PERGAMON

Nonlinear Analysis 52 (2003) 703–723

Nonlinear Analysis

www.elsevier.com/locate/na

Relaxed controls for a class of strongly nonlinear delay evolution equations☆

X. Xiang[a,*], P. Sattayatham[b], Wei Wei[a]

[a]*Department of Mathematics, Guizhou University, Guiyang, Guizhou 550025, People's Republic of China*

[b]*School of Mathematics, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand*

Received 30 April 2001; accepted 22 November 2001

Abstract

Relaxed controls for a class of strongly nonlinear delay evolution equations are investigated. Existence of solutions of strongly nonlinear delay equations is proved and properties of original and relaxed trajectories are discussed. The existence of optimal relaxed controls and relaxation result are also presented. For illustration, two examples are given.

MSC: 49A20

Keywords: Delay evolution equations; Nonlinear monotone operator; Evolution triple; Measurable multifunction; Relaxation

1. Introduction

It is well-known that in the study of the existence of optimal controls the convexity condition (or more precisely Cesari property) is very important. Many authors working on variational and optimal control problem convexified finite-dimensional control system for existence of optimal controls. This problem (called relaxation) has already been studied in literature (see [7,8,5]). For infinite-dimensional systems, some authors discussed a series of questions on relaxation for semilinear or some nonlinear evolution systems (see [1,11]). However, to our knowledge, few authors studied the problem

☆ Partly Supported by Thailand Research Fund, Grand No. BRG/11/2544.

* Corresponding author. Tel.: +86-851362-0057.

E-mail addresses: xlxiang@tmail.gzu.edu.cn (X. Xiang), pairote@ccs.sut.ac.th (P. Sattayatham), wei66@yahoo.com (W. Wei).

PII: S0362-546X(02)00127-X

on relaxed controls of systems governed by delay evolution equations. Particularly for strongly nonlinear delay evolution equations.

Semilinear delay evolution equations have been studied by many authors including us (see [9,12,10]). Most results are concerned with semigroup. In this paper, we investigate relaxed controls for a class of strongly nonlinear delay evolution equations. Existence of solutions for such a class of strongly nonlinear delay evolution equations is proved and both result and method are different from others. In addition, the properties of the set of solutions for corresponding control systems are discussed. By introducing measure-valued controls, the original control systems are convexified and relaxed control systems are obtained. Under some reasonable assumptions we prove that the set of original trajectories is dense in the set of relaxed trajectories in appropriate space (see Theorem 3.C). The approximation result showing the relation between original control systems and relaxed control systems is very significant for control theory and application.

For relaxed system, the existence of optimal controls is obtained under some regularity hypotheses on cost functional. Finally, we show that the optimal values of original and relaxed problems are equal, namely relaxation theorem. Two examples are presented for illustration.

This paper is organized as follows. In Section 2, existence of solutions of strongly nonlinear delay systems and some properties of set solutions for corresponding control systems are presented. We give relaxed systems and prove approximate results on relaxed trajectories in Section 3. Section 4 contributes to existence of optimal relaxed control and relaxation theorem. Two examples concerning delay partial differential equations are given in last section.

2. Nonlinear delay evolution equations and controlled system

Let $V \hookrightarrow H \hookrightarrow V^*$ be evolution triple and the embedding $V \hookrightarrow H$ be compact. The system model considered here is based on this evolution triple (see Chapter 23 of [13]).

Let $\langle x, y\rangle$ denote the pairing of an element $x \in V^*$ and an element $y \in V$. If $x, y \in H$, then $\langle x, y\rangle = (x, y)$, where $(,)$ is the scalar product on H. The norm in any Banach space X will be denoted by $\|\cdot\|_X$. We denote by $C = C([-r,0],H)$ the Banach space of all continuous maps from $[-r,0]$ into H with the usual supremum norm, here $r \geqslant 0$.

Let $0 < t \leqslant T < +\infty$, $I_t \equiv (0,t)$, $I \equiv (0,T)$, and let $p,q \geqslant 1$ such that $1/p+1/q=1$ and $2 \leqslant p < \infty$. For p,q satisfying the preceding conditions, it follows from reflexivity of V that both $L_p(I_t, V)$ and $L_q(I_t, V^*)$ are reflexive Banach Spaces (see Theorem 1.1.17 of [2]). The pairing between $L_p(I_t, V)$ and $L_q(I_t, V^*)$ is denoted by $\langle\langle,\rangle\rangle_t$. In particular, for $t = T$, we use $\langle\langle,\rangle\rangle = \langle\langle,\rangle\rangle_T$. Clearly, for $u, v \in L_2(I,H)$, $\langle\langle u, v\rangle\rangle = ((u, v))$, where $((\cdot,\cdot))$is the scalar product in Hilbert space $L_2(I,H)$.

Define

$$W_{pq} = \{x\colon x \in L_p(I,V),\ \dot{x} \in L_q(I,V^*)\},$$

$$\|x\|_{W_{pq}} = \|x\|_{L_p(I,V)} + \|\dot{x}\|_{L_q(I,V^*)},$$

where $\dot{x}$ denotes the derivative of x in the generalized sense. $\{W_{pq}, \|.\|_{W_{pq}}\}$ is a Banach space and the embedding $W_{pq} \hookrightarrow C(I,H)$ is continuous. If the embedding $V \hookrightarrow H$ is compact, the embedding $W_{pq} \hookrightarrow L_p(I,H)$ is also compact (see Problem 23.13b of [13]).

Consider the following basic initial value problem

$$(2.1)\begin{cases} \dot{x}(t) + A(t,x(t)) = g(t,x(t),x_t), & t\in(0,T),\\ x(t)=\phi(t), & -r \leqslant t \leqslant 0,\end{cases}$$

where for given $t\in[0,T]$ and $x\in C([-r,T],H)$, $x_t\colon [-r,0] \to H$ is defined as $x_t(\theta)=x(t+\theta)$, $x_t\in C$, $\theta\in[-r,0]$.

We will need the following hypotheses on the data of problem (2.1)

(A) $A: I\times V \to V^*$ is an operator such that
  (1) $t \to A(t,x)$ is measurable;
  (2) $x \to A(t,x)$ is monotone and hemicontinuous; i.e.,

$$\langle A(t,x_1)-A(t,x_2), x_1-x_2\rangle \geqslant 0 \quad \forall x_1,x_2\in V,\ t\in I;$$

$$A(t,x+sy)\xrightarrow{w} A(t,x) \quad \text{in } V^*\ \forall x,y\in V \quad \text{as } s\to 0;$$

  (3) There exist constants $c_1>0$, $c_2\geqslant 0$, $c_3\geqslant 0$ and a nonnegative function $c_4(\cdot)\in L_q(I)$, such that

$$\langle A(t,x),x\rangle \geqslant c_1\|x\|_V^p - c_2 \quad \text{for all } x\in V,\ t\in I,$$

$$\|A(t,x)\|_{V^*} \leqslant c_4(t) + c_3\|x\|_V^{p-1} \quad \text{for all } x\in V,\ t\in I;$$

(G) $g: I\times H\times C \to H$ is an operator such that
  (1) $t\to g(t,\xi,\eta)$ is measurable,
  $(\xi,\eta)\to g(t,\xi,\eta)$ is continuous
  (2) There exist constants $\alpha,\beta\geqslant 0$ and a nonnegative function $h(\cdot)\in L_2(I)$ such that

$$\|g(t,\xi,\eta)\|_H \leqslant h(t) + \alpha\|\xi\|_H^{k-1} + \beta\|\eta\|_C^{2/q}$$

  where $2\leqslant k < p$.

Remark 1. Define $\bar{A}(t,x)=A(t,x-x_0)$ for some $x_0\in V$. It is easy to check that $\bar{A}$ satisfies assumption (A).

Consider the following problem

$$\dot{x}(t)+A(t,x(t)) = g(t,x(t),x_t) \quad 0\leqslant t\leqslant T,$$

$$x(t)=\phi(t) \quad -r\leqslant t\leqslant 0, \tag{1}$$

where $\phi\in C([-r,0],H)$, $\phi(0)\in V$. By Remark 1, we can assume $\phi(0)=0$ and $A(t,0)=0$ without loss of generality. Define

$$W_{pq}([-r,T]) = \{x\colon x\in C([-r,T],H),\ x|_I\in W_{pq}\}.$$

Seek a function $x \in W_{pq}([-r,T])$ such that (1) is satisfied in weak sense (see Chapter 30 of [13]). Since the trajectories of the system belong to $W_{pq}([-r,T])$ and $W_{pq}([-r,T]) \hookrightarrow C([-r,T],H)$ (see Proposition 23.23 of [13]), the initial condition $x_0 = \phi$ makes sense.

Furthermore, define

$$W_{pq}^0 = \{x \in W_{pq}([-r,T]),\ x(t) = \phi(t),\ -r \leqslant t \leqslant 0\},$$

where $\phi \in C([-r,0],H)$ is the initial data.

Theorem 2.A. *Under assumptions* (A) *and* (G), *problem* (1) *has a solution* $x \in W_{pq}^0$.

Proof. (1) Set

$$B = \{y \mid y \in C([0,T],H),\ y(0) = 0\}.$$

Obviously, B is a Banach space. For any $y \in B$, we define $\hat{y}:[-r,T] \to H$ by

$$\hat{y}(t) = \begin{cases} \phi(t) & \text{for } t \in [-r,0], \\ y(t) & \text{for } t \in [0,T]. \end{cases}$$

The operator F is defined on B by letting $y = Fx$ be a solution of the following Cauchy problem:

$$\begin{cases} \dot{y}(t) + A(t,y(t)) = g(t,x(t),\hat{x}_t) & t \in I, \\ y(0) = 0. \end{cases}$$

By assumption (G), $G(x)(t) = g(t,x(t),\hat{x}_t)$ is measurable and $G(x)(\cdot) \in L_2(I,H) \subset L_q(I,V^*)$. F is well defined and $y \in W_{pq} \hookrightarrow C(I,H)$ (see Theorem 30.A of [13]). Hence F maps B into itself.

(2) $F: B \to B$ is continuous.

Suppose $x_n \to x$ in B as $n \to \infty$. This means

$$\sup_{0 \leqslant t \leqslant T} \|x_n(t) - x(t)\|_H \to 0,$$

and

$$\|(\hat{x}_n)_t - \hat{x}_t\|_C \to 0$$

uniformly with respect to $t \in [0,T]$ as $n \to \infty$. Hence, there exists a constant $M > 0$ such that

$$\|\hat{x}_n\|_{C([-r,T],H)} \leqslant M \quad \text{and} \quad \|\hat{x}\|_{C([-r,T],H)} \leqslant M.$$

By virtue of assumption (G),

$$G(x_n)(t) \to G(x)(t) \quad \text{in } H$$

for a.e. $t \in I$ as $n \to \infty$ and there exists a constant $M_1 > 0$ such that

$$\|G(x_n)(t)\|_H \leqslant h(t) + M_1 \quad \text{and} \quad \|G(x)(t)\|_H \leqslant h(t) + M_1.$$

It follows from majorized convergence principle that

$$G(x_n) \to G(x) \quad \text{in } L_2(I,H)$$

as $n \to \infty$.

For $0 \leqslant t \leqslant T$, $Fx_n = y_n$ and $Fx = y$ satisfy the following equations, respectively:

$$\dot{y}_n(t) + A(t, y_n(t)) = G(x_n)(t),$$

$$\dot{y}(t) + A(t, y(t)) = G(x)(t).$$

We have

$$\dot{y}_n(t) - \dot{y}(t) + A(t, y_n(t)) - A(t, y(t)) = G(x_n)(t) - G(x)(t). \tag{2}$$

Integrating by parts and using the assumptions and Cauchy inequality, one can obtain that

$$\begin{aligned}\frac{1}{2}\|y_n(t) - y(t)\|_H^2 &\leqslant \|G(x_n) - G(x)\|_{L_2(I_t,H)}\|y_n - y\|_{L_2(I_t,H)}\\ &\leqslant \frac{1}{2}\|G(x_n) - G(x)\|_{L_2(I,H)}^2 + \frac{1}{2}\int_0^t \|y_n(\tau) - y(\tau)\|_H^2\,d\tau.\end{aligned}$$

Thanks to Gronwall's Lemma, it is easy to show that

$$y_n \to y \quad \text{in } B \quad \text{as } n \to \infty.$$

(3) F is a compact operator on B.

Let $\{x_n\}$ be a bounded sequence in B. That is, there is a constant $M_2 > 0$ such that

$$\|x_n\|_{C([0,T],H)} \leqslant M_2.$$

Again, by assumption (G), there exist constants $M_3, M_4 > 0$ such that

$$\|G(x_n)(t)\|_H \leqslant h(t) + M_3 \quad \text{and} \quad \|G(x_n)\|_{L_2(I,H)} \leqslant M_4.$$

$y_n = Fx_n$ is a solution of the following equation:

$$\dot{y}_n(t) + A(t, y_n(t)) = G(x_n)(t). \tag{3}$$

Integrating by parts in (3) and using assumption (A)(3), one can obtain

$$\frac{1}{2}\|y_n(t)\|_H^2 + C_1\|y_n\|_{L_p(I_t,V)}^p \leqslant \|G(x_n)\|_{L_2(I_t,H)}\|y_n\|_{L_2(I_t,H)} + C_2.$$

It follows from Cauchy inequality that there exist constants $\gamma > 0$ and $K > 0$ such that

$$\frac{1}{2}\|y_n(t)\|_H^2 + \gamma\|y_n\|_{L_p(I_t,V)}^p \leqslant K\|G(x_n)\|_{L_2(I,H)}^q + C_2.$$

Hence $\{y_n\}$ is bounded in $C(I,H) \cap L_p(I,V)$. Again by assumption (A3), (G3), and Eq. (3), we get $\{\dot{y}_n\}$ is bounded in $L_q(I,V^*)$. Therefore $\{y_n\}$ is bounded in W_{pq}.

Since $W_{pq} \hookrightarrow L_p(I,H)$ is compact, there exists a subsequence, relabelled $\{y_n\}$, such that

$$y_n \to y \quad \text{in } L_p(I,H) \quad \text{as } n \to \infty,$$

and therefore $\{y_n\}$ is Cauchy sequence in $L_p(I,H)$. Hence there exists a constant $M_5 > 0$ such that

$$\begin{aligned}\frac{1}{2}\|y_n(t) - y_m(t)\|_H^2 &\leqslant \|G(x_n) - G(x_m)\|_{Lq(I,H)}\|y_n - y_m\|_{L_p(I,H)} \\ &\leqslant M_5\|y_n - y_m\|_{L_p(I,H)}.\end{aligned}$$

This inequality implies that $\{y_n\}$ is a Cauchy sequence in B. Since B is closed, F is compact.

(4) A priori estimate on fixed points.

Suppose $x \in B$ and $x = \sigma Fx$ where $\sigma \in [0,1]$. This implies that x satisfies the following Cauchy problem:

$$\begin{cases} \dot{x}(t) + A(t,x(t)) = g(t, \sigma x(t), \sigma\hat{x}_t) \quad t \in I, \\ x(0) = 0. \end{cases} \tag{4}$$

We will show that there exists a $Q > 0$ such that

$$\|x\|_{C([0,T],H)} \leqslant Q.$$

Using the same arguments and assumption (A) and (G), we have

$$\begin{aligned}\frac{1}{2}\|x(t)\|_H^2 + C_1\|x\|_{L_p(I_t,V)}^p &\leqslant \int_0^t \langle g(\tau, \sigma x(\tau), \sigma\hat{x}_\tau), x(\tau)\rangle \, d\tau + C_2 \\ &\leqslant \left(\int_0^t \|g(\tau, \sigma x(\tau), \sigma\hat{x}_\tau)\|_{V^*}^q \, d\tau\right)^{1/q} \\ &\quad \times \left(\int_0^t \|x(\tau)\|_V^p \, d\tau\right)^{1/p} + C_2,\end{aligned}$$

hence

$$\frac{1}{2}\|x(t)\|_H^2 + \gamma\|x\|_{L_p(I_t,V)}^p \leqslant a_1 + b_1\int_0^t \|x(\tau)\|_H^{(k-1)q} \, d\tau + d_1\int_0^t \|\hat{x}_\tau\|_C^2 \, d\tau,$$

where γ, a_1, b_1, d_1 are positive constants. Further, we have the following inequality:

$$\frac{1}{2}\|x(t)\|_H^2 + \gamma\|x\|_{L_p(I_t,V)}^p \leqslant a_1 + b_2\|x\|_{L_p(I_t,V)}^{(k-1)q} + d_1\int_0^t \|\hat{x}_\tau\|_C^2 \, d\tau$$

where $b_2 > 0$ is a constant.

Consider the real function

$$h(\xi) = \gamma\xi^p - (a_1 + b_2\xi^{(k-1)q})$$

with $(k-1)q < p$. There exists a constant $\xi_0 > 0$ such that

$$h(\xi) \geqslant 0 \quad \text{for all } \xi \geqslant \xi_0.$$

Hence there exists $a_2 \geqslant 0$ such that

$$a_1 + b_1 \xi^{(k-1)q} \leqslant a_2 \quad \text{for all } 0 \leqslant \xi \leqslant \xi_0.$$

It implies that

$$\frac{1}{2}\|x(t)\|_H^2 \leqslant a_2 + d_1 \int_0^t \|\hat{x}_\tau\|_C^2 \, d\tau.$$

We denote

$$k(t) = 2\left(a_2 + d_1 \int_0^t \|\hat{x}_\tau\|_C^2 \, d\tau\right).$$

It is obvious that $k(t)$ is continuous increasing function in R. Hence

$$\begin{aligned}
\|\hat{x}_t\|_C^2 &= \sup_{-r\leqslant s\leqslant 0} \|\hat{x}(t+s)\|_H^2 \\
&\leqslant \sup_{-r\leqslant \tau\leqslant 0} \|\phi(\tau)\|_H^2 + \sup_{0\leqslant \tau\leqslant t} \|x(\tau)\|_H^2 \\
&\leqslant \sup_{-r\leqslant \tau\leqslant 0} \|\phi(\tau)\|_H^2 + \sup_{0\leqslant \tau\leqslant t} k(\tau) \\
&\leqslant \sup_{-r\leqslant \tau\leqslant 0} \|\phi(\tau)\|_H^2 + k(t) \\
&\leqslant a_3 + d_2 \int_0^t \|\hat{x}_\tau\|_C^2 \, d\tau
\end{aligned}$$

where a_3 and d_2 are positive constants. That is,

$$\|\hat{x}_t\|_C^2 \leqslant a_3 + d_2 \int_0^t \|\hat{x}_\tau\|_C^2 \, d\tau \quad \text{for all } t \in [0, T].$$

Gronwall's Lemma (see [10]) implies that

$$\|x\|_{C([0,T],H)} \leqslant Q.$$

By the Leray–Schauder's fixed point theorem, F has a fixed point x^* in B. x^* is just a solution of (1). □

Remark 2. We can assume g maps bounded set to bounded set instead of assumption (G)(2).

Remark 3. It follows from the proof of Theorem 2.A, we also obtain that if x is a solution of (1) then

$$\|x\|_{W_{pq}} \leqslant Q.$$

Remark 4 (Uniqueness). The Leray–Schauder's Theorem guarantees the existence but not uniqueness. In order to obtain uniqueness, we have to impose some more strong assumption on g.

(G)(3) g is locally Lipschitz continuous with respect to ξ and η, i.e., for any $\rho > 0$, there exists a constant $L(\rho)$ such that

$$\|g(t,\xi_1,\eta_1) - g(t,\xi_2,\eta_2)\|_H \leqslant L(\rho)(\|\xi_1 - \xi_2\|_H + \|\eta_1 - \eta_2\|_C), \quad \forall t \in I$$

provided for $\xi_1, \xi_2 \in H$, $\eta_1, \eta_2 \in C$, and $\|\xi_1\|_H, \|\xi_2\|_H, \|\eta_1\|_C, \|\eta_2\|_C \leqslant \rho$.

In fact, if problem (1) has two solutions x_1, x_2, then

$$\frac{1}{2}\|x_1(t) - x_2(t)\|_H^2 \leqslant \|G(x_1) - G(x_2)\|_{L_2(I_t,H)}\|x_1 - x_2\|_{L_2(I_t,H)}$$

$$\leqslant \frac{1}{2}\|G(x_1) - G(x_2)\|_{L_2(I_t,H)}^2 + \frac{1}{2}\|x_1 - x_2\|_{L_2(I_t,H)}^2.$$

By assumption (G)(3), there exist constants $C_1^* > 0$ and $C_2^* > 0$ such that

$$\|x_1(t) - x_2(t)\|_H^2 \leqslant C_1^* \int_0^t \|x_1(\tau) - x_2(\tau)\|_H^2 \, \mathrm{d}\tau + C_2^* \int_0^t \|(x_1)_\tau - (x_2)_\tau\|_C^2 \, \mathrm{d}\tau.$$

Thanks to Gronwall's Lemma (see [12]), it implies

$$x_1(t) = x_2(t) \quad \text{for all } t \in [0,T].$$

That is,

$$x_1 = x_2.$$

Now let us consider the corresponding control system. For any topological space $\mathscr{X}$, $2^{\mathscr{X}} \setminus \emptyset$ will denote the space of nonempty subset of $\mathscr{X}$ and $P_c(\mathscr{X})$ denotes the class of nonempty closed convex subset of $\mathscr{X}$. Let (Ω, Σ) be an arbitrary measurable space and $\mathscr{X}$ be a metric space. A multifunction $F : \Omega \to P_c(\mathscr{X})$ is said to be measurable if for all $z^* \in \mathscr{X}$, $\omega \to d(z^*, \omega) = \inf\{d(z^*, z) : z \in F(\omega)\}$ is measurable (see Theorem 2.11 of Chapter 3 of [6]). We will use S_F to denote the set of measurable selectors of F.

Assume:

(U) Z is a Polish space.
$U : I \to P_c(Z)$ is a measurable multifunction satisfying $U(t) \subseteq M$, a.e., $t \in [0,T]$, where M is a fixed weakly compact convex subset of Z. For the admissible controls, we choose the set $U_{\text{ad}} = S_U$.

(G1) $g : [0,T] \times H \times C \times Z \to H$
(1) $t \mapsto g(t,\xi,\eta,\zeta)$ is measurable,
$(\xi,\eta,\zeta) \mapsto g(t,\xi,\eta,\zeta)$ is continuous on $H \times C \times Z$.
(2) There exist constants a, b, $\mathrm{d} \geqslant 0$ such that

$$\|g(t,\xi,\eta,\zeta)\|_H \leqslant a + b\|\xi\|_H^{k-1} + \mathrm{d}\|\eta\|_C^{2/q}$$

for all $\xi \in H$, $\eta \in C$, $t \in [0,T]$, and $\zeta \in Z$, where $2 \leqslant k < p$.

By famous selection theorem and assumption (U), $S_U \neq \emptyset$ (see Theorem 2.23, Chapter 3 of [6]).

We consider the following control systems

$$\dot{x}(t) + A(t,x(t)) = g(t,x(t),x_t,u(t))$$

$$x(t) = \phi(t), \quad -r \leqslant t \leqslant 0, \quad u(\cdot) \in U_{ad}. \tag{5}$$

By Theorem 2.A, we immediately obtain the following existence theorem.

Theorem 2.B. *Suppose the assumptions* (A), (U), *and* (G1) *hold. For every* $u \in U_{ad}$ *Eq.* (5) *has a solution* $x(u) \in W_{pq}$.

Remark 5. Imposing the local Lipschitz condition, one can prove the uniqueness of solutions for (5).

Define

$$X_0 = \{x \in W_{pq}^0 \mid x \text{ is a solution of (5) corresponding to } u, u \in U_{ad}\}.$$

X_0 is called the set of original trajectories. Set

$$A_0 = \{(x(u), u) \in W_{pq} \times S_U \mid x(u) \text{ is a solution of (5) corresponding to } u\}.$$

A_0 is called the set of admissible state control pairs.

It can be seen from the proof of Theorem 2.A, the following conclusions are true.

Theorem 2.C. *Under assumptions of Theorem* 2.B, X_0 *is weakly precompact in* W_{pq} *and precompact in* $L_p([-r,T],H)$ *and* $C([-r,T],H)$.

Proof. Using assumptions (A), (G), and the same arguments as in the step (4) of the proof of Theorem 2.A, one can verify that there exists a constant Q such that

$$\|x\|_{W_{pq}} \leqslant Q \quad \text{for all } x \in X_0.$$

This implies that X_0 is weakly precompact in W_{pq}. It is easy to assert from compactness of embedding $V \hookrightarrow H$ that X_0 is precompact in $L_p(I,H)$. Hence every sequence $\{x_n\}$ of X_0 has a subsequence $\{x_{n_k}\}$ which is Cauchy sequence in $L_p(I,H)$. By the arguments similar to the step (3) of the proof of Theorem 2.A, one can show that $\{x_{n_k}\}$ is also a Cauchy sequence in $C(I,H)$. This means X_0 is precompact in $C(I,H)$. □

3. Relaxed systems

We consider the following optimal control problem:

$$(P) \quad \inf\left\{J(x,u) = \int_0^T L(t,x(t),x_t,u(t))\,dt\right\}$$

subject to Eq. (5).

It is well known that the convexity conditions or more precisely the Cesari property play a central role in the study of existence of optimal controls. If the convexity hypothesis is no longer satisfied, to have optimal control solution, we need to pass to a larger systems, in which the orientor field have been convexified. For the purpose here we introduce the relaxed controls and the corresponding relaxed systems.

Let Z be a Polish space (i.e., a separable complete metric space) and $B(Z)$ be its Borel σ-field. We will denote the space of probability measures on Z by $M_+^1(Z)$. Let (Ω, Σ) be a measurable space. Transition probability is a function $\lambda : \Omega \times B(Z) \to [0,1]$ such that for every $\xi \in B(Z)$, $\lambda(\cdot,\xi)$ is $\sum$-measurable and for every $\omega \in \Omega$, $\lambda(\omega,\cdot) \in M_+^1(Z)$. We use $R(\Omega, Z)$ to denote the set of all transition probabilities from $(\Omega, \sum, \mu)$ into $(Z, B(Z))$. Following Balder [3] (see also Warga [8]), we can define a topology on $R(\Omega, Z)$ as follows. Let $f : \Omega \times Z \to R$ be a Caratheodory function (i.e. $\omega \to f(\omega,x)$ is measurable, $x \to f(\omega,x)$ is continuous and $|f(\omega,x)| \leqslant \alpha(\omega)$ is μ-almost everywhere with $\alpha(\cdot) \in L_1(\Omega)$) and $I_f(\lambda) = \int_\Omega \int_Z f(\omega,z)\lambda(\omega)(\mathrm{d}z)\,\mathrm{d}\mu(\omega)$. The weakest topology on $R(\Omega, Z)$ that makes the above functional continuous is called the weak topology on $R(\Omega, Z)$.

Suppose $\Omega = I = [0,T]$ and Z is a compact Polish space, then Caratheodory integrands on $I \times Z$ can be identified with the Lebesgue Bochner space $L_1(C(Z))$. We know $[C(Z)]^* = M(Z)$ is the space of all bounded Borel measures on $B(Z)$. $M(Z)$ is a separable Banach space and hence has the Radon–Nikodym property which tell us that

$$(L_1(C(Z)))^* = L_\infty(M(Z)).$$

So the weak topology on $R(I,Z)$ coincides with the relative $w^*(L_\infty(M(Z)), L_1(C(Z)))$ topology (see [8,4]).

Now we introduce some assumptions imposed on U_{ad}.

(U1) Z is a compact Polish space. $U : I \to P_c(Z)$ is a measurable multifunction.

Define $\sum(t) = \{\lambda \in M_+^1(Z),\ \lambda(U(t)) = 1\}$ and S_Σ is the set of transition probabilities that are measurable selectors of $\Sigma(\cdot)$. Since $\delta(U(t)) \subseteq \Sigma(t)$, then $u \in S_U$ implies $\delta(u) \in S_\Sigma$ i.e., $S_U \subseteq S_\Sigma$.

The following lemma are crucial in discussing relaxation problem (for proofs one can see [8,3]).

Lemma 3.1. *Suppose Z is a compact Polish space. Then S_Σ is convex and sequentially compact.*

Lemma 3.2. *S_U is dense in S_Σ.*

Lemma 3.3. *Suppose $h : I \times H \times C \times Z \to R$ satisfying*

1. *$t \mapsto h(t,\xi,\eta,\sigma)$ is measurable, $(\xi,\eta,\sigma) \mapsto h(t,\xi,\eta,\sigma)$ is continuous.*
2. *$|h(t,\xi,\eta,\sigma)| \leqslant \psi_R(t) \in L_1(I)$ provided with $\|\xi\|_H \leqslant R$, $\|\eta\|_C \leqslant R$, and $\sigma \in Z$.*

If

$$x_n \to x \quad \text{in } C([-r,T],H)$$

then

$$h_n(\cdot,\cdot) \to h(\cdot,\cdot) \quad \text{in } L_1(C(Z))$$

as $n \to \infty$, *where*

$$h_n(t,\sigma) = h(t,x_n(t),(x_n)_t,\sigma),$$

$$h(t,\sigma) = h(t,x(t),x_t,\sigma).$$

Proof. It is obvious that

$$h_n, h \in L_1(C(Z)).$$

The convergence

$$x_n \to x \quad \text{in } C([-r,T],H)$$

implies

$$(x_n)_t \to x_t \quad \text{in } C$$

uniformly with respect to $t \in I$, as $n \to \infty$.

For fixed $t \in I$, we show that $h_n(t,\cdot) \to h(t,\cdot)$ in $C(Z)$ as $n \to \infty$.

In fact, for fixed $t \in I$, $n \in N$, there exists $\sigma_n \in Z$ such that

$$\sup_{\sigma \in Z} |h_n(t,\sigma) - h(t,\sigma)| = |h_n(t,\sigma_n) - h(t,\sigma_n)|.$$

Since Z is compact, we assume $\sigma_n \to \sigma^*$. By continuity of h, one can verify that

$$\begin{aligned}
&\sup_{\sigma \in Z} |h_n(t,\sigma) - h(t,\sigma)| \\
&\quad = |h_n(t,\sigma_n) - h(t,\sigma_n)| \\
&\quad \leqslant |h_n(t,\sigma_n) - h_n(t,\sigma^*)| + |h_n(t,\sigma^*) - h(t,\sigma^*)| + |h(t,\sigma_n) - h(t,\sigma^*)| \\
&\quad \to 0 \quad \text{as } n \to \infty,
\end{aligned}$$

i.e.,

$$\|h_n(t,\cdot) - h(t,\cdot)\|_{C(Z)} \to 0 \quad \text{as } n \to \infty.$$

By assumption (2) and Majorized convergent principle, we have

$$h_n(\cdot,\cdot) \to h(\cdot,\cdot) \quad \text{in } L_1(C(Z)) \quad \text{as } n \to \infty. \qquad \square$$

We consider the relaxed system

$$\begin{aligned}
&\dot{x}(t) + A(t,x(t)) = \int_Z g(t,x(t),x_t,\sigma)\,\mathrm{d}\mu_t(\sigma) \quad 0 \leqslant t \leqslant T, \\
&x(t) = \phi(t) \quad -r \leqslant t \leqslant 0, \ \mu \in S_\Sigma.
\end{aligned} \tag{6}$$

Similar to the arguments of Theorem 2.B, one can verify the following existence theorem.

Theorem 3.A. *Suppose the assumptions* (A), (G1), (U1) *hold. For each* $\mu \in S_\Sigma$, *the relaxed system* (6) *has a solution.*

Define

$$X_r = \{x \in W^0_{pq} \,|\, x \text{ is a solution of (6) corresponding to } \mu, \mu \in S_\Sigma\}.$$

$$A_r = \{(x,\mu) \in W_{pq} \times S_\Sigma \,|\, x \text{ is a solution of (6) corresponding to } \mu, \mu \in S_\Sigma\}.$$

X_r is called set of relaxed trajectories and A_r is called set of admissible relaxed pairs. Theorem 3.A shows that

$$X_r \neq \emptyset$$

and combining with $S_U \subset S_\Sigma$ we have

$$X_0 \subseteq X_r.$$

Similar to Theorem 2.C, we have the following conclusion.

Theorem 3.B. *Under assumptions of Theorem* 3.A, X_r *is weakly precompact in* W_{pq} *and precompact in* $L_p([-r,T],H)$ *and* $C([-r,T],H)$.

This then raises the fundamental question of how much we enlarged the set of trajectories of original system. The next theorem answers this question by stating that this process does not essentially alter the original solution set.

Theorem 3.C. *If assumption* (A), (G1), (U1) *hold, then* $X_r = \overline{X_0}$ (*the closure is taken in* $C(I,H)$) *provided that the solution of* (6) *is unique.*

Proof. Suppose $\tilde{y} \in X_r$, i.e., $(\tilde{y},\lambda) \in A_r$ for some $\lambda \in S_\Sigma$.

By virtue of density result Lemma 3.2, there exists a sequence $\{u_n\} \subset S_U$ such that

$$\delta_{u_n} \to \lambda \quad \text{in } R(I,Z).$$

We have sequence $\{(y_n,u_n)\} \subset A_0$. Since X_0 is bounded in W_{pq} and $C([-r,T],H)$, there exists a constant $M > 0$ such that

$$\|y_n\|_{C([-r,T],H)} \leqslant M,$$

$$\|y_n\|_{W_{pq}(I)} \leqslant M.$$

It follows from Theorem 2.C that there exist $y \in W_{pq}$ and $w \in L_q(I,V^*)$ such that

$$y_n \xrightarrow{w} y \text{ in } L_p(I,V), \quad \dot{y}_n \xrightarrow{w} \dot{y} \text{ in } L_q(I,V^*),$$

$$A(\cdot,y(\cdot)) \xrightarrow{w} w \quad \text{in } L_q(I,V^*),$$

$$y_n \to y \text{ in } L_p(I,H), \quad y_n \to y \text{ in } C([-r,T],H),$$

$$(y_n)_t \to y_t \quad \text{in } C \text{ uniformly for all } t \in I,$$

as $n \to \infty$ (if necessary passing to subsequence). Consider the following equations:

$$\dot{y}_n(t) + A(t,y_n(t)) = \int_Z g(t,y_n(t),(y_n)_t,\sigma)\delta_{u_n}(t)(\mathrm{d}\sigma), \quad 0 < t \leqslant T,$$

$$y_n(t) = \phi(t), \quad -r \leqslant t \leqslant 0. \tag{7}$$

Set

$$G(y_n(t),\sigma) = g(t, y_n(t), (y_n)_t, \sigma), \quad G(y(t),\sigma) = g(t, y(t), y_t, \sigma),$$

$$G_n^*(t) = \int_Z G(y_n(t),\sigma)\delta_{u_n}(t)(\mathrm{d}\sigma), \quad G^*(t) = \int_Z G(y(t),\sigma)\lambda(t)(\mathrm{d}\sigma).$$

For each $\psi \in C^\infty(I)$ and $v \in V$, define

$$\bar{G}_n(t,\sigma) = ((Gy_n(t),\sigma), \psi(t)v), \quad \bar{G}(t,\sigma) = ((Gy(t),\sigma), \psi(t)v).$$

For fixed $t \in I$, it follows from assumption (G1) that

$$\bar{G}_n(t,\cdot), \bar{G}(t,\cdot) \in C(Z),$$

furthermore

$$\bar{G}_n(\cdot,\cdot), \bar{G}(\cdot,\cdot) \in L_1(C(Z)).$$

Since Z is compact Polish space, and

$$y_n \to y \quad \text{in } C(I,H) \quad \text{as } n \to \infty,$$

by assumption (G1) and Lemma 3.3, we have

$$\bar{G}_n \to \bar{G} \quad \text{in } L_1(C(Z)) \quad \text{as } n \to \infty.$$

In the topology of $R(I,Z)$, we assert

$$\int_I \int_Z \bar{G}_n(t,\sigma)\delta_{u_n}(t)(\mathrm{d}\sigma)\,\mathrm{d}t \to \int_I \int_Z \bar{G}(t,\sigma)\lambda(t)(\mathrm{d}\sigma)\,\mathrm{d}t \quad \text{as } n \to \infty.$$

That is,

$$\int_I \langle G_n^*(t), \psi(t)v\rangle\,\mathrm{d}t \to \int_I \langle G^*(t), \psi(t)v\rangle\,\mathrm{d}t.$$

i.e.,

$$\langle\langle G_n^*, \psi v\rangle\rangle \to \langle\langle G^*, \psi v\rangle\rangle.$$

It implies that

$$G_n^* \xrightarrow{W} G^* \quad \text{in } L_q(I, V^*).$$

We conclude that y satisfies the following equation

$$\begin{cases} \dot{y}(t) + w(t) = \int_Z g(t, y(t), y_t, \sigma)\lambda(t)(\mathrm{d}\sigma) & 0 \leqslant t \leqslant T, \\ y(t) = \phi(t) & -r \leqslant t \leqslant 0. \end{cases}$$

Observe that

$$\langle\langle G^*(y_n), y_n\rangle\rangle = \int_I \int_Z (G(y_n(t),\sigma), y_n(t))\delta_{u_n}(t)(\mathrm{d}\sigma)\,\mathrm{d}t$$

and

$$(G(y_n(t),\sigma), y_n(t)) = (G(y_n(t),\sigma) - G(y(t),\sigma), y_n(t)) + (G(y(t),\sigma), y_n(t) - y(t)) + (G(y(t),\sigma), y(t)).$$

Similar to the above procedures, one can verify that

$$(G(y_n(\cdot),\cdot) - G(y(\cdot),\cdot), y_n(\cdot)) \to 0 \quad \text{in } L_1(C(Z)).$$

By growth condition, we have,

$$|(G(y(t),\sigma), y_n(t) - y(t))| \leqslant (a + b\|y(t)\|_H^{k-1} + d\|y_t\|_C^{2/q})\|y_n(t) - y(t)\|_H.$$

Hence,

$$(G(y(\cdot),\cdot), y_n(\cdot) - y(\cdot)) \to 0 \quad \text{in } L_1(C(Z)).$$

We conclude that

$$\langle\langle G^*(y_n), y_n\rangle\rangle \to \langle\langle G^*(y), y\rangle\rangle.$$

By integration by parts, it follows from (7) that

$$\overline{\lim_{n\to\infty}} \langle\langle A(y_n), y_n\rangle\rangle \leqslant \tfrac{1}{2}(\|y(0)\|_H^2 - \|y(T)\|_H^2) + \langle\langle G^*(y), y\rangle\rangle$$
$$= \langle\langle w, y\rangle\rangle.$$

Since A satisfies condition (M) (see p. 474 of [13]), we have

$$A(y) = w.$$

Now we can say that y is the solution of following equation:

$$\begin{cases} \dot{y} + A(y) = \int_Z g(t, y(t), y_t, \sigma)\lambda(t)(\mathrm{d}\sigma) & 0 \leqslant t \leqslant T, \\ y(t) = \phi(t) & -r \leqslant t \leqslant 0. \end{cases}$$

Uniqueness implies

$$y = \tilde{y}.$$

This means that

$$y_n \to \tilde{y} \quad \text{in } C(I,H).$$

As $\{y_n\} \subset X_0$ and $\tilde{y} \in X_r$, we proved that

$$\overline{X_0} \supseteq X_r \quad \text{in } C(I,H).$$

Since $R(I,Z)$ is sequentially compact, by the same procedure one can show that

$$\overline{X_0} \subseteq X_r = \overline{X_r}.$$

Hence

$$X_r = \overline{X_0}$$

where closure is taken in $C(I,H)$. □

The following corollaries can be obtained from the proof of Theorem 3.B.

Corollary 3.C. *Under assumptions of Theorem* 3.B, X_r *is sequentially compact in* $C(I,H)$.

Corollary 3.D. *Under assumptions of Theorem* 3.B, $\lambda \to x(\lambda) \in X_r$ *is continuous from* $S_\Sigma \subset R(I,Z)$ *into* $C(I,H)$.

Proof. Let $\lambda_n \to \lambda$ in $R(I,Z)$. For corresponding sequence $\{x_n\}$ of solutions, we have a subsequence $\{x_{n_k}\} \subset \{x_n\}$ such that

$$x_{n_k} \to x \quad \text{in } C(I,H),$$

where x is the unique solution of (6) corresponding to λ. By convergence principle, we assert that

$$x_n \to x \quad \text{in } C(I,H).$$

The proof is completed. □

4. Relaxation theorem

As we have already mentioned the introduction of the larger relaxed system guarantees the existence of an optimal solution. This is illustrated by the following general result.

Consider the following problem (P_r):

$$J(y,\lambda) = \int_I \int_Z \ell(t, y(t), y_t, \sigma)\lambda(t)(\mathrm{d}\sigma)\,\mathrm{d}t = \min$$

subject to system (6).

We make the following hypotheses concerning the integrand $\ell(\cdot,\cdot,\cdot,\cdot)$.

(L) $\ell : I \times H \times C \times Z \to \bar{R} = R \cup \{+\infty\}$
  (1) $(t,\xi,\eta,\sigma) \to \ell(t,\xi,\eta,\sigma)$ is measurable,
  (2) $(\xi,\eta,\sigma) \to \ell(t,\xi,\eta,\sigma)$ is lower semicontinuous,
  (3) $\psi(t) \leqslant \ell(t,\xi,\eta,\sigma)$ almost everywhere with $\psi(\cdot) \in L^1$.

Let $m_r = \inf\{J(y,\lambda),\ (y,\lambda) \in A_r\}$. We have the following existence of relaxed optimal control:

Theorem 4.A. *Suppose assumptions* (A), (G1), (U), (L) *hold and Z is compact Polish space, then there exists* $(x,\lambda) \in A_r$ *such that* $J(x,\lambda) = m_r$.

Proof. Let $\{(x_n, \lambda_n)\}$ be a minimizing sequence in A_r. Recall that S_Σ is w^*-compact, by passing to a subsequence if necessary, we may assume $\lambda_n \to \lambda$ in $R(I,Z)$.

Invoking Theorem 3.B, we may assume

$$x_n \to x \quad \text{in } C(I,H),$$

$(x,\lambda) \in A_r$.

Recalling that every lower semicontinuous measurable integrand is the limit of an increasing sequence of Caratheodory integrands, there exists increasing sequence of Caratheodory integrands $\{l_k\}$ such that

$$l_k(t,\xi,\eta,\sigma) \uparrow l(t,\xi,\eta,\sigma) \quad \forall (t,\xi,\eta,\sigma) \in I \times H \times C \times Z.$$

Invoking the definition of the weak topology and Lemma 3.3, we have

$$
\begin{aligned}
J(x,\lambda) &= \int_I \int_Z l(t,x(t),x_t,\sigma)\lambda(t)(\mathrm{d}\sigma)\,\mathrm{d}t \\
&= \lim_{k\to\infty} \int_I \int_Z l_k(t,x(t),x_t,\sigma)\lambda(t)(\mathrm{d}\sigma)\,\mathrm{d}t \\
&= \lim_{k\to\infty}\lim_{n\to\infty} \int_I \int_Z l_k(t,x_n(t),(x_n)_t,\sigma)\lambda_n(t)(\mathrm{d}\sigma)\,\mathrm{d}t \\
&\leqslant \lim_{n\to\infty} \int_I \int_Z l(t,x_n(t),(x_n)_t,\sigma)\lambda_n(t)(\mathrm{d}\sigma)\,\mathrm{d}t \\
&= m_r.
\end{aligned}
$$

However, by definition of m_r, it is obvious that $J(x,\lambda) \geqslant m_r$. Hence

$$J(x,\lambda) = m_r.$$

This implies that (x,λ) is an optimal pair. □

If $J_0(x,u) = \int_I \ell(t,x(t),x_t,u(t))\,\mathrm{d}t$ is the cost functional for the original problem and $m = \inf\{J_0(x,u),\ (x,u)\in A_0\}$. In general we have $m_r \leqslant m$. It is desirable that $m_r = m$, i.e., our relaxation is reasonable. We have the following relaxation theorem. For this, we need stronger hypotheses on ℓ than (L):

(L1) $\ell : I \times H \times C \times Z \to R$ is an integrality such that
- (1) $(t) \to \ell(t,\xi,\eta,\sigma)$ is measurable,
- (2) $(\xi,\eta,\sigma) \to \ell(t,\xi,\eta,\sigma)$ is continuous,
- (3) $|\ell(t,\xi,\eta,\sigma)| \leqslant \theta_R(t)$ for all almost $t\in I$ provided $\|\xi\|_H, \|\eta\|_C \leqslant R$, $\sigma\in Z$, and $\theta_R \in L^1$.

Theorem 4.B. *If assumptions* (A), (G1), (U), (L1) *hold and Z is compact, then* $m = m_r$ *provided that the solution of* (6) *is unique.*

Proof. Theorem 4.A shows that there exists $(x,\lambda)\in A_r$ such that

$$J(x,\lambda) = m_r.$$

By Lemma 3.2, there exists a sequence $\{u_n\} \subset S_U$ such that $\delta_{u_n} \to \lambda$ in $R(I,Z)$.

Let x_n be the solution of (6) corresponding to u_n. Passing to a subsequence if necessary, we may assume that

$$x_n \to x \quad \text{in } C([-r,T],H)$$

(see Theorem 2.C).

Similarly, by using Lemma 3.3, one can verify that

$$\ell(\cdot,x_n(\cdot),(x_n)_\cdot,\cdot) \to \ell(\cdot,x(\cdot),x_\cdot,\cdot) \quad \text{in } L_1(C(Z)).$$

By definition of weak topology on $R(I,Z)$, we have

$$J_0(x_n,u_n) = J(x_n,\delta_{u_n}) = \int_I \int_Z \ell(t,x_n(t),(x_n)_t,\sigma)\delta_{u_n}(t)(\mathrm{d}\sigma)\,\mathrm{d}t$$

$$\to \int_I \int_Z \ell(t,x(t),x_t,\sigma)\lambda(t)(\mathrm{d}\sigma)\,\mathrm{d}t = J(x,\lambda) = m_r.$$

This implies

$$m = m_r. \quad \square$$

5. Examples

In this section we present two examples of delay evolution equations to which our general theory applies.

Let Ω be an bounded domain in R^n with smooth boundary $\partial\Omega$, $Q_T = (0,T) \times \Omega$, $0 < T < \infty$. Let $\alpha = (\alpha_1,\alpha_2,\dots,\alpha_n)$ be a multi-index with $\{\alpha_i\}$ nonnegative integers and $|\alpha| = \sum_{i=1}^n \alpha_i$. Suppose $p \geqslant 2$ and $q = p/(p-1)$, $W^{m,p}(\Omega)$ denotes the standard Sobolev space with the usual norm:

$$\|\varphi\|_{W^{m,p}} = \left(\sum_{|\alpha| \leqslant m} \|D^\alpha \varphi\|^p_{L^p(\Omega)} \right)^{1/p}, \quad m = 0,1,2,\dots .$$

Let $W_0^{m,p}(\Omega) = \{\varphi \in W^{m,p} \mid D^\beta\varphi|_{\partial\Omega} = 0,\ |\beta| \leqslant m-1\}$. It is well known that $C_0^\infty(\Omega) \hookrightarrow W_0^{m,p}(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow W^{-m,p}(\Omega)$and the embedding $W_0^{m,p}(\Omega) \hookrightarrow L^2(\Omega)$ is compact. Denote $V \equiv W_0^{m,p}(\Omega)$, $H \equiv L_2(\Omega)$, then $V^* \equiv W^{-m,q}(\Omega)$.

Example 1. We consider the following initial-boundary value problem of $2m$-order quasi-linear delay parabolic control system:

$$\frac{\partial}{\partial t} y(t,x) + \sum_{|\alpha| \leqslant m} (-1)^{|\alpha|} D^\alpha A_\alpha(t,x,\eta(y)(t,x)) = g(t,x,y(t,x),y(t-r,x),u) \text{ on } Q_T,$$

$$D^\beta y(t,x) = 0 \quad \text{on } [0,T] \times \partial\Omega \quad \text{for all } \beta\text{: } |\beta| \leqslant m-1,$$

$$y(s,x) = \phi(s,x), \quad \text{on } \Omega,\ -r \leqslant s \leqslant 0. \tag{8}$$

where $\eta(y) \equiv \{(D^\gamma y),\ |\gamma| \leqslant m\}, \phi(t,x)$ is given function, $\phi \in C([-r,0],L_2(\Omega))$ and $M = (n+m)!/n!m!$.

Suppose that $\beta_{1i}(\cdot), \beta_{2i}(\cdot)$ $(1 \leqslant i \leqslant M_1)$ are continuous functions from $[0,T]$ to R and satisfy

$$\beta_{1i}(t) < \beta_{2i}(t) \quad \text{for all } t \in [0,T],\ 1 \leqslant i \leqslant M_1.$$

There exists a constant $a > 0$ such that

$$-a \leqslant \beta_{1i}(t) < \beta_{2i}(t) \leqslant a \quad \text{for all } t \in [0,T], 1 \leqslant i \leqslant M_1.$$

Set $Z = [-a,a]^{M_1} \subset R^{M_1}$. Then Z is a compact Polish space. Set

$$U(t) \equiv \{(w_i(t)) \in R^{M_1},\ \beta_{1i}(t) \leqslant w_i(t) \leqslant \beta_{2i}(t),\ 1 \leqslant i \leqslant M_1\}.$$

It is clearly that $U : I \to P_c(Z)$. The set of admissible controls $U_{\rm ad}$ is chosen as

$$U_{\rm ad} \equiv S_U$$
$$= \{u : I \to R^{M_1} \text{ is measurable } u(t) \in U(t) \text{ a.e. } t \in [0,T]\}.$$

Assume that the function $l_0 : Q_T \times R \times R \times R^{M_1} \to \bar{R} = R \cup \{+\infty\}$ is continuous and

$$|l_0(t,x,\xi,\eta,u)| \leqslant d_1|\xi|^2 + d_2|\eta|^2 + \theta(t,x)$$

with constants d_1, d_2, and the function $\theta \in L^1(Q_T)$ for a.e. $(t,x) \in Q_T$ and $u \in Z$. For $\xi, \eta \in L_2(\Omega)$, $u \in Z$, we define

$$l(t,\xi,\eta,u) = \int_\Omega l_0(t,x,\xi(x),\eta(x),u)\,\mathrm{d}x. \tag{9}$$

The cost functional is given by

$$J(u) = \int_I l(t,y(t),y(t-r),u(t))\,\mathrm{d}t.$$

The optimal control problem (P^*) is to find $u_0 \in U_{\rm ad}$ s.t.

$$J(u_0) \leqslant J(u) \quad \text{for all } u \in U_{\rm ad}$$

subject to system (8).

For $y_1, y_2 \in W_0^{m,p}(\Omega)$, $t \in I$, we set

$$a(t,y_1,y_2) = \int_\Omega \sum_{|\alpha| \leqslant m} A_\alpha(t,x,\eta(y_1)(t,x)) D^\alpha y_2\,\mathrm{d}x$$

and assume that for all α with $|\alpha| \leqslant m$, the function $A_\alpha : Q_T \times R^M \to R$ satisfies the following properties.

($\tilde{A}$) (1) $(t,x) \to A_\alpha(t,x,\eta)$ is measurable on Q_T for $\eta \in R^M$, $\eta \to A_\alpha(t,x,\eta)$ is continuous on R^M for a.e. $(t,x) \in Q_T$;
(2) For $\eta = (\eta_\alpha) \in R^M$, $\tilde{\eta} = (\tilde{\eta}_\alpha) \in R^M$, there exist positive constants c, c_1, c_2, c_3, and c_4 such that

$$\sum_{|\alpha| \leqslant m} (A_\alpha(t,x,\eta) - A_\alpha(t,x,\tilde{\eta}))(\eta_\alpha - \tilde{\eta}_\alpha) \geqslant 0,$$

$$\sum_{|\alpha| \leqslant m} A_\alpha(t,x,\eta)\eta_\alpha \geqslant c_1 \sum_{|\gamma| \leqslant m} |\eta_\gamma|^p - c_2,$$

$$|A_\alpha(t,x,\eta)| \leqslant c_4 + c_3 \sum_{|\gamma| \leqslant m} |\eta_\gamma|^{p-1}.$$

It is not difficult to verify that under the above assumption, for each $y_1 \in V$ and $t \in [0,T]$, $y_2 \to a(t,y_1,y_2)$ is a continuous linear form on V. Hence there exists an operator $A : I \times V \to V^*$ such that

$$\langle A(t,y_1), y_2\rangle_{V^*,V} = a(t,y_1,y_2).$$

Under the given assumption ($\tilde{\mathrm{A}}$), it is easy to see that A satisfies our assumption (A) of Section 2.

Assume the function $g : Q_T \times R \times R \times R^{M_1} \to R$ satisfies the following properties.

($\tilde{\mathrm{G}}$) (1) $(t,x) \to g(t,x,\xi,\eta,u)$ is measurable on Q_T for all $(\xi,\eta,u) \in R \times R \times R^{M_1}$;
(2) $(\xi,\eta,u) \to g(t,x,\xi,\eta,u)$ is continuous on $R \times R \times R^{M_1}$ for almost all $(t,x) \in Q_T$;
(3) There exist constants $b_1 > 0$, $b_2 > 0$ and $b_3 \in L_2(Q_T)$ s.t.

$$|g(t,x,\xi,\eta,u)| \leqslant b_1|\xi|^{2/q} + b_2|\eta|^{2/q} + b_3(t,x),$$

for almost all $(t,x) \in Q_T$ uniformly in $u \in U(t)$.

For $\phi_1, \phi_2 \in H$, $\upsilon \in Z$, $t \in I$, set

$$b^{\upsilon}(t,\phi_1,\phi_2,\psi) = \int_{\Omega} g(t,x,\phi_1,\phi_2,\upsilon)\psi \, dx.$$

Then $\psi \to b^{\upsilon}(t,\phi_1,\phi_2)$ is a continuous linear form on H. Hence there exists an operator $G : [0,T] \times H \times H \times Z \to H$ s.t.

$$b^{\upsilon}(t,\phi_1,\phi_2,\psi) = (G(t,\phi_1,\phi_2,\upsilon),\psi).$$

Noting that $y_t(\theta) = y_t(r)$ for all $-r \leqslant \theta \leqslant 0$ and ($\tilde{\mathrm{G}}$), one can verify that G satisfies assumption (G1) of Section 2.

Using the operators A and G as defined above, Eq. (8) can be written as the abstract evolution equation

$$\dot{y}(t) + A(t,y(t)) = G(t,y(t),y(t-r),u), \quad 0 < t < T,$$

$$x(t) = \phi(t), \quad t \in (-r,0).$$

In addition, for the cost functional l which is defined in (9), we get that $l : [0,T] \times H \times H \times Z \to R$ satisfies assumption (L1) of Section 4.

Hence our result can be used to this model.

Example 2. Consider the system of reaction diffusion equations with delay:

$$\frac{\partial \psi}{\partial t} = D(t,x)\Delta\psi + f(t,x,\psi(t,x),\psi(t-r,x),u) \quad \text{on } Q_T,$$

$$\psi(t,x) = 0 \quad \text{on } [0,T] \times \partial\Omega,$$

$$\psi(t,x) = \varphi(t,x) \quad \text{on } \Omega,\ -r \leqslant t < 0, \tag{10}$$

where ψ is an k-vector-valued function on Q_T, $\varphi(t,x)$ is a given k-vector-valued function, $\varphi \in C([-r,0], L_2(\Omega,R^k))$. $D = \mathrm{diag}(d_1,d_2,\ldots,d_k)$ is the diffusion matrix and $d_i > 0$, $i = 1,2,\ldots,k$.

For the mathematical setting, we take $p = q = 2$, $H = L_2(\Omega, R^k)$, $V = W_0^{1,2}(\Omega, R^k)$ with $V^* = W^{-1,2}(\Omega, R^k)$ being the dual of V.

For the nonlinear term, we assume that $f(t,x,\xi,\eta,u)$ is a k-vector-valued function defined on $Q_T \times R^k \times R^k \times Z \to R^k$ which satisfies the following properties:

(1) f is continuous in all the variables,
(2) There exist a function $b_1 \in L_2(\Omega)$ and constants $b_2 \geqslant 0$, $b_3 \geqslant 0$ such that

$$|f(t,x,\xi,\eta,u)|_{R^k} \leqslant b_1(t,x) + b_2|\xi|_{R^k} + b_3|\eta|_{R^k} \quad \text{a.e. in } Q_T \text{ uniformly in } u \in Z.$$

For $\phi_1, \phi_2 \in L^2(\Omega), u \in Z$, $t \in I$, set

$$f^u(t,\phi_1,\phi_2,\psi) = \int_\Omega f(t,x,\phi_1,\phi_2,u)\psi \, dx.$$

Then $\psi \to f^v(t,\phi_1,\phi_2)$ is a continuous linear form on H. Hence there exists an operator $F : [0,T] \times H \times H \times Z \to H$ s.t.

$$f^u(t,\phi_1,\phi_2,\psi) = (F(t,\phi_1,\phi_2,u),\psi)$$

and F satisfies assumption (G1) of Section 2.

Let $A = -D\Delta$, it is obvious that the operator $A \in L(V,V^*)$ and it is coercive and hence monotone. A satisfies the assumption (A) of Section 2.

The same as example 1, we choose that $Z = [-a,a]^{M_1} \subset R^{M_1}$. Then Z is a compact Polish space.

Set

$$U(t) \equiv \{v \in R^{M_1}: \beta_{1i}(t) \leqslant v_i(t) \leqslant \beta_{2i}(t),\ i = 1,2,\ldots,M_1\}$$

The admissible controls are given by

$$U_{ad} \equiv \{u : I \to R^{M_1} \text{ is measurable, } u(t) \in U(t) \text{ a.e.}\}.$$

For the cost integrand l, one may choose the quadratic function $l : I \times H \times H \times Z \to R$ with

$$\begin{aligned} l(t,\psi(t),\psi(t-r),u) = &\int_\Omega (C_1(t,x)(\psi(t,x) - y_1(t,x)),(\psi(t,x) - y_1(t,x)))_{R^k}\, dx \\ &+ \int_\Omega (C_2(t,x)(\psi(t-r,x) - y_2(t,x)), \\ &(\psi(t-r,x) - y_2(t,x)))_{R^k}\, dz \\ &+ \int_\Omega (C_3(t,x)u(t,x),u(t,x))_{R^{M_1}}\, dx \end{aligned}$$

with C_i, $i = 1,2,3$ are positive semidefinite matrix-valued functions on $Q_T, y_i(i = 1,2) \in R^k$ are target states. Then l satisfies assumption (L1) of Section 4. Our problem takes the following abstract form:

$$\inf_{U_{\rm ad}}\{J(u)=\int_0^T l(t,\psi(t),\psi(t-r),u(t))\,\mathrm{d}t\}=m,$$

$$\begin{cases}\text{s.t. } \dot{\psi}(t)+A\psi(t)=F(t,\psi(t),\psi(t-r),u(t)),\\ \psi(s)=\varphi(s), \quad -r\leqslant s\leqslant 0,\end{cases}$$

$$u\in U_{\rm ad}.$$

Since all the assumptions of our abstract result are satisfied in this particular case, our result can be applied in this model.

References

[1] N.U. Ahmed, Properties of relaxed trajectories for a class of nonlinear evolution equations on a Banach Space, SIAM J. Control Optim. 21 (1983) 953–967.
[2] N.U. Ahmed, K.L. Teo, Optimal Control of Distributed Parameter Systems, North-Holland, New York, Oxford, 1981.
[3] E. Balder, A general denseness result for relaxed control theory, Bull. Austral. Math. Soc. 20 (1984) 463–475.
[4] H.O. Fattorini, Infinite Dimensional Optimization and Control Theory, Cambridge University Press, New York, USA, 1999.
[5] A. Filippov, On certain questions in the theory of optimal control, SIAM J. Control Ser. A. 1 (1962) 76–84.
[6] X. Li, J. Yong, Optimal Control Theory for Infinite Dimensional System, Birkhäuser, Boston, 1995.
[7] J. Warga, Relaxed variational problems, J. Math. Anal. Appl. 4 (1962) 111–128.
[8] J. Warga, Optimal Control of Differential and Functional Equations, Academic Press, New York, 1972.
[9] J. Wu, Theory and Applications of Partial Functional Differential Equations, Springer, New York, 1996.
[10] X. Xiang, Existence of periodic solutions for delay partial differential equations, Acta Math. Sci. 14 (1991) 73–81.
[11] X. Xiang, N.U. Ahmed, Properties of relaxed trajectories of evolution equations and optimal control, SIAM J. Control Optim. 31 (1993) 1135–1142.
[12] X. Xiang, H. Kuang, Delay systems and optimal control, Acta Math. Appl. Sinica 16 (2000) 27–35.
[13] E. Zeidler, Nonlinear Functional Analysis and its Applications II, Springer, New York, 1990.