilon_0(t_0) + \int_{t_0}^{t} G(t - s)B_{u_\epsilon(s)}M^\epsilon_0(s) ds,
\]
(20)
\( N^\epsilon_{\tau}(t) \) and \( M^\epsilon_0(t) \) are two “mild solutions” giving by the formulas (19) and (20), for the following system, with the initial condition \( z(t_0) \):
\[
\dot{z}(t) = \Delta z(t) + B_{u_\epsilon(t)}z(t).
\]
(21)
Consequently
\[
\begin{cases}
\dot{N}^\epsilon_{\tau}(t) = \Delta N^\epsilon_{\tau}(t) + B_{u_\epsilon(t)}N^\epsilon_{\tau}(t), \\
N^\epsilon_{\tau}(t_0) = x_{\tau(t_0)} - x_\epsilon(t_0),
\end{cases}
\]
(22)
and
\[
\begin{cases}
\dot{M}^\epsilon_0(t) = \Delta M^\epsilon_0(t) + B_{u_\epsilon(t)}M^\epsilon_0(t), \\
M^\epsilon_0(t_0) = \lim_{\tau \to 0} N^\epsilon_{\tau}(t_0).
\end{cases}
\]
(23)
As from the result in [5, p. 47], we deduce that if \( M^\epsilon_0(t_0) = \lim_{\tau \to 0} N^\epsilon_{\tau}(t_0) \) then \( M^\epsilon_0(t) = \lim_{\tau \to 0} N^\epsilon_{\tau}(t) \) exists and is solution of Eq. (21), for all \( t > t_0 \). In (16), we set \( p_\epsilon(T) = d\phi(x_\epsilon(T)) \), where \( p_\epsilon(t) \) is the “mild solution” of the following system:
\[
\begin{align*}
\dot{p}_\epsilon(t) &= -\Delta^* p_\epsilon(t) - (B_{u_\epsilon(t)})^* p_\epsilon(t), \\
p_\epsilon(T) &= d\phi(x_\epsilon(T)).
\end{align*}
\] (24)

This solution exists and is unique from [26, p. 41 and p. 185].

Now consider the problem of regularity on the systems (23) and (24). If the mild and strong solutions coincide (see [22], for definitions of mild and strong solutions), then for all \(t \in [t_0, T]\), we have

\[
\frac{d}{dt} \langle p_\epsilon(t), M_0^\epsilon(t) \rangle = \langle \dot{p}_\epsilon(t), M_0^\epsilon(t) \rangle + \langle p_\epsilon(t), \Delta M_0^\epsilon(t) + B_{u_\epsilon(t)} M_0^\epsilon(t) \rangle = 0.
\]

Therefore, for all \(t \in [t_0, T]\),

\[
\langle p_\epsilon(t), M_0^\epsilon(t) \rangle = \text{cst}.
\]

In particular

\[
\langle p_\epsilon(t_0), M_0^\epsilon(t_0) \rangle = \langle p_\epsilon(T), M_0^\epsilon(T) \rangle.
\] (25)

Thus, we distinguish two cases:

1st case: \(u \in C^1([t_0, T]; K)\).

If \((M_0^\epsilon(t_0), p_\epsilon(T)) \in D \times D\), then the mild and strong solutions coincide (see [12] or [22]), so we get (25). More general, for \((M_0^\epsilon(t_0), p_\epsilon(T)) \in E \times E\), by density of \(D\) in \(E\), there exist two sequences \((M_0^{\epsilon,n}(t_0), p_\epsilon^n(T)) \in D \times D\) which converge to \((M_0^\epsilon(t_0), p_\epsilon(T))\) in \(E\), when \(n\) tends to \(+\infty\). From the result in [5, p. 47], we have \((M_0^{\epsilon,n}(t), p_\epsilon^n(t)) \in D \times D\) converge to \((M_0^\epsilon(t), p_\epsilon(t))\) in \(E\), \(\forall t \geq t_0\) a.e., when \(n\) tends to \(+\infty\). It results that \((M_0^{\epsilon,n}(T), p_\epsilon^n(t_0)) \in D \times D\) converge to \((M_0^\epsilon(T), p_\epsilon(t_0))\) in \(E\). Therefore,

\[
\langle p_\epsilon^n(t_0), M_0^{\epsilon,n}(t_0) \rangle = \langle p_\epsilon^n(T), M_0^{\epsilon,n}(T) \rangle
\]

by passing to the limit, we get (25).

2nd case: \(u \in L^2([t_0, T]; K)\).

By density, there exist a sequence \(u_n \in C^1([t_0, T]; K)\) converging to \(u\) in \(L^2([t_0, T]; K)\). It is easy to prove that \(M_0^{\epsilon,n}(t, u_n, M_0^\epsilon(t_0))\) and \(p_\epsilon^n(t, u_n, p_\epsilon(T))\) converge to \(M_0^\epsilon(t, u, M_0^\epsilon(t_0))\) and \(p_\epsilon(t, u, p_\epsilon(T))\) in \(C([t_0, T]; K)\) respectively (see proof of Lemma 2.2). As in the first case, we obtain (25).

Now, \(M_0^\epsilon(t_0) = \frac{d}{d\tau} x_\tau(t_0)|_{\tau=0}\) and \(M_0^\epsilon(T) = \frac{d}{d\tau} x_\tau(T)|_{\tau=0}\). Consequently (18) becomes,

\[
\frac{d}{d\tau} \phi(x_\tau(T))|_{\tau=0} = \left\{ \frac{d}{d\tau} x_\tau(T)|_{\tau=0}, p_\epsilon(T) \right\} = \left\{ \frac{d}{d\tau} x_\tau(t_0)|_{\tau=0}, p_\epsilon(t_0) \right\} = \left\{ A(t_0)u_0 + B(u_0, x_\epsilon(t_0)) - (A(t_0)u_\epsilon(t_0) + B(u_\epsilon(t_0), x_\epsilon(t_0))) \right\}, p_\epsilon(t_0)\}
\]

Finally compute \(\frac{d}{d\tau} y_\tau(T)|_{\tau=0}\).
\[ y_\tau(T) = \int_0^T L(u_\tau(s), x_\tau(s)) \, ds \]
\[ = \int_{t_0-\tau}^{t_0} L(u_\epsilon(s), x_\epsilon(s)) \, ds + \int_{t_0}^T L(u_0, x_\tau(s)) \, ds + \int_{t_0}^T L(u_\epsilon(s), x_\epsilon(s)) \, ds \]
\[ = y_\epsilon(t_0 - \tau) + \int_{t_0-\tau}^{t_0} L(u_0, x_\tau(s)) \, ds + \int_{t_0}^T L(u_\epsilon(s), x_\epsilon(s)) \, ds \]
\[ = y_\epsilon(t_0) - \tau \frac{d}{dt} y_\epsilon(t_0) + \tau L(u_0, x_\epsilon(t_0)) + \mathcal{O}(\tau^2) + \int_{t_0}^T L(u_\epsilon(s), x_\epsilon(s)) \, ds \]
\[ = y_\epsilon(t_0) + \tau \left[ L(u_0, x_\epsilon(t_0)) - L(u_\epsilon(t_0), x_\epsilon(t_0)) \right] + \int_{t_0}^T L(u_\epsilon(s), x_\epsilon(s)) \, ds + \mathcal{O}(\tau^2), \]
\[ = y_\epsilon(t_0) + \int_{t_0}^T L(u_\epsilon(s), x_\epsilon(s)) \, ds + \tau \left[ L(u_0, x_\epsilon(t_0)) - L(u_\epsilon(t_0), x_\epsilon(t_0)) \right] + \mathcal{O}(\tau^2), \]
\[ = y_\epsilon(T) + \tau \left[ L(u_0, x_\epsilon(t_0)) - L(u_\epsilon(t_0), x_\epsilon(t_0)) \right] + \mathcal{O}(\tau^2), \quad (26) \]

where
\[ \frac{d}{d\tau} y_\tau(T) \bigg|_{\tau=0} = L(u_0, x_\epsilon(t_0)) - L(u_\epsilon(t_0), x_\epsilon(t_0)). \]

Finally we obtain
\[ \frac{d}{d\tau} g(x_\tau(T), y_\tau(T)) \bigg|_{\tau=0} = \left[ A(t_0)u_0 + B(u_0, x_\epsilon(t_0)), p_\epsilon(t_0) \right] + L(u_0, x_\epsilon(t_0)) \]
\[ - \left[ A(t_0)u_\epsilon(t_0) + B(u_\epsilon(t_0), x_\epsilon(t_0)), p_\epsilon(t_0) \right] \]
\[ - L(u_\epsilon(t_0), x_\epsilon(t_0)) \]
\[ \geq -\epsilon. \quad (27) \]

As the inequality (27) is true almost everywhere on \([0, T]\) and as \(u_0\) is any element of \(K\), the Theorem 2.1 is proved. \(\Box\)

3. “Exact” Maximum Principle in infinite dimension

In the particular cases when the set of control is contained in a closed bounded convex subset with operators defining its dynamics are compact, or when it is contained in a finite dimensional space, we get the following version of a maximum principle:
Theorem 3.1. Among all controls $u \in L^2([0, T]; K)$, where $K$ is a compact subset of a finite dimensional space $F$ (resp. $u \in L^2([0, T]; K)$, where $K$ is a closed bounded convex subset, $B$ and $A(t)$ are compact for all $t \in [0, T]$), there exists a control $\bar{u}$ which minimizes the functional $\Psi$ and moreover, $\bar{u}$ satisfies the following relation for almost all $t \in [0, T]$:

\[
\left\{ A(t)\bar{u}(t) + B(\bar{u}(t), \bar{x}(t)), \bar{p}(t) \right\} + L(\bar{u}(t), \bar{x}(t)) \leq \min_{v \in K} \left\{ A(t)v + B(v, \bar{x}(t)), \bar{p}(t) \right\} + L(v, \bar{x}(t)),
\]

(28)

where $\bar{x}(t)$ is the trajectory associated to $\bar{u}(t)$ and where $\bar{p}(t)$ is a mild solution of the adjoint system:

\[
\begin{cases}
    \dot{\bar{p}}(t) = -\Delta^* \bar{p}(t) - (B\bar{u})^* \bar{p}(t), \\
    \bar{p}(T) = d\phi(\bar{x}(T)).
\end{cases}
\]

Proof. At first assume that $\dim F = k$.

For $\epsilon = 1/n$ take a control $u_n$ with property described in Theorem 2.1. As the sequence $(u_n)$ is bounded in $L^2([0, T]; K)$, without loss of generality, we can suppose that $u_n$ converges weakly to a point $\bar{u} \in L^2([0, T]; K)$. In this case, after taking a sub-sequence, $u_n(t)$ converges to $\bar{u}(t)$ almost everywhere.

First choose a fixed basis in $F$, we can write:

\[
A(t)u = u_1 A_1(t) + \cdots + u_k A_k(t) \quad \text{and} \quad B(u, x) = u_1 B_1(x) + \cdots + u_k B_k(x),
\]

where $A_i(t)$ are vectors in $E$ and $B_i$ are bounded operator of $E$, $i = 1, \ldots, k$.

Denote by $u^i_n$ and by $\bar{u}^i$, $i = 1, \ldots, k$, the components of $u_n$ and $\bar{u}$ respectively.

By using same arguments as in the proof of Lemma 2.2 in relations (9) and (10), we obtain

\[
\| x_n(t) - \bar{x}(t) \|_E \leq \sum_{i=1}^k \left\| \int_0^t \left[ (u^i_n(s) - \bar{u}^i(s)) G(t - s) \left( A_i(s) + B_i(\bar{x}(s)) \right) \right] ds \right\| 
\]

\[
+ \left\| \int_0^t G(t - s) B(u_n(s), x_n(s) - \bar{x}(s)) ds \right\| 
\]

\[
\leq Me^{CT} \sum_{i=1}^k \left\| \int_0^t \left[ (u^i_n(s) - \bar{u}^i(s)) G(t - s) \left( A_i(s) + B_i(\bar{x}(s)) \right) \right] ds \right\|. \quad (29)
\]

We fix the time $t$. Since the sequence $u_n$ converges weakly to $\bar{u}$, then from [21], for all continuous functions $\psi : [0, t] \rightarrow E$, we have

\[
\lim_{n \rightarrow \infty} \left\| \int_0^t \left[ u^i_n(s) - \bar{u}^i(s) \right] \psi(s) ds \right\| = 0,
\]

for all $i = 1, \ldots, k$.

It follows that $x_n$ converges to $\bar{x}(t)$ on $[0, T]$ . By using continuity property of $\Psi$ in relation (4), we obtain the property of minimality of $\bar{u}$. From the same type of previous arguments we can show that $p_n(t)$ converges also to $\bar{p}(t)$ on $[0, T]$. 

Now, consider a fixed measurable set $M$ in $[0, T]$ and a point $v$ in $K$. From Theorem 2.1, we have

$$
\begin{equation}
\int_M \left( \sum_{i=1}^{k} \left( (v_i - u^i_n(s)) (A_i(s) + B_i(x_n(s))) , p_n(s) \right) \right) ds + \int_M \left( L(v, x_n(s)) - L(u_n(s), x_n(s)) \right) ds + \frac{T}{n} \geq 0.
\end{equation}
$$

By using continuity relative to the weak topology on $L^2([0, T]; F)$, the usual topology on $E$ and Lebesgue’s theorem, we have the following convergence properties:

$$
\lim_{n \to \infty} \int_M \left[ \sum_{i=1}^{k} \left( A_i(s)v_i + B_i(x_n(s))v_i, p_n(s) \right) \right] ds + \left( L(v, x_n(s)) - L(u_n(s), x_n(s)) \right) \right] ds = \int_M \left[ \sum_{i=1}^{k} \left( A_i(s)v_i + B_i(x(s))v_i, \bar{p}(s) \right) \right] ds + \left( L(v, x(s)) - L(\bar{u}(s), \bar{x}(s)) \right) \right] ds.
$$

On the other hand

$$
\left| \sum_{i=1}^{k} \int_M \left[ \left( A_i(s)u^i_n(s) + B_i(x_n(s))u^i_n(s), p_n(s) \right) - \left( A_i(s)\bar{u}_i(s) + B_i(\bar{x}(s))\bar{u}_i(s), \bar{p}(s) \right) \right] ds \right|
\leq K \int_M \sum_{i=1}^{k} \left| \left( B_i(x_n(s)), p_n(s) \right) - \left( B_i(\bar{x}(s)), \bar{p}(s) \right) \right| ds
$$

$$
+ \sum_{i=1}^{k} \int_M \left| \left( \bar{u}^i_n(s) - \bar{u}^i(s) \right) (A_i(s) + B_i(\bar{x}(s))), \bar{p}(s) \right| ds
$$

$$
+ \sum_{i=1}^{k} \int_M \left| \left( A_i(s)u^i_n(s), p_n(s) \right) - \left( A_i(s)u^i_n(s), \bar{p}(s) \right) \right| ds.
$$

We deduce that:

$$
\lim_{n \to \infty} \sum_{i=1}^{k} \int_M u^i_n(s)(A_i(s) + B_i(x_n(s)), p_n(s)) ds = \sum_{i=1}^{k} \int_M \bar{u}^i(s)(A_i(s) + B_i(\bar{x}(s)), \bar{p}(s)) ds.
$$

It follows from (30) that:

$$
0 \leq \int_M \sum_{i=1}^{k} \left[ (v_i - \bar{u}_i(s)) (A_i(s) + B_i(\bar{x}(s))), \bar{p}(s) \right] ds + L(v, x(s)) - L(\bar{u}(s), \bar{x}(s)) \right] ds.
$$

As the inequality above is true for all measurable subset $M$ of $[0, T]$, it follows that:

$$
\{ A(t)\bar{u}(t) + B(\bar{u}(t), \bar{x}(t)), \bar{p}(t) \} - L(\bar{u}(t), \bar{x}(t))
\leq \{ A(t)v + B(v, x(t)), \bar{p}(t) \} + L(v, x(t)),
$$

for almost $t \in [0, T]$ and for any $v \in K$. 
Now assume that $K$ is a closed bounded convex subset of $F$ and suppose that the operators $B$ and $A(t)$ defining the dynamics of control are compact for all $t \geq 0$.

For $\epsilon = 1/n$, take a control $u_n$ with property described in Theorem 2.1. As the sequence $(u_n)$ is bounded in $L^2([0, T]; K)$, without loss of generality, we can suppose that $u_n$ converges weakly to a point $\bar{u} \in L^2([0, T]; K)$.

We then have

$$\|x_n(t) - \bar{x}(t)\|_E \leq \left\| \int_0^t G(t - s)[A(s)(u_n(s) - \bar{u}(s)) + B(u_n(s), x_n(s)) - B(\bar{u}(s), \bar{x}(s))] ds \right\|$$

$$\leq \left\| \int_0^t G(t - s)A(s)(u_n(s) - \bar{u}(s)) ds \right\|$$

$$+ M \left\| \int_0^t B(u_n(s), x_n(s)) - B(u_n(s), \bar{x}(s)) ds \right\|$$

$$+ \left\| \int_0^t G(t - s)(B(u_n(s), \bar{x}(s)) - B(\bar{u}(s), \bar{x}(s))) ds \right\|$$

$$\leq Me^{CT} \left\| \int_0^t G(t - s)A(s)(u_n(s) - \bar{u}(s)) ds \right\|$$

$$+ \left\| \int_0^t G(t - s)(B(u_n(s), \bar{x}(s)) - B(\bar{u}(s), \bar{x}(s))) ds \right\|. \quad (31)$$

Now, we will show that

$$\int_0^t G(t - s)A(s)(u_n(s) - \bar{u}(s)) ds,$$

converges to 0 in $E$, when $u_n$ converges weakly to $\bar{u}$ in $L^2([0, T]; K)$. For that, we consider the following lemma:

**Lemma 3.1.** For all $T > 0$, the operator

$$\Pi u(.) = \int_0^t G(\cdot - s)A(s)(u(s)) ds$$

from $L^2([0, T]; K)$ into $L^2([0, T]; E)$ is compact.
Proof. We can check that the adjoint operator $L^* \pi$ from $L^2([0, T]; E)$ into $L^2([0, T]; K)$ is given by:

$$L^* \pi (v(.)) = A^* (.) \int_0^T G^* (s - .) v(s) \, ds,$$

this operator is compact, since $A^* (.)$ is compact and the operator $v(.) \to \int_0^T G^* (s - .) v(s) \, ds$ is continuous and then $L^* \pi$ is compact.

With same arguments, we show that, $\int_0^T G(t - s) (B(u_n(s), \bar{x}(s)) - B(\bar{u}(s), \bar{x}(s))) \, ds \to 0$, when $n \to + \infty$. We deduce that, $x_n(t)$ converges to $\bar{x}(t)$ in $E$ for all $t \in [0, T]$. From continuity property of $\Psi$ in (4), this leads to the minimality of $\bar{u}$.

With same arguments as we used for the inequality (9), after taking a sub-sequence, we can show that $p_n(t)$ converges to $\bar{p}(t)$ almost everywhere in $[0, T]$.

Now, consider a fixed measurable set $M$ in $[0, T]$ and a point $v$ in $K$. From Theorem 2.1, we have,

$$\int_M \left( (A(s)v - B(\bar{u}(s), \bar{x}(s))) + B(B(u_n(s), x(n(s))), p_n(s)) \right) ds + \int_M \left( L(v, x(n(s))) - L(u_n(s), x(n(s))) \right) ds + \frac{T}{n} \geq 0. \tag{32}$$

By using continuity to the weak topology on $L^2([0, T]; F)$, usual topology on $E$ and from Lebesgue’s theorem, we have the following convergence properties:

$$\lim_{n \to \infty} \int_M \left[ (A(s)v + B(v, x(n(s))), p_n(s)) + (L(v, x(n(s))) - L(u_n(s), x(n(s)))) \right] ds$$

$$= \int_M \left[ (A(s)v + B(v, \bar{x}(s)), \bar{p}(s)) + (L(v, \bar{x}(s)) - L(\bar{u}(s), \bar{x}(s))) \right] ds.$$

On the other hand

$$\left| \int_M \left[ (A(s)u_n(s) + B(u_n(s), x(n(s))), p_n(s)) - (A(s)\bar{u}(s) + B(\bar{u}(s), \bar{x}(s)), \bar{p}(s)) \right] ds \right|$$

$$\leq \left| \int_M \left( A(s)u_n(s) - A(s)\bar{u}(s), \bar{p}(s) \right) ds \right| + \left| \int_M \left( A(s)u_n(s), p_n(s) - \bar{p}(s) \right) ds \right|$$

$$+ \left| \int_M \left( B(u_n(s) - \bar{u}(s), x), p_n(s) \right) ds \right| + \left| \int_M \left( B(\bar{u}(s), \bar{x}(s)), p_n(s) - \bar{p}(s) \right) ds \right|$$

$$+ \left| \int_M \left( B(u_n(s), x(n(s) - \bar{x}(s))), p_n(s) \right) ds \right|.$$

Since $B$ and $A(t)$ are compact, we deduce from the same type of arguments as in Theorem 3.1 that,
\[
\langle A(t)u(t) + B(u(t), x(t)), p(t) \rangle - L(u(t); x(t)) - L(v; x(t)) + \langle A(t)v + B(v, x(t)), p(t) \rangle,
\]
for almost \( t \in [0, T] \) and for any \( v \in K \) and then Theorem 3.1 is proved. \( \Box \)

4. Application in sub-Riemannian geometry

Consider \( A \in L(F; E) \) and \( B \in L(F \times E; E) \). For \( u \in F \), we will denote by \( A(u) \) the image \( A(u) \) and by \( B_u \) the linear operator \( L(E, E): x \to B(u, x) \). Let \( \{f_i: i \in \mathbb{N}\} \) be a fixed Hilbertian basis of \( F \). We set

\[
X_i(x) = A f_i + B f_i x.
\]

We will denote by \( \mathcal{F} \) the bilinear distribution spanned by \( \{X_i: i \in \mathbb{N}\} \). Consider the associated system (\( \Sigma \)):

\[
\dot{x} = Au + B(u, x).
\]

For a given bilinear distribution \( \mathcal{F} \), with the previous notations, for all horizontal curves \( \gamma : [0, T] \to E \), that is tangent to \( \mathcal{F} \), there exists a control \( u : [0, T] \to F \) such that,

\[
\dot{\gamma} = Au + B(u, \gamma).
\]

On the opposite to finite dimensional sub-Riemanniann theory (see [25]) in this context, to any control \( u \), we cannot associate a horizontal curve. However, if \( A \) and \( B \) are Hilbert–Schmidt operators, and if \( \mathcal{F} \) is the associated distribution, to any control \( u \in L^2([0, T], F) \) and to any \( x_0 \in E \) there exists an unique associated horizontal curve \( \gamma \) such that \( \gamma(0) = x_0 \) and \( \dot{\gamma} = Au + B(u, \gamma) \), where \( L^2([0, T], F) \) is identified to the space \( L^2([0, T], l^2(\mathbb{N})) \) via the Hilbertian basis \( \{f_i\}_{i \geq 1} \) of \( F \) (see [9,10,13]). In fact, after changing the parametrization if necessary, we can always assume that all horizontal curves are defined on \([0, 1]\).

When \( A \) and \( B \) are Hilbert–Schmidt operators, the associated distribution \( \mathcal{F} \) is called an Hilbert–Schmidt distribution. In the whole paragraph, \( \mathcal{F} \) will be an Hilbert–Schmidt distribution, associated to given fixed Hilbert–Schmidt operators \( A \) and \( B \).

Then, for any horizontal curve \( \gamma \) we can define its length \( \mathcal{L}(\gamma) \) by

\[
\mathcal{L}(\gamma) = \int_0^1 \|u(s)\| \, ds,
\]

where \( u \in L^2([0, 1], F) \) is the unique control associated to \( \gamma \) and where \( \|\cdot\| \) denotes the Hilbertian norm on \( F \). Given two points \( x_0 \) and \( x_1 \) of \( E \), as in sub-Riemannian geometry in finite situation, we can look for the optimal problem:

\[\text{(P) minimize } \mathcal{L} \text{ among all horizontal curves } \gamma \text{ such that } \gamma(0) = x_0 \text{ and } \gamma(1) = x_1.\]

It is well known that for an infinite dimensional Riemannian Hilbertian manifold, this problem can have no solution, although the set of curves which join two given points \( \gamma(0) = x_0 \) and \( \gamma(1) = x_1 \) is not empty (see [18] for instance). Of course, the same is true in general sub-Riemannian geometry in infinite dimensional situation and, moreover, the set of horizontal curves which join two given points \( \gamma(0) = x_0 \) and \( \gamma(1) = x_1 \) can be empty. However, in the previous situation we can prove:
Property 4.1. Let $F$ be an Hilbert–Schmidt distribution and $A(x_0)$ the set of points $x \in E$ for which there exists an horizontal curve $\gamma : [0, 1] \to E$ with $\gamma (0) = x_0$ and $\gamma (1) = x$. The set $A(x_0)$ is closed and moreover, if $x_1 \in A(x_0)$, there exists a minimal length horizontal curve, that is a minimum for the functional $L$ on the set of horizontal curves $\gamma$ such that $\gamma (0) = x_0$ and $\gamma (1) = x_1$.

Proof. Let $S_R$ be the ball of radius $R$ in $F$. Consider a horizontal curve $\gamma : [0, 1] \to E$ whose associated control $u$ belongs to $L^2([0, 1], S_R)$. Set $\sigma (t) = \int_0^t \|u(s)\| \, ds$. The map $t \mapsto \sigma (t)$ is a nondecreasing and continuous map from $[0, 1]$ into $[0, L]$ if $L = L(\gamma)$. Moreover, if for some $t < t'$ we have $\sigma (t) = \sigma (t')$ then $u$ is zero almost everywhere on $[t, t']$ and so $\gamma (s) = \gamma (t)$ for all $s \in [t, t']$. As we have $\sigma (t') - \sigma (t) \leq R(t' - t)$, $\sigma$ is a Lipschitz function and then it is differentiable almost everywhere and its derivative is equal to $\|u\|$. Denote by $C$ the union of closed intervals on which $\sigma$ is constant and by $O$ its complementary in $[0, 1]$. Consider the map $\tau : [0, L] \to [0, 1]$ defined by $\tau (s) = t$ if and only if $t$ is the infimum of the set $\{t' : \sigma (t') = s\}$. So we have $\sigma \circ \tau = id$ and $\tau$ is inverse map of the map which is the restriction of $\sigma$ to each connected component of $O$. So $\tau$ is differentiable almost everywhere and we have $\frac{d\tau}{ds} = \frac{1}{\|u\|}$ almost everywhere.

Consequently, if we set $v(s) = \tau (Ls)$, the curve $\tilde{\gamma} = \gamma \circ v$ is a reparametrization of $\gamma$ whose associated control is $v(s) = L \frac{u(\gamma(s))}{\|u(\gamma(s))\|}$ with norm $\|v(s)\| = L$ almost everywhere. Obviously, we have $L(\gamma) = L(\tilde{\gamma}) = L$.

We will say that $\tilde{\gamma}$ is geometric parametrization of $\gamma$.

Let $x_1$ be a point of the closure $\overline{A}(x_0)$ of $A(x_0)$ and consider a sequence $\gamma_n : [0, 1] \to E$ of horizontal curves with origin $\gamma_n(0) = x_0$ and such that the sequence of ends $z_n = \gamma_n(1)$ converges to $x_1$. Denote by $L_{\gamma_n}$ the lower-limit of the real sequence $L(\gamma_n)$ and by $L_{x_1}$ the lower bound of the set of real numbers $L(\gamma_n)$ for all sequences $\gamma_n$ with the previous properties.

For given $\epsilon > 0$, denote by

$$H_\epsilon = \{ u \in F/L_{x_1} - \epsilon \leq \|u\| \leq L_{x_1} + 2\epsilon \},$$

and by

$$F(u) = \int_0^1 \|u(t)\| \, dt + \|x_1 - \gamma (1)\|^2,$$

if $\gamma$ is the trajectory with origin $x_0$ and which is associated to a control $u$. For all $u \in L^2([0, 1], H_\epsilon)$, we have:

$$F(u) \geq L_{x_1} - \epsilon.$$

On the other hand, it follows from the definition of $L_{x_1}$ that there exists a geometric parameterized curve $\gamma$ whose length satisfies the inequality:

$$L_{x_1} - \epsilon \leq L(\gamma) \leq L_{x_1} + \epsilon,$$

and such that,

$$\|x_1 - \gamma (1)\|^2 \leq \epsilon.$$

If we denote by $u$ the associated control of $\gamma$, then $u \in L^2([0, 1], H_\epsilon)$ and

$$L_{x_1} - \epsilon \leq F(u) \leq L_{x_1} + 2\epsilon.$$
It follows that the lower bound \( \inf F \) of the set of values of \( F \) on \( L^2([0, 1], H_\varepsilon) \) satisfies,

\[ L_{x_1} - \varepsilon \leq \inf F \leq L_{x_1} + 2\varepsilon. \]

From Theorem 2.1, we conclude that there exists an horizontal curve \( \gamma_\varepsilon \) associated to a control \( u_\varepsilon \in L^2([0, 1], H_\varepsilon) \) such that:

\[
\begin{align*}
L_{x_1} - \varepsilon &\leq L(\gamma_\varepsilon) + \|x_1 - \gamma_\varepsilon(1)\|^2 \leq L_{x_1} + 3\varepsilon, \\
(Au_\varepsilon(t) + Bu_\varepsilon(t), \gamma_\varepsilon(t)) &+ p_\varepsilon(t) + \|u_\varepsilon(t)\| \\
&\leq \min_{v \in H_\varepsilon} \{(Av + B(v, \gamma_\varepsilon(t)), p_\varepsilon(t)) + \|v\| + \varepsilon, \\
&\quad \|x_1 - \gamma_\varepsilon(1)\| \}
\end{align*}
\]

(33)

where \( p_\varepsilon(t) \) is a solution of the adjoint system:

\[
\begin{align*}
\dot{p}_\varepsilon(t) &= -(B_{u_\varepsilon})^* p_\varepsilon(t), \\
p_\varepsilon(T) &= 2\|x_1 - \gamma_\varepsilon(1)\|.
\end{align*}
\]

(34)

If we take \( \varepsilon = 1/n \), we construct a sequence of horizontal curves \( \gamma_n \) whose length tends to \( L_{x_1} \), whose end tends to \( x_1 \). As the associated control \( u_n \) is bounded and by taking subsequence if necessary, we can suppose that the sequence \( (u_n) \) converges weakly to some \( v \). In fact, as \( u_n \in H_{1/n} \), we have \( \|v\|_{L^2} \leq L_{x_1} \). Denote by \( \psi \) the horizontal curve associated to the control \( v \) and whose origin is \( x_0 \). From the relation (10), we deduce that \( \gamma_n(1) \) converges to \( \psi(1) \) and then that \( \gamma(1) = x_1 \). So, from the definition of \( L_{x_1} \), we must have \( L\psi \geq L_{x_1} \). On the other hand, by Schwartz inequality, we have \( L\psi \leq \|v\|_{L^2} \). Finally we conclude that \( \|v\|_{L^2} = L\psi = L_{x_1} \) and in fact, \( u_n \) converges strongly to \( v \).

On the other hand, \( p_n(t) \) also converges almost everywhere to the adjoint vector \( p(t) = 0 \), which is solution of (34) with control \( v \). By passing to the limit in (33), for almost every \( t \in [0, 1] \), we must have \( \|v(t)\| = L_{x_1} \) almost everywhere, which means that \( \gamma \) has a geometric parametrization. \( \square \)

Let \( G \) be an Hilbertian Lie group and \( G \) be its Lie algebra. If we denote by \([ , \) the Lie bracket, the center of \( G \) is the greatest subspace \( Z \) such that \([ Z, G ] = 0 \).

For some integer \( k \geq 1 \) set \( G^k = [G, G^{k-1}] \) and \( G^0 = G \). Recall that a Lie algebra is nilpotent if there exists \( k \geq 1 \) such that \( G^k = [0] \). A nilpotent Lie algebra \( G \) is of degree \( r \) if \( G^r = [0] \) with \( G^{r-1} \neq 0 \). An Hilbertian Lie group is nilpotent of degree \( r \) if and only if its Lie algebra is also nilpotent of degree \( r \).

Now, let be \( G \) a nilpotent Hilbertian Lie group of degree \( r \) connected and simply connected and let be \( Z \) the center of its Lie algebra \( G \), of course, we have \( Z \subset G^1 \).

As \( Z \) is a closed Lie sub-algebra of \( G \) then:

\[
G = F \oplus Z,
\]

where \( F = Z^{\perp} \) is the orthogonal of \( Z \) according to the scalar product defined on \( G \). The Lie bracket induces a bilinear skew-symmetric map \( \Lambda : G \times G \to G \), which satisfies the Jacobi’s identity, that is:

\[
\Lambda(\Lambda(u, v), w) + \Lambda(\Lambda(v, w), u) + \Lambda(\Lambda(w, u), v) = 0 \quad \text{for all } u, v, w \in G,
\]

and we have \( \operatorname{Im} \Lambda \subset Z \). If \( G \) is of degree 2 then, the Jacobi’s identity for \( \Lambda \) is trivially satisfied. Conversely, given a such skew-symmetric operator \( \Lambda \) on an Hilbert space \( G \) defines a unique structure of Lie nilpotent algebra of degree 2 on \( G \).

As \( G \) is connected and simply connected, the exponential map \( \exp : G \to G \) is a \( C^\infty \) diffeomorphism. So, this map allows us to identify \( G \) to \( G \) as topological space. In this chart,
we have \( d_x \exp(u) = u + \frac{1}{2} [x, u] \) for \( u \in F \) and \( x \in G \), there exists an unique bilinear operator \( B \in L(F \times G; F) \) defined by:
\[
\{ v, B(u, x) \} = \{ [u, v], x \} = \{ \Lambda(u, v), x \} \quad \text{for all} \ v \in F.
\]
The left invariant distribution \( \mathcal{F} \) on \( G \) spanned by \( F \) is exactly \( d \exp(F) \). In fact, \( \mathcal{F} \) is the \textit{bilinear distribution} associated to the operators \( A = \text{Id}_F \) and previous \( B \). In particular, if \( \Lambda \) is an Hilbert–Schmidt operator \( G \) then \( B \) is also an Hilbert–Schmidt operator and so the previous distribution is a \textit{bilinear Hilbert–Schmidt distribution}; in this situation, we will say that \( G \) is a \textit{nilpotent Hilbert–Schmidt Lie group of degree 2}. With the previous notations, let \( \mathcal{H} \) be the closed Lie sub-algebra spanned by \( F \) in \( G \). Of course, we have,
\[
\mathcal{H} = F \oplus \overline{\text{Im} \Lambda}.
\]
We will denote by \( H \) the closed connected Lie sub-group of \( G \) whose Lie algebra is \( \mathcal{H} \). The left invariant distribution on \( G \) associated to \( \mathcal{H} \) is integrable and the associated foliation is in fact the partition \( \{ gH, g \in G \} \) of \( G \). We then have:

**Theorem 4.1.** Let \( \mathcal{F} \) be the left invariant distribution defined on a nilpotent Hilbert–Schmidt Lie group \( G \) of degree 2 which is connected and simply connected. For all \( g \in G \), the set \( \mathcal{A}(g) \) is the closed Hilbertian submanifold \( gH \) of \( G \). Moreover, the map \( d : \mathcal{A}(g) \times \mathcal{A}(g) \to \mathbb{R} \) defined by
\[
\quad d(x, y) = \inf \{ \mathcal{L}(\gamma), \, \gamma : [0, 1] \to G \text{ horizontal such that } g(0) = x, \, \gamma(1) = y \},
\]
is a distance on \( \mathcal{A}(g) \). In particular, if \( \overline{\text{Im} \Lambda} = \mathbb{Z} \), we have \( \mathcal{A}(g) = G \) and \( d \) is a distance on \( G \).

**Proof.** Let be \( g \in G \), from [9, Theorem 3.1], \( \mathcal{A}(g) \) is a dense set in \( gH \) and from Proposition 4.1, it follows that we have \( \mathcal{A}(g) = gH \). As the minimum of the length is always reached on \( \mathcal{A}(g) \), (Proposition 4.1), with classical arguments we can show that:
\[
\quad d(x, y) = \inf \{ \mathcal{L}(\gamma), \, \gamma : [0, 1] \to G \text{ horizontal such that } \gamma(0) = x, \, \gamma(1) = y \},
\]
defines a distance on \( \mathcal{A}(g) \). \( \square \)

5. Application to the heat equation

We will apply Theorem 3.1 in the context of classical system derived by a heat equation.

Consider a metal bar with length \( l = 1 \) which, at time \( t_0 = 0 \), is heated at a point \( x_0 \in ]0, 1[ \). We want to study the evolution of the temperature \( \phi(x, t) \) at position \( x \) at time \( t \), when the temperature is constant at \( x = 0 \) and \( x = 1 \), zero for instance (Dirichlet’s conditions).

Denote by \( \Omega \) the interval \([0, 1]\), by \( E = L^2(\Omega, \mathbb{R}) \) the state space and by \( \Delta = \frac{\partial^2}{\partial x^2} \) the classical Laplacian, whose domain is \( D(\Delta) = H^2([0, 1]) \cap H^1_0([0,1]) \).

In \( E \), consider the possible model for the temperature distribution:
\[
\begin{aligned}
\dot{\phi}(x, t) &= \Delta \phi(x, t) + u(t) \phi(x, t) + u(t) F(t) \gamma(x), \\
\phi(x, 0) &= \alpha(x), \quad x \in ]0, 1[, \\
\phi(0, t) &= \phi(1, t) = 0, \quad t \in ]0, T[,
\end{aligned}
\tag{35}
\]
where \( F : \mathbb{R} \to \mathbb{R} \) is a bounded analytic function not identically zero and where \( \gamma(.) \) and \( \alpha(.) \) are given points in \( D(\Delta) \).

With the previous notations, the system (35) is bilinear and we have,
Now, we will give the solution of the system (35). We can choose the following Hilbert basis: \( \{e_k(x)\}_{k \geq 1} \) are eigenfunctions of \( \Delta \), with Dirichlet’s conditions, that is:

\[
\Delta e_k = \lambda_k e_k \quad \text{on} \quad \Omega \quad \text{and} \quad e_k = 0 \quad \text{on the frontier of} \quad \Omega,
\]

where \( \lambda_k = -k^2 \pi^2 \) are the eigenvalue of \( \Delta \), for all \( k \geq 1 \). In these conditions, there exists a unique solution \( \phi(x, t) \) to (35) (see [11, p. 205]). By using the Hilbert basis of \( E \), we can write any solution of system (35) in the following way:

\[
\phi(x, t) = \sum_{k \geq 1} a_k(t) e_k(x),
\]

with:

\[
\begin{align*}
\dot{a}_k(t) &= \lambda_k a_k(t) + u(t) a_k(t) + F(t) u(t) \langle \gamma, e_k \rangle_{L^2(\Omega)}, \\
\dot{a}_k(0) &= \langle \alpha, e_k \rangle_{L^2(\Omega)},
\end{align*}
\]

for all \( k \geq 1 \) and all \( t \in [0, T] \).

The solution to (36) can be readily obtained by a “separation of variables” approach. If we set,

\[
\gamma_k = \langle \gamma, e_k \rangle_{L^2(\Omega)}, \quad \alpha_k = \langle \alpha, e_k \rangle_{L^2(\Omega)},
\]

we obtain:

\[
a_k(t) = e^{(\int_0^T u(s) ds + \lambda_k t)} \left( \alpha_k + \gamma_k \int_0^t F(\tau) u(\tau) e^{-\int_0^\tau u(s) ds + \lambda_k \tau} d\tau \right).
\]  

(37)

Now, look the solution for adjoint system:

\[
\begin{align*}
\dot{p}(x, t) &= -\Delta^* p(x, t) - u(t) p(x, t), \\
p(x, T) &= d g(\phi(x, T)), \quad x \in [0, 1], \\
p(0, t) &= p(1, t) = 0, \quad t \in [0, T],
\end{align*}
\]

(38)

where \( g \) is a \( C^1 \) map from \( E \) to \( \mathbb{R} \) and lower bounded.

Take the function \( \phi(.) \mapsto g(\phi(.)) = \frac{1}{2} \| \phi(.) - \beta(.) \|^2_{L^2(\Omega)} \), where \( \beta \in E \) is a given function.

Then we have

\[
d g(\phi(x, T)) = \phi(x, T) - \beta(x).
\]

Since \( \Delta^* = \Delta \), as for the system (35), the solution of the adjoint system (38) can be written in the following way:

\[
p(x, t) = \sum_{k \geq 1} p_k(t) e_k(x),
\]

where

\[
p_k(t) = e^{\int_0^T u(s) ds + \lambda_k T + \int_0^T u(s) ds + \lambda_k (T-t)}
\]

\[
\times \left[ \alpha_k + \gamma_k \int_0^T F(\tau) u(\tau) e^{-\int_0^\tau u(s) ds + \lambda_k \tau} d\tau \right] - \beta_k e^{\int_0^T u(s) ds + \lambda_k (T-t)},
\]

(39)

with \( \lambda_k = -k^2 \pi^2 \), \( \forall k \geq 1 \).
On the other hand, if we set,
\[ S_t : E \to E, \]
\[ \sum_{k \geq 1} \gamma_k e_k(x) \mapsto \sum_{k \geq 1} \gamma_k e^{-tk^2/2\pi^2} e_k(x), \]
then \( S_t \in \mathcal{L}(E) \) is a strongly continuous semigroup on \( E \), whose generator is \( \Delta \) and we also have,
\[ \| S(t) \|_{\mathcal{L}(E)} \leq 1 \text{ for all } t \in [0, +\infty[. \]

5.1. Optimality conditions

According to the previous presentation, we can apply Theorem 3.1 to this situation: there exists a control \( u \in L^2([0, T]; K) \) such that the associated trajectory \( \phi \) satisfies:
\[ g(\phi(., T)) = \inf g(\phi(., T)), \]
\[ \{ A(t)u(t) + u(t)\phi(., t), p(., t) \} = \min_{v \in K} \{ A(t)v + v\phi(., t), p(., t) \}, \quad (40) \]
where \( p \) is a solution of the adjoint system:
\[ \begin{cases} \dot{p}(x, t) = -t(\Delta + u(t))p(x, t), \\ p(x, T) = dg(\phi(x, T)). \end{cases} \]

Now, we will analyze the conditions (40), which give information on optimal controls.
For all control \( u \), we set:
\[ H_u(t) = \{ A(t)\cdot + \phi(., t), p(., t) \} = \{ F(t)\gamma(\cdot) + \phi(., t), p(., t) \}, \]
so, we consider three situations:

1st case, \( H_u(t) > 0 \): since \( u(t) \in [a, b] \), we always have \( u(t) = a \).

2nd case, \( H_u(t) < 0 \): in the same way, we obtain \( u(t) = b \).

We can conclude, that in the set \( \{ t \in [0, T] / H_u(t) \neq 0 \} \), the optimal control \( u(t) \) takes constant values \( a \) or \( b \) only. The optimal control \( u(t) \), \( 0 \leq t \leq T \), is a piecewise constant function, which takes values \( a \) or \( b \) and has intervals of constancy according to the sign of \( H_u(t) \).

3rd case, \( K(t) = 0 \): for this case, we will give some conditions for which \( H_u(t) \) is zero on \( [0, T] \). Those conditions give more information on a such control. We have the following lemma:

**Lemma 5.1.**

1. \( H_u \) is derivable almost everywhere and we have
\[ \frac{d}{dt} \{ H_u \}(t) = e^{T-t} u(s) ds \frac{d}{dt} \{ F f_u \}(t), \]
where \( f_u(t) = \langle S(T-t)\gamma, \phi(., T) - \beta \rangle_{L^2(\Omega)} \) on \( [0, T] \).

2. Let \( u \) be a control.
   i. If \( \phi(., T) = \beta(\cdot) \) then, \( H_u(t) \) is zero on \( [0, T] \).
   ii. If \( \phi(., T) \neq \beta(\cdot) \) and if \( H_u(t) \) is zero on a sub-interval \( [t_0, t_1] \) then, \( H_u \) is zero on \( [0, T] \) if and only if we have the relation
\[ \{ S(T-t)\gamma, \phi(., T) - \beta \rangle_{L^2(\Omega)} = 0, \quad (41) \]
on \( [0, T] \).
iii. If $\phi(., T) \neq \beta(.)$ and if the condition (41) is not satisfied on all sub-interval $[0, T]$, then $H_u(t)$ is zero on a compact subset $C$ of $[0, T]$ whose interior is empty.

**Proof.**

1. On $[0, T]$, we have

\[
H_u(t) = F(t)[\gamma, p(., t)]_{L^2(\Omega)} + \langle \phi(., t), p(., t) \rangle_{L^2(\Omega)},
\]

after derivation, we obtain

\[
\frac{d}{dt}\{H_u\}(t) = \dot{F}(t)[\gamma, p(t)]_{L^2(\Omega)} - F(t)[\Delta \gamma, p(t)]_{L^2(\Omega)}.
\]

From the expression (39), we have:

\[
p(., t) = e^{\int_t^T u(s)ds} q(., t),
\]

where $q(., t) = \sum_{k=1}^{\infty} e^{\lambda_k(T-t)}(a_k(T) - \beta_k) e_k(.)$ and $a_k(T)$ is the component of $\phi(., T)$ on $e_k$.

So we conclude that on $[0, T]$, (43) is equivalent to:

\[
\frac{d}{dt}\{H_u\}(t) = e^{\int_t^T u(s)ds} \left[ \dot{F}(t) f_u(t) + F(t) \dot{f}_u(t) \right],
\]

where

\[
f_u(t) = \sum_{k=1}^{\infty} e^{\lambda_k(T-t)}(a_k(T) - \beta_k) \gamma_k = [\gamma, p(., t)] e^{-\int_t^T u(s)ds}.
\]

In fact, $f_u$ admits an extension on $]-\infty, T]$ defined by $f_u(t) = \langle S(T-t)\gamma, \phi(., T) - \beta \rangle_{L^2(\Omega)}$; so the first part of the lemma is proved.

2i. If $\phi(., T) = \beta(.)$, then $dg(\phi(., T)) = 0$ and consequently the associated adjoint vector $p$ is identically zero on $[0, T]$ and the same follows for $H_u$.

ii. Now suppose that $H_u(t) = 0$ on $[t_0, t_1]$. From part 1, we will have,

\[
F(t) f_u(t) = \text{cte} \quad \text{on} \quad [t_0, t_1].
\]

Since $f_u(t) = \langle S(T-t)\gamma, \phi(., T) - \beta \rangle_{L^2(\Omega)}$, on $]-\infty, T]$ it follows that $\lim_{t \to -\infty} f_u(t) = 0$. From the relation (45), analyticity property of $F$ and since $F$ is not identically zero, we can conclude that $f_u$ is an analytic function on $]-\infty, T]$. So, $H_u(t)$ is identically zero on $[0, T]$ and the relation (45) is true on $]-\infty, T]$. Using the fact that $\lim_{t \to -\infty} f_u(t) = 0$ and that $F$ is bounded, the constant in (45) must be zero and then $f_u(t)$ is identically zero.

Finally,

\[
\langle S(T-t)\gamma, \phi(., T) - \beta \rangle_{L^2(\Omega)} = 0 \quad \text{for all} \quad t \in [0, T].
\]

Trivially, the converse is true.

iii. This case is a consequence of the preceding case ii. \square

**Summary**

Now, by using Lemma 5.1, we can analyse the case where $H_u(t) = 0$:

1. $\phi(., T) = \beta$, in this situation, $H_u$ is identically zero and $g(\phi(., T)) = 0$ and it is also a minimum of $g$.
2. $\phi(., T) \neq \beta$, but $\phi(., T)$ satisfies the condition (41) on a sub-interval of $[0, T]$, then $H_u(t)$ is zero on $[0, T]$. 
3. \( \phi(., T) \neq \beta \), and the condition (41) is not satisfied on all sub-interval \([0, T]\), then \( H_u(t) \) is zero on a compact subset \( C \subset [0, T] \) whose interior is empty.

**Remark 5.1.** Let \( \beta \in E \) be such that \( \langle \gamma, \beta \rangle_{L^2(\Omega)} \neq 0 \) and such that \( \beta \) does not belongs to the accessibility set of \( \alpha \) but belongs to its closure. For all \( \epsilon > 0 \) there always exists a control \( u_\epsilon \) which satisfies the properties of Theorem 2.1 and for which, the associated function \( H_{u, \epsilon} \), can be zero, at most on a compact set on \([0, T]\) whose interior is empty. Indeed, let \( (u_n) \) a sequence of controls which satisfy the conclusions of Theorem 2.1.

By taking a subsequence, we can suppose that \( \phi_n(., T) \) converges weakly to \( \beta(., T) \). It follows that:

\[
\langle \gamma, \phi_n(., T) - \beta(., T) \rangle_{L^2(\Omega)} \neq 0.
\]

5.2. Example of synthesis

In this section, we look for the situation where \( \gamma = \sum_{k=1}^{N} \gamma_k e_k \) for fixed \( N \geq 1 \). For any control \( u \), we will denote by \( f_u(., T) \) the function defined by:

\[
f_u(t) = \langle S(T-t)\gamma, \phi(., T) - \beta \rangle_{L^2(\Omega)}.
\]

In these conditions, the function \( f_u \) is always an analytic function. We have two situations:

1st case: there exists a trajectory \( \hat{\phi}(., t) \) which is a solution of (35) and whose associated control \( \hat{u} \) satisfies \( f_u(t) = 0 \) on \([0, T]\). In this situation: either \( \hat{\phi}(., T) = \beta \) and then such a trajectory trivially satisfies the conclusions of Theorem 2 and \( \hat{\phi}(., T) \) is the minimum of \( g \); or \( \hat{\phi}(., T) \neq \beta \), in this last situation, the associated adjoint vector \( \hat{p}(., t) \) is orthogonal to \( \gamma \) for all \( t \in [0, T] \).

If moreover we have \( \| \hat{\phi}(., T) - \beta(.) \|_{L^2(\Omega)} < \epsilon \), then the hypothesis of Theorem 3.1 are satisfied, but we have no more information on a such control.

**Example 5.1.** Consider the case where \( F(t) = \cos t \), \( \alpha = \gamma = e_k \) and \( \beta = e_j \) with \( j \neq k \) for all fixed integers \( k \geq 1 \) and \( j \geq 1 \) (likewise, we can study the case where \( j = k \)). We have then \( f_u(t) = a_k(T)e^{\lambda_k(t-T)} \). According to those information, \( \hat{\phi}(., T) = a_k(T)e_k(., T) \) and then \( \hat{\phi}(., T) \neq \beta(., T) \).

This first case occurs if and only if \( a_k(T) = 0 \), or equivalently, the associated control \( \hat{u} \) is a solution of the equation:

\[
\int_{0}^{T} \cos(t) \hat{u}(t) e^{-\int_{0}^{t} \hat{u}(s) ds + \lambda_k v} dt = -1. \tag{46}
\]

\( a_k(T) = 0 \) is equivalent to \( \hat{\phi}(., T) = 0 \), and then \( \| g(\hat{\phi}(., T)) - \beta \|^2 = \frac{1}{2} \).

Or, \( g(\hat{\phi}(., T)) = \frac{1}{2} a_k^2(T) + \frac{1}{2} \), consequently \( \hat{\phi}(., T) = 0 \) is the minimum of \( g \).

**Remark 5.2.** Note for instance, that this case do not occur if \( a \geq 0 \) and \( T \leq \frac{\pi}{2} \).

2nd case: there exists no trajectory which is a solution of (35) and such that the associated function \( f_u \) is identically zero. In this situation consider a trajectory \( \hat{\phi} \) which satisfies the conclusions of Theorem 2.2. Then, from Lemma 5.1 it follows that the associated function \( H_{\hat{u}} \) can be zero on a finite or infinite number of points and then, the associated control \( \hat{u} \) is piecewise constant and take the value \( a \) (resp. \( b \)) on each sub-interval on which the sign of \( H_{\hat{u}} \) is positive (resp. negative). Note that, this case can occur if we are in the case of Remark 5.2.
Example 5.2. (Continuation of Example 5.1) Again, consider the situation where \( F(t) = \cos t \), \( \alpha = \gamma = e_k \) and \( \beta = e_j \) with \( j \neq k \) for all fixed integers \( k \geq 1 \) and \( j \geq 1 \) (likewise we can study the case where \( j = k \)). We have \( f_{\bar{a}}(t) = a_k(T) e^{\lambda_k(T-t)} \) where,
\[
a_k(T) = e^{(\int_0^T \bar{u}(s) ds + \lambda_k T)} \left( 1 + \int_0^T \cos(t)\bar{u}(t) e^{-(\int_0^t \bar{u}(s) ds + \lambda_k t)} dt \right).
\]
Since by assumption, \( a_k(T) \neq 0 \) then we have \( f_{\bar{a}}(T) \neq 0 \). The derivative of \( Ff_{\bar{a}} \) is equal to:
\[
\frac{d}{dt}[Ff_{\bar{a}}](t) = e^{-\int_t^T \bar{u}(s) ds} \frac{d}{dt}[H_{\bar{a}}](t) = a_k(T) e^{\lambda_k(T-t)} (-\sin t + k^2 \pi^2 \cos t),
\]
therefore we have,
\[
\frac{d}{dt}[H_{\bar{a}}](t) = e^{\int_t^T \bar{u}(s) ds + \lambda_k(T-t)} (-\sin t + k^2 \pi^2 \cos t).
\]
For \( T \leq \frac{\pi}{2} \) and \( k = 1 \) for instance, we will study the sign of \( \frac{d}{dt}[H_{\bar{a}}](t) \). Several sub-cases occur:

1st sub-case, \( a_1(T) > 0 \): in this situation,
\[
\text{sign} \left( \frac{d}{dt}[H_{\bar{a}}](t) \right) = \text{sign}(-\sin t + \pi^2 \cos t).
\]
(a) If \( t = \frac{\pi}{2} \), then either \( a_1(\frac{\pi}{2}) > 0 \) and the optimal control is \( \bar{u} = a \); or \( a_1(\frac{\pi}{2}) < 0 \) and the optimal control is \( \bar{u} = b \).
(b) If \( t \neq \frac{\pi}{2} \), then \( \tan(t) = \pi^2 \), and then \( t = \arctan(\pi^2) = t_0 \approx 1.471 \), and then we have two possibilities:

(b.1) if \( t \in [0, t_0[ \), then \( H_{\bar{a}}(t) \) is creasing, because \( H_{\bar{a}}(0) > 0 \) and so that the optimal control is \( \bar{u}(t) = a \),

(b.2) if \( t \in ]t_0, T[ \), then \( H_{\bar{a}}(t) \) is decreasing, so that \( H_{\bar{a}}(T) \geq H_{\bar{a}}(t) \geq H_{\bar{a}}(t_0) \), we have then two another possibilities:

(b.2.1) \( H_{\bar{a}}(T) > 0 \); in this case, the optimal control is \( \bar{u}(t) = a \) on \( ]t_0, T[ \).

(b.2.2) \( H_{\bar{a}}(T) < 0 \); then, \( H_{\bar{a}} \) changes sign at most once on \( ]t_0, T[ \). Consequently, either \( \bar{u}(t) = b \) on \( ]t_0, T[ \); or there exists \( \tau \in ]t_0, T[ \) such that \( \bar{u}(t) = a \) on \( ]t_0, \tau[ \) and \( \bar{u}(t) = b \) on \( ]\tau, T[ \).

2nd sub-case, \( a_1(T) < 0 \): We can study this case, in the same way as in the first sub-case.

References


