

OPTIMAL CONTROL FOR QUASILINEAR RETARDED PARABOLIC SYSTEMS

LIPING PAN¹ and JIONGMIN YONG¹

(Received 19 February 1998; revised 29 July 1998)

Abstract

We study an optimal control problem for a quasilinear parabolic equation which has delays in the highest order spatial derivative terms. The cost functional is Lagrange type and some terminal state constraints are presented. A Pontryagin-type maximum principle is derived.

1. Introduction

The classical Fourier's law

$$\mathbf{q}(t, x) = -k\nabla y(t, x), \quad (1.1)$$

relating the *heat flux* $\mathbf{q}(t, x)$ and the *temperature* $y(t, x)$, and the conservation of energy equation

$$cy_t(t, x) + \nabla \cdot \mathbf{q}(t, x) = 0, \quad (1.2)$$

lead to the classical heat equation:

$$y_t(t, x) = \frac{k}{c} \Delta y(t, x). \quad (1.3)$$

Equation (1.3) has an infinite propagation speed for a finite thermal pulse, which is not physical. A modified Fourier's law was introduced in [9] by Gurtin and Pipkin for general heat conduction problems of materials with memory (see [16] also):

$$\mathbf{q}(t, x) = - \int_0^\infty a(s) \nabla y(t-s, x) ds, \quad (1.4)$$

¹Laboratory of Mathematics for Nonlinear Sciences and Department of Mathematics, Fudan University, Shanghai 200433, China.

© Australian Mathematical Society 2001, Serial-fee code 0334-2700/01

where $a(s)$ is called the heat-flux relaxation function. If we assume a little more general memory pattern for the materials, the above might even be replaced by

$$\mathbf{q}(t, x) = - \int_0^\infty \nabla y(t - s, x) \mu(ds), \quad (1.5)$$

for some Borel measure $\mu(\cdot)$. Combining (1.5) with (1.2), we obtain

$$y_t(t, x) = \frac{k}{c} \int_0^\infty \Delta y(t - s, x) \mu(ds). \quad (1.6)$$

This is a parabolic equation with delays in the highest spatial derivative term. See [5, p. 176] (and [17] also) for an earlier study of a modified Fourier's law.

If further taking into account the source/sink of the heat (which could be regarded as a control), we will have a general nonhomogeneous linear parabolic equation having delays in the highest order spatial derivative terms. In the case that the material is nonlinear, the equation can be nonlinear. For extensive studies of such equations, see [2, 7, 20] and the recent book [18] by Tanabe.

In this paper, we study an optimal control problem for an abstract delay equation in a Hilbert space, motivated by the heat conduction problems discussed above. The following is a typical example of such systems:

$$\left\{ \begin{array}{l} y_t(t, x) - y_{xx}(t, x) \\ \quad = \int_{-r}^0 f(t, y_{xx}(t + \theta, x), u(t, x)) d\theta \\ \quad + \sum_{i \geq 1} g_i(t, y_{xx}(t - r_i, x), u(t, x)), \quad (t, x) \in (0, T) \times (0, 1), \\ y(t, x) = \varphi(t, x), \quad (t, x) \in [-r, 0] \times (0, 1), \\ y(t, 0) = y(t, 1) = 0, \quad t \in (0, T), \end{array} \right. \quad (1.7)$$

where the f and g_i 's are some maps and the $r_i \in (0, r]$ are some positive constants. In the above, $y(\cdot, \cdot)$ is regarded as the state and $u(\cdot, \cdot)$ is the control. The cost functional could be, say,

$$J(u(\cdot)) = \int_0^T \int_0^1 \left[\int_{-r}^0 f^0(t, y_{xx}(t + \theta, x), u(t, x)) d\theta + \sum_{i \geq 1} g_i^0(t, y_{xx}(t - r_i, x), u(t, x)) \right] dx dt. \quad (1.8)$$

We may also allow the system and the cost functional to be more general, say, in f , g_i and f^0 , g_i^0 , so that the terms $y(t, x)$, $y_x(t, x)$, $y_{xx}(t, x)$ and $y(t + \theta, x)$, $y_x(t + \theta, x)$,

$y(t - r_i, x), y_x(t - r_i, x)$ appear as well. On the other hand, we are allowed to have some terminal state constraints, of which the following is a typical example:

$$\int_0^1 \left\{ \int_{-r}^0 |y_{xx}(T + \theta, x)|^2 d\theta + |y(T, x)|^2 \right\} dx \leq 1. \tag{1.9}$$

Note that in (1.7), delays appear in the highest order spatial derivative terms and there are possibly countably many discrete delays. Moreover, the highest spatial derivatives also appear in the cost functional (1.8) and in the terminal state constraint (1.9). In [20], the authors established the existence and uniqueness as well as some properties of the solutions of equations more general than (1.7) (in an abstract setting). Based on the results of [20], we are going to derive a Pontryagin-type maximum principle for the corresponding optimal control problem in this paper. The major mathematical difficulty comes from the fact that since the control domain U is not necessarily convex, we have to use spike perturbation of the control. On the other hand, since the presented terminal state constraint is in some “small” space (relative to the space $M^2 \equiv H \times L^2(0, T; H)$, see Section 2), we need some fine estimates on the perturbed trajectories under the spike-type variation of the control. This will be achieved by an approximation technique. We point out that the usual vector measure technique used in [11, 12] cannot be directly applied here.

There is an extensive literature on the optimal control of infinite dimensional retarded evolution equations (see [1, 6, 13–15, 19, 21] and the references cited therein). Most of the above-mentioned works, however, deal with systems with delays appearing in the lower order spatial derivatives, except [6], which studies an LQ problem for a linear system with one (discrete) delay in the highest order spatial derivative term. For finite dimensional cases, see [3, 4]; and for general results of infinite dimensional optimal control theory, see [12] (as well as [8, 10, 11]).

The rest of the paper is organized as follows. In Section 2, we formulate our optimal control problem. In Section 3, we state the main result of this paper—the maximum principle. Section 4 is devoted to some auxiliary results which are crucial to the proof of the main result. A proof of the maximum principle is given in Section 5.

2. Formulation of the problem

Roughly speaking, we consider here the following abstract control system:

$$\begin{cases} \dot{x}(t) = Ax(t) + F(t, x_t, u(t)), & t \in (0, T), \\ x(0) = x_0, \quad x(t) = \varphi(t), & t \in [-r, 0). \end{cases} \tag{2.1}$$

The cost functional is given by

$$J(u(\cdot)) = \int_0^T F^0(t, x_t, u(t)) dt, \tag{2.2}$$

and the terminal state constraint takes the form:

$$\begin{pmatrix} x(T; u(\cdot)) \\ x_T(\cdot; u(\cdot)) \end{pmatrix} \in Q. \tag{2.3}$$

The optimal control problem is to minimize (2.2) subject to (2.1) and (2.3), over the set of all admissible controls.

We now make the above precise. Let H be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and the induced norm $|\cdot| \equiv |\cdot|_H$. Let $M^2 = H \times L^2(-r, 0; H)$ be endowed with an inner product

$$\left\langle \begin{pmatrix} \xi \\ \varphi(\cdot) \end{pmatrix}, \begin{pmatrix} \hat{\xi} \\ \hat{\varphi}(\cdot) \end{pmatrix} \right\rangle_{M^2} \triangleq \langle \xi, \hat{\xi} \rangle + \int_{-r}^0 \langle \varphi(\theta), \hat{\varphi}(\theta) \rangle d\theta, \quad \forall \begin{pmatrix} \xi \\ \varphi(\cdot) \end{pmatrix}, \begin{pmatrix} \hat{\xi} \\ \hat{\varphi}(\cdot) \end{pmatrix} \in M^2. \tag{2.4}$$

Then M^2 is a Hilbert space. Next, we let $A : \mathcal{D}(A) \subseteq H \rightarrow H$ generate an exponentially stable analytic semigroup e^{At} on H . Thus, without loss of generality, we may assume the following:

(H1) There exist constants $\delta_0, \omega, M_0, M_1, \hat{M} > 0$, such that

$$\rho(A) \supseteq \Sigma_{\delta_0} \triangleq \{ \lambda \in \mathbb{C} \mid |\arg \lambda| \leq \pi/2 + \delta_0 \} \cup \{0\}, \tag{2.5}$$

where $\rho(A)$ is the resolvent set of A ;

$$\|(\lambda I_H - A)^{-1}\|_{\mathcal{L}(H)} \leq \frac{M_0}{1 + |\lambda|}, \quad \forall \lambda \in \Sigma_{\delta_0}; \tag{2.6}$$

$$\|A e^{At}\|_{\mathcal{L}(H)} \leq M_1/t, \quad \forall t > 0; \tag{2.7}$$

$$\|e^{At}\|_{\mathcal{L}(H)} \leq \hat{M} e^{-\omega t}, \quad \forall t \geq 0. \tag{2.8}$$

In what follows, we endow $\mathcal{D}(A)$ with the inner product

$$\langle x, \hat{x} \rangle_A = \langle Ax, A\hat{x} \rangle, \quad \forall x, \hat{x} \in \mathcal{D}(A). \tag{2.9}$$

Then $(\mathcal{D}(A), \langle \cdot, \cdot \rangle_A)$ is a Hilbert space. We denote the induced norm by $|\cdot|_A$. Next, we introduce the space

$$X = \left\{ x \in H \mid \int_0^\infty |A e^{At} x|^2 dt < \infty \right\}, \tag{2.10}$$

and define

$$\langle x, y \rangle_X \triangleq \langle x, y \rangle + \int_0^\infty \langle A e^{At} x, A e^{At} y \rangle dt, \quad \forall x, y \in X. \tag{2.11}$$

Then $(X, \langle \cdot, \cdot \rangle_X)$ is a Hilbert space. We denote the induced norm by $|\cdot|_X$.

Since e^{A^t} is exponentially stable, the following operator $P \in \mathcal{L}(H)$ is well-defined:

$$P = \int_0^\infty e^{A^*t} e^{At} dt, \tag{2.12}$$

which is positive definite and is the unique solution of the Lyapunov equation

$$PA + A^*P = -I. \tag{2.13}$$

Consequently,

$$\int_0^\infty \langle e^{At}x, e^{At}y \rangle dt = \langle Px, y \rangle, \quad \forall x, y \in H. \tag{2.14}$$

Thus we have $\mathcal{D}(A) \subseteq X \subseteq H$ and

$$|x|_H \leq |x|_X \leq |x|_H + \|P^{1/2}\| |x|_A, \quad \forall x \in \mathcal{D}(A). \tag{2.15}$$

Note that $\mathcal{D}(A) \neq X \neq H$.

On the other hand, for any $\delta > 0$, we have (under (H1))

$$\begin{aligned} |x|_{X,\delta} &\equiv \left(|x|^2 + \int_0^\delta |Ae^{At}x|^2 dt \right)^{1/2} \leq |x|_X \leq \left(|x|^2 + \int_0^\delta |Ae^{At}x|^2 dt + \frac{M_1^2}{\delta} |x|^2 \right)^{1/2} \\ &\leq \left(1 + \frac{M_1}{\sqrt{\delta}} \right) |x|_{X,\delta}, \quad \forall x \in X. \end{aligned} \tag{2.16}$$

Thus $|\cdot|_X$ and $|\cdot|_{X,\delta}$ are equivalent.

Now let U be a metric space and $T > r > 0$. Denote

$$\mathcal{U} \equiv \{u(\cdot) : [0, T] \rightarrow U \mid u(\cdot) \text{ is measurable}\}. \tag{2.17}$$

Any element $u(\cdot) \in \mathcal{U}$ is called a control. We let $\Omega \equiv [0, T] \times [-r, 0] \times \mathcal{D}(A) \times \mathcal{D}(A) \times U$, and any element in Ω be denoted by (t, θ, x, y, u) . We introduce functions $f : \Omega \rightarrow H$ and $f^0 : \Omega \rightarrow \mathbb{R}$ which satisfy the following conditions.

(H2) $f \in C(\Omega; H)$, $f^0 \in C(\Omega; \mathbb{R})$; f and f^0 are Fréchet differentiable in x and y and

$$f_x, f_y \in C(\Omega; \mathcal{L}(\mathcal{D}(A); H)), \quad f_x^0, f_y^0 \in C(\Omega; \mathcal{D}(A)'). \tag{2.18}$$

Moreover, there exist constants $L_0, L_1, L_2 \geq 0$, such that for all $(t, \theta, x, y, u) \in \Omega$,

$$\begin{cases} \|f_x(t, \theta, x, y, u)\|_{\mathcal{L}(\mathcal{D}(A); H)} \leq L_1, \\ \|f_y(t, \theta, x, y, u)\|_{\mathcal{L}(\mathcal{D}(A); H)} \leq L_2, \\ \|f_x^0(t, \theta, x, y, u)\|_{\mathcal{D}(A)'}, \|f_y^0(t, \theta, x, y, u)\|_{\mathcal{D}(A)'} \leq L_0, \\ |f(t, \theta, 0, 0, u)|_H, |f^0(t, \theta, 0, 0, u)| \leq L_0, \end{cases} \tag{2.19}$$

where $\mathcal{D}(A)'$ is the dual of $(\mathcal{D}(A), |\cdot|_A)$ (as a Banach space).

Now we let μ and ν be nonnegative Radon measures on $\mathcal{B}[-r, 0]$, the Borel σ -field of $[-r, 0]$. Then for any $z(\cdot) \in C([-r, 0]; \mathcal{D}(A))$, $(t, u) \in [0, T] \times U$, we define

$$\begin{cases} F(t, z, u) = \int_{[-r,0]} f(t, \theta, z(0), z(\theta), u)\mu(d\theta), \\ F^0(t, z, u) = \int_{[-r,0]} f^0(t, \theta, z(0), z(\theta), u)\nu(d\theta). \end{cases} \tag{2.20}$$

These are the functions which appeared in (2.1) and (2.2).

The following result is a refinement of Lemma 3.2 found in [20]. As in [20], we may assume all the functions involved are Borel measurable (if necessary, we may make Borel measurable modifications).

LEMMA 2.1. *Let (H2) hold. Then, for any $x(\cdot) \in L^2(-r, T; \mathcal{D}(A))$ and $u(\cdot) \in \mathcal{U}$, the map $t \mapsto F(t, x_t, u(t))$ is in $L^2(0, T; H)$, where x_t stands for $x(t+\theta)$, $\theta \in [-r, 0]$. Moreover, for any $t_0 \in [0, T]$ and $\delta \in [0, r]$ with $t_0 + \delta \leq T$, we have that*

$$\begin{aligned} & \left(\int_{t_0}^{t_0+\delta} |F(t, x_t, u(t))|_H^2 dt \right)^{1/2} \\ & \leq \{L_1\mu([-r, 0]) + L_2\mu([-r, 0])^{1/2}\mu([-r, 0])^{1/2}\} |x(\cdot)|_{L^2(t_0, t_0+\delta; \mathcal{D}(A))} \\ & \quad + L_0\sqrt{\delta}\mu([-r, 0]) + L_2\mu([-r, 0])|x(\cdot)|_{L^2(t_0-r, t_0; \mathcal{D}(A))}. \end{aligned} \tag{2.21}$$

PROOF. The first conclusion follows from [20]. We now prove the estimate (2.21). By (2.19) and Lemma 2.2 of [20], we have

$$\begin{aligned} & \left(\int_{t_0}^{t_0+\delta} |F(t, x_t, u(t))|_H^2 dt \right)^{1/2} \\ & \leq L_0\sqrt{\delta}\mu([-r, 0]) + L_1\mu([-r, 0])|x(\cdot)|_{L^2(t_0, t_0+\delta; \mathcal{D}(A))} + (I), \end{aligned} \tag{2.22}$$

where

$$\begin{aligned} (I) & \equiv \left\{ \int_{t_0}^{t_0+\delta} \left[L_2 \int_{[-r,0]} |Ax(t+\theta)|\mu(d\theta) \right]^2 dt \right\}^{1/2} \\ & \leq L_2\mu([-r, 0])^{1/2} \left\{ \int_{[-r,0]} \int_{t_0+\theta}^{t_0+\theta+\delta} |Ax(t)|^2 dt \mu(d\theta) \right\}^{1/2} \\ & \leq L_2\mu([-r, 0])^{1/2} \left\{ \left(\int_{t_0}^{t_0+\delta} \int_{[-(t_0+\delta-t), 0]} |Ax(t)|^2 \mu(d\theta) dt \right)^{1/2} \right. \\ & \quad \left. + \left(\int_{t_0-r}^{t_0} \int_{[-(r \wedge (t_0+\delta-t)), t-t_0]} |Ax(t)|^2 \mu(d\theta) dt \right)^{1/2} \right\} \end{aligned}$$

$$\begin{aligned} &\leq L_2\mu([-r, 0])|x(\cdot)|_{L^2(t_0-r, t_0; \mathcal{D}(A))} \\ &\quad + L_2\mu([-r, 0])^{1/2}\mu([- \delta, 0])^{1/2}|x(\cdot)|_{L^2(t_0, t_0+\delta; \mathcal{D}(A))}. \end{aligned} \tag{2.23}$$

Then (2.21) follows.

It is easy to see that a similar result holds for the map F^0 . From [20], we have the following existence and uniqueness result for (2.1).

PROPOSITION 2.2. *Let (H1) and (H2) hold. Let the following also hold:*

(H3) *The nonnegative Radon measure μ satisfies*

$$\overline{\lim}_{s \downarrow 0} (M_0 + 1)\mu([-r, 0])^{1/2} [L_1\mu([-r, 0])^{1/2} + L_2\mu([-s, 0])^{1/2}] < 1. \tag{2.24}$$

Then, for any $x_0 \in X$, $\varphi(\cdot) \in L^2(-r, 0; \mathcal{D}(A))$ and $u(\cdot) \in \mathcal{U}$, there exists a unique strong solution $x(\cdot) \in L^2(0, T; \mathcal{D}(A)) \cap W^{1,2}([0, T]; H) \hookrightarrow C([0, T]; X)$ of (2.1).

Hereafter, the pair $(x_0, \varphi(\cdot)) \in X \times L^2(-r, 0; \mathcal{D}(A))$ will be fixed. Sometimes we write $x(\cdot; u(\cdot))$ to emphasize the dependence of the solution $x(\cdot)$ on the control $u(\cdot) \in \mathcal{U}$. Next, we let Q be a convex and closed subset in $X \times L^2(-r, 0; \mathcal{D}(A))$. Then, by Proposition 2.2, we see that the terminal state constraint (2.3) is meaningful. We let \mathcal{U}_{ad} be the set of all $u(\cdot) \in \mathcal{U}$ such that the corresponding trajectory $x(\cdot; u(\cdot))$ satisfies the constraint (2.3).

Clearly, under (H1)–(H3), for any $u(\cdot) \in \mathcal{U}$, the cost functional (2.2) is well-defined. The nonemptiness of the set \mathcal{U}_{ad} is the problem of controllability which will not be discussed here. In this paper, we consider the optimal control problem associated with the system (2.1), the terminal constraint (2.3) and the cost functional (2.2), assuming that \mathcal{U}_{ad} is nonempty. Our optimal control problem can be stated as follows.

PROBLEM (C). Find a control $\bar{u}(\cdot) \in \mathcal{U}_{ad}$, such that

$$J(\bar{u}(\cdot)) = \min_{u(\cdot) \in \mathcal{U}_{ad}} J(u(\cdot)). \tag{2.25}$$

Any $\bar{u}(\cdot) \in \mathcal{U}_{ad}$ satisfying (2.25) is referred to as an optimal control, the corresponding trajectory $x(\cdot; \bar{u}(\cdot))$ and the pair $(x(\cdot; \bar{u}(\cdot)), \bar{u}(\cdot))$ are called an optimal trajectory and pair, respectively. We will derive some necessary conditions for optimal pairs in the next section.

3. Main result

In this section, we state our main result and make some simple remarks. Let $\bar{u}(\cdot) \in \mathcal{U}_{ad}$ be an optimal control and $\bar{x}(\cdot) \equiv x(\cdot; \bar{u}(\cdot))$ be the corresponding optimal

trajectory. We now introduce the variational system along the pair $(\bar{x}(\cdot), \bar{u}(\cdot))$. For any $u(\cdot) \in \mathcal{U}$, let

$$\begin{cases} \dot{\xi}(t) = A\xi(t) + (DF)(t, \bar{x}_t, \bar{u}(t))\xi_t \\ \quad + F(t, \bar{x}_t, u(t)) - F(t, \bar{x}_t, \bar{u}(t)), & t \in (0, T), \\ \xi(t) = 0, & t \in [-r, 0], \text{ in } H, \end{cases} \quad (3.1)$$

where, for any $(x(\cdot), u(\cdot)) \in L^2(-r, T; \mathcal{D}(A)) \times \mathcal{U}$ and $\xi(\cdot) \in L^2(-r, T; \mathcal{D}(A))$,

$$\begin{aligned} (DF)(t, x_t, u(t))\xi_t &= \int_{[-r,0]} f_x(t, \theta, x(t), x(t + \theta), u(t))\xi(t)\mu(d\theta) \\ &\quad + \int_{[-r,0]} f_y(t, \theta, x(t), x(t + \theta), u(t))\xi(t + \theta)\mu(d\theta). \end{aligned} \quad (3.2)$$

We also define

$$\begin{aligned} (DF^0)(t, x_t, u(t))\xi_t &= \int_{[-r,0]} f_x^0(t, \theta, x(t), x(t + \theta), u(t))\xi(t)\nu(d\theta) \\ &\quad + \int_{[-r,0]} f_y^0(t, \theta, x(t), x(t + \theta), u(t))\xi(t + \theta)\nu(d\theta). \end{aligned} \quad (3.3)$$

Clearly (3.1) has the same form as (2.1). Thus, under (H1)–(H3), for any $u(\cdot) \in \mathcal{U}$, there exists a unique solution $\xi(\cdot) \equiv \xi(\cdot; u(\cdot))$ of (3.1), which is in $L^2(0, T; \mathcal{D}(A)) \cap W^{1,2}(0, T; H) \hookrightarrow C([0, T]; X)$. Set

$$\mathcal{R} = \left\{ \begin{pmatrix} \xi(T; u(\cdot)) \\ \xi_T(\cdot; u(\cdot)) \end{pmatrix} \mid u(\cdot) \in \mathcal{U} \right\} \subseteq X \times L^2(0, T; \mathcal{D}(A)). \quad (3.4)$$

We call \mathcal{R} the reachable set of the variational system (3.1). Next, we introduce the adjoint system of the variational system. To this end, let us first define the maps $(DF)'$ and $(DF^0)'$. For any $(x(\cdot), u(\cdot)) \in L^2(-r, T; \mathcal{D}(A)) \times \mathcal{U}$, $t \in [0, T]$ and $\psi(\cdot) \in L^2(0, T + r; H)$,

$$\begin{aligned} (DF)'(t, x_t, u(t))\psi' &= \int_{[-r,0]} f_x(t, \theta, x(t), x(t + \theta), u(t))'\psi(t)\mu(d\theta) \\ &\quad + \int_{[-r,0]} f_y(t - \theta, \theta, x(t - \theta), x(t), u(t - \theta))'\psi(t - \theta)\mu(d\theta), \end{aligned} \quad (3.5)$$

$$\begin{aligned} (DF^0)'(t, x_t, u(t)) &= \int_{[-r,0]} f_x^0(t, \theta, x(t), x(t + \theta), u(t))\nu(d\theta) \\ &\quad + \int_{[-r,0]} f_y^0(t - \theta, \theta, x(t - \theta), x(t), u(t - \theta))\nu(d\theta). \end{aligned} \quad (3.6)$$

In the above, ψ' stands for $\psi(t - \theta), \theta \in [-r, 0]$. We note that

$$f_x(t, \theta) \equiv f_x(t, \theta, x(t), x(t + \theta), u(t)) : \mathcal{D}(A) \rightarrow H \tag{3.7}$$

and

$$\|f_x(t, \theta)\|_{\mathcal{L}(\mathcal{D}(A); H)} \leq L_1. \tag{3.8}$$

Thus we may let $f_x(t, \theta)' : H' = H \rightarrow \mathcal{D}(A)'$ be the adjoint operator of $f_x(t, \theta)$. Clearly, for any $\psi(\cdot) \in L^2(0, T + \gamma; H)$, the first term on the right-hand side of (3.5) is well-defined. Similarly, the second term is also well-defined. In (3.5) and (3.6), we have taken the convention that the functions and measures are extended to be zero out of their domains. For example,

$$f_y(s, \theta, x, y, u) = 0, \quad f^0(s, \theta, x, y, u) = 0, \quad \text{if } s > T.$$

Clearly, we have

$$\begin{cases} (DF)'(\cdot, x, u(\cdot))\psi' \in L^2(0, T; \mathcal{D}(A)'), & \forall \psi(\cdot) \in L^2(0, T + r; H), \\ (DF^0)'(\cdot, x, u(\cdot)) \in L^2(0, T; \mathcal{D}(A)'). \end{cases} \tag{3.9}$$

By some routine calculations, we further have

$$\begin{cases} \int_0^T \langle (DF)(t, x_t, u(t))\xi_t, \psi(t) \rangle_H dt = \int_0^T \langle (DF)'(t, x_t, u(t))\psi', \xi(t) \rangle_{\mathcal{D}(A)', \mathcal{D}(A)} dt, \\ \int_0^T (DF^0)(t, x_t, u(t))\xi_t dt = \int_0^T \langle (DF^0)'(t, x_t, u(t)), \xi(t) \rangle_{\mathcal{D}(A)', \mathcal{D}(A)} dt. \end{cases} \tag{3.10}$$

On the other hand, since $A \in \mathcal{L}(\mathcal{D}(A); H)$, we may define $A' \in \mathcal{L}(H; \mathcal{D}(A)')$ to be the adjoint operator of A . Finally, we let $\psi^0 \in [-1, 0]$ and $\varphi(\cdot) \in L^2(-r, 0; \mathcal{D}(A)')$ (defining $\varphi(t)$ to be zero for $t \notin [-r, 0]$). Now the adjoint system (along the optimal pair $(\bar{x}(\cdot), \bar{u}(\cdot))$) reads as follows:

$$\begin{cases} \dot{\psi}(t) + A'\psi(t) + (DF)'(t, \bar{x}_t, \bar{u}(t))\psi^t \\ \quad + \psi^0(DF^0)'(t, \bar{x}_t, \bar{u}(t)) + \varphi(t - T) = 0, & t \in (0, T), \text{ in } \mathcal{D}(A)', \\ \psi(T) \in X', \quad \psi(t) = 0, & t > T. \end{cases} \tag{3.11}$$

Let us now define the Hamiltonian: $\forall(t, z(\cdot), u) \in [0, T] \times C([-r, 0]; \mathcal{D}(A)) \times U$ and $(\psi^0, \psi) \in \mathbb{R} \times H$,

$$H(t, z(\cdot), u, \psi^0, \psi) = \psi^0 F^0(t, z(\cdot), u) + \langle \psi, F(t, z(\cdot), u) \rangle. \tag{3.12}$$

As with F and F^0 , we see that for any $z \in L^2(-r, 0; \mathcal{D}(A))$, the above is also well-defined. We are now ready to state our main result.

THEOREM 3.1 (Maximum Principle). *Let (H1)–(H3) hold and let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be an optimal pair. Let the following also hold:*

(H4) *The set (recall (3.4) for the definition of \mathcal{R})*

$$\mathcal{R} - Q \equiv \{p - q \mid p \in \mathcal{R}, q \in Q\} \tag{3.13}$$

is finite codimensional ([11, 12]) in $X \times L^2(-r, 0; \mathcal{D}(A))$.

Then there exists a nontrivial pair $(\psi^0, \psi(\cdot)) \in [-1, 0] \times (L^2(0, T; H) \cap W^{1,2}([0, T]; \mathcal{D}(A)'))$, such that the adjoint system (3.11) with the data $(\psi(T), \varphi(\cdot)) \in X' \times L^2(-r, 0; \mathcal{D}(A))$ satisfies the transversality condition

$$\langle \psi(T), q_0 - \bar{x}(T) \rangle_{X', X} + \int_{-r}^0 \langle \varphi(\theta), q_1(\theta) - \bar{x}(T + \theta) \rangle_{\mathcal{D}(A)', \mathcal{D}(A)} d\theta \geq 0, \tag{3.14}$$

$$\forall (q_0, q_1(\cdot)) \in Q,$$

and the maximum condition holds:

$$H(t, \bar{x}_t, \bar{u}(t), \psi^0, \psi(t)) = \max_{u \in U} H(t, \bar{x}_t, u, \psi^0, \psi(t)), \quad \text{a.e. } t \in [0, T]. \tag{3.15}$$

REMARK 3.2. From [11, 12], we know that (H4) holds provided one of \mathcal{R} and Q is finite codimensional. The condition that Q be finite codimensional should be relatively easy to check.

4. Auxiliary lemmas

In this section, we present some lemmas which will play crucial roles in the proof of our main result.

LEMMA 4.1. *Let $h(\cdot) \in L^2(0, T; H)$. Then for any $\rho \in (0, 1)$ there exists a measurable set $E_\rho \subseteq [0, T]$ with*

$$|E_\rho| \triangleq \text{meas } E_\rho = \rho T, \tag{4.1}$$

such that the function

$$\eta_\rho(t) \equiv \int_0^t e^{A(t-\tau)} i \left(\frac{1}{\rho} \chi_{E_\rho}(\tau) - 1 \right) h(\tau) d\tau, \quad \tau \in [0, T], \tag{4.2}$$

satisfies

$$|\eta_\rho(\cdot)|_{L^2(0, T; \mathcal{D}(A)) \cap C([0, T]; X)} = o(1), \quad \text{as } \rho \rightarrow 0. \tag{4.3}$$

REMARK 4.2. Let us take $h(t) \equiv h_0 \in \mathcal{D}(A)$, with $|h_0| = 1$. Then

$$\dot{\eta}_\rho(t) = A\eta_\rho(t) + (\chi_{E_\rho}(t)/\rho - 1)h_0. \tag{4.4}$$

Since

$$\int_0^T \left| \left(\frac{1}{\rho} \chi_{E_\rho}(t) - 1 \right) h_0 \right|^2 dt = \frac{T}{\rho} - T, \tag{4.5}$$

we see that the following equation does not hold:

$$|\eta_\rho(\cdot)|_{L^2(0,T;\mathcal{D}(A)) \cap W^{1,2}([0,T];H)} = o(1), \quad \text{as } \rho \rightarrow 0. \tag{4.6}$$

Thus we could not expect to obtain

$$|\eta_\rho(\cdot)|_{C([0,T];X)} = o(1), \quad \text{as } \rho \rightarrow 0, \tag{4.7}$$

via proving (4.6) and applying the embedding $L^2(0, T; \mathcal{D}(A)) \cap W^{1,2}([0, T]; H) \hookrightarrow C([0, T]; X)$. However (4.7) is necessary for treating the state constraint $Q \subset X \times L^2(-r, 0; \mathcal{D}(A))$. Hence we have to prove (4.7) directly.

PROOF OF LEMMA 4.1. For any given $\rho \in (0, 1)$, we let $\tilde{h}(\cdot) \in C([0, T]; \mathcal{D}(A))$ such that

$$|h(\cdot) - \tilde{h}(\cdot)|_{L^2(0,T;H)} \leq \rho^2. \tag{4.8}$$

Since $A\tilde{h}(\cdot) \in C([0, T]; H)$ and $h(\cdot) \in L^2(0, T; H)$, as in [11, 12], we can find a measurable set $E_\rho \subseteq [0, T]$ with property (4.1), such that

$$\left| \frac{1}{\rho} \int_0^\cdot e^{A(\cdot-\tau)} A\tilde{h}(\tau) \chi_{E_\rho}(\tau) d\tau - \int_0^\cdot e^{A(\cdot-\tau)} A\tilde{h}(\tau) d\tau \right|_{C([0,T];H)} = o(1) \tag{4.9}$$

and

$$\left| \frac{1}{\rho} \int_0^\cdot h(\tau) \chi_{E_\rho}(\tau) d\tau - \int_0^\cdot h(\tau) d\tau \right|_{C([0,T];H)} = o(1). \tag{4.10}$$

Thus

$$\begin{aligned} |\eta_\rho(\cdot)|_{L^2(0,T;\mathcal{D}(A))} &\leq \left| \frac{1}{\rho} \int_0^\cdot e^{A(\cdot-\tau)} [h(\tau) - \tilde{h}(\tau)] \chi_{E_\rho}(\tau) d\tau \right|_{L^2(0,T;\mathcal{D}(A))} \\ &\quad + \left| \int_0^\cdot e^{A(\cdot-\tau)} [h(\tau) - \tilde{h}(\tau)] d\tau \right|_{L^2(0,T;\mathcal{D}(A))} \\ &\quad + \left| \frac{1}{\rho} \int_0^\cdot e^{A(\cdot-\tau)} \tilde{h}(\tau) \chi_{E_\rho}(\tau) d\tau - \int_0^\cdot e^{A(\cdot-\tau)} \tilde{h}(\tau) d\tau \right|_{L^2(0,T;\mathcal{D}(A))} \\ &\leq (M_0 + 1) \left(\frac{1}{\rho} + 1 \right) |h(\cdot) - \tilde{h}(\cdot)|_{L^2(0,T;H)} + o(1) = o(1). \end{aligned} \tag{4.11}$$

On the other hand, we see that $\eta_\rho(\cdot)$ is the strong solution of

$$\dot{\eta}_\rho(t) = A\eta_\rho(t) + (\chi_{E_\rho}(t)/\rho - 1)h(t), \quad \eta_\rho(0) = 0. \tag{4.12}$$

Hence, by (4.10) and the proved (4.11),

$$\begin{aligned}
 |\eta_\rho(t)| &\leq \int_0^T |A\eta_\rho(\tau)| d\tau + \left| \frac{1}{\rho} \int_0^t h(\tau)\chi_{E_\rho}(\tau) d\tau - \int_0^t h(\tau) d\tau \right| \\
 &\leq \sqrt{T} |\eta_\rho(\cdot)|_{L^2(0,T;\mathcal{D}(A))} + o(1) = o(1), \quad \text{uniformly in } t \in [0, T]. \quad (4.13)
 \end{aligned}$$

Take $\delta \in (0, T]$ to be fixed. We now look at the following:

$$\begin{aligned}
 &\left\{ \int_0^\delta |Ae^{As}\eta_\rho(t)|^2 ds \right\}^{1/2} \\
 &\leq \left\{ \int_0^\delta \left| A \int_0^t e^{A(s+t-\tau)} [h(\tau) - \tilde{h}(\tau)] d\tau \right|^2 ds \right\}^{1/2} \\
 &\quad + \left\{ \int_0^\delta \left| A \int_0^t e^{A(s+t-\tau)} \frac{1}{\rho} [h(\tau) - \tilde{h}(\tau)]\chi_{E_\rho}(\tau) d\tau \right|^2 ds \right\}^{1/2} \\
 &\quad + \left\{ \int_0^\delta \left| e^{As} \left[\frac{1}{\rho} \int_0^t e^{A(t-\tau)} A\tilde{h}(\tau)\chi_{E_\rho}(\tau) d\tau - \int_0^t e^{A(t-\tau)} A\tilde{h}(\tau) d\tau \right] \right|^2 ds \right\}^{1/2} \\
 &\leq \left\{ \int_t^{\delta+t} \left| A \int_0^s e^{A(s-\tau)} [h(\tau) - \tilde{h}(\tau)]\chi_{[0,t]}(\tau) d\tau \right|^2 ds \right\}^{1/2} \\
 &\quad + \left\{ \int_t^{t+\delta} \left| A \int_0^s e^{A(s-\tau)} \frac{1}{\rho} [h(\tau) - \tilde{h}(\tau)]\chi_{E_\rho}(\tau)\chi_{[0,t]}(\tau) d\tau \right|^2 ds \right\}^{1/2} \\
 &\quad + \hat{M}\sqrt{T} \left| \frac{1}{\rho} \int_0^t e^{A(-\tau)} A\tilde{h}(\tau)\chi_{E_\rho}(\tau) d\tau - \int_0^t e^{A(-\tau)} A\tilde{h}(\tau) d\tau \right|_{C([0,T];H)} \\
 &\leq (M_0 + 1) \left(\frac{1}{\rho} + 1 \right) \left\{ \int_0^{t+\delta} |h(\tau) - \tilde{h}(\tau)|^2 \chi_{[0,T]}(\tau) d\tau \right\}^{1/2} + o(1) \\
 &\leq (M_0 + 1) \left(\frac{1}{\rho} + 1 \right) |h(\cdot) - \tilde{h}(\cdot)|_{L^2(0,T;H)} + o(1) = o(1), \quad (4.14)
 \end{aligned}$$

uniformly in $t \in [0, T]$. Hence, by the observation we made in Section 2 about the norm of X , we have

$$|\eta_\rho(\cdot)|_{C([0,T];X)} = o(1), \quad \text{as } \rho \rightarrow 0. \quad (4.15)$$

Combining (4.11) with the above, we obtain (4.3).

Next, let us present a result concerned with the continuous dependence of the solution of evolution equations on the data.

LEMMA 4.3. Let the operator-valued function $G(\cdot) : L^2(-r, T; \mathcal{D}(A)) \rightarrow L^2(0, T; H)$ satisfy the following: $\forall t_0 \in [0, T)$ and $\delta > 0$ with $t_0 + \delta \leq T$,

$$\left(\int_{t_0}^{t_0+\delta} |G(t)\hat{\zeta}_t|_H^2 dt \right)^{1/2} \leq \bar{\omega}(\delta)|\hat{\zeta}(\cdot)|_{L^2(t_0, t_0+\delta; \mathcal{D}(A))} + K|\hat{\zeta}(\cdot)|_{L^2(t_0-r, t_0; \mathcal{D}(A))}, \quad (4.16)$$

$$\forall \hat{\zeta}(\cdot) \in L^2(-r, T; \mathcal{D}(A)),$$

with $K > 0$ and $\bar{\omega}(\cdot) \in C(\mathbb{R}^+; \mathbb{R}^+)$, $(M_0 + 1)\bar{\omega}(0) < 1$, independent of $\hat{\zeta}(\cdot)$. Let $h(\cdot) \in L^2(0, T; H)$, $(x, \varphi(\cdot)) \in X \times L^2(-r, 0; \mathcal{D}(A))$ and let $\zeta(\cdot) \in L^2(0, T; \mathcal{D}(A)) \cap W^{1,2}(0, T; H)$ be the strong solution of

$$\begin{cases} \dot{\zeta}(t) = A\zeta(t) + G(t)\zeta_t + h(t), & t \in (0, T], \text{ in } H, \\ \zeta(0) = x, \quad \zeta(t) = \varphi(t), & t \in [-r, 0). \end{cases} \quad (4.17)$$

Then there exists a constant $C > 0$, such that

$$\begin{aligned} & |\zeta(\cdot)|_{L^2(0, T; \mathcal{D}(A))} + |\zeta(\cdot)|_{C([0, T]; X)} \\ & \leq C \left\{ \left| \int_0^{\cdot} e^{A(\cdot-\tau)} h(\tau) d\tau \right|_{L^2(0, T; \mathcal{D}(A))} + \left| \int_0^{\cdot} e^{A(\cdot-\tau)} h(\tau) d\tau \right|_{C([0, T]; X)} \right. \\ & \quad \left. + |h(\cdot)|_{L^1(0, T; H)} + |x|_X + |\varphi(\cdot)|_{L^2(-r, 0; \mathcal{D}(A))} \right\}. \end{aligned} \quad (4.18)$$

PROOF. Take $\delta \equiv T/N > 0$ such that

$$\omega_0 \equiv (M_0 + 1)\bar{\omega}(\delta) < 1. \quad (4.19)$$

For $t \in [i\delta, (i + 1)\delta]$, we have

$$\zeta(t) = e^{A(t-i\delta)}\zeta(i\delta) + \int_{i\delta}^t e^{A(t-\tau)}G(\tau)\zeta_\tau d\tau + \int_{i\delta}^t e^{A(t-\tau)}h(\tau) d\tau. \quad (4.20)$$

Then

$$\begin{aligned} & |\zeta(\cdot)|_{L^2(i\delta, (i+1)\delta; \mathcal{D}(A))} \\ & \leq \left(\int_{i\delta}^{(i+1)\delta} |Ae^{A(t-i\delta)}\zeta(i\delta)|^2 dt \right)^{1/2} + (M_0 + 1) \left(\int_{i\delta}^{(i+1)\delta} |G(t)\zeta_t|^2 dt \right)^{1/2} \\ & \quad + \left(\int_{i\delta}^{(i+1)\delta} \left| A \int_{i\delta}^t e^{A(t-\tau)} h(\tau) d\tau \right|^2 dt \right)^{1/2} \\ & \leq \left(\int_0^\delta |Ae^{A^1}\zeta(i\delta)|^2 dt \right)^{1/2} + \omega_0 |\zeta(\cdot)|_{L^2(i\delta, (i+1)\delta; \mathcal{D}(A))} \\ & \quad + (M_0 + 1)K|\zeta(\cdot)|_{L^2(i\delta-r, i\delta; \mathcal{D}(A))} + I_i, \end{aligned} \quad (4.21)$$

where

$$\begin{aligned}
 I_i &= \left(\int_{i\delta}^{(i+1)\delta} \left| A \int_{i\delta}^t e^{A(t-\tau)} h(\tau) d\tau \right|^2 dt \right)^{1/2} \\
 &\leq \left(\int_{i\delta}^{(i+1)\delta} \left| A \int_0^t e^{A(t-\tau)} h(\tau) d\tau \right|^2 dt \right)^{1/2} \\
 &\quad + \left(\int_{i\delta}^{(i+1)\delta} \left| A \int_0^{i\delta} e^{A(t-\tau)} h(\tau) d\tau \right|^2 dt \right)^{1/2} \\
 &\leq \left| \int_0^{i\delta} e^{A(\cdot-\tau)} h(\tau) d\tau \right|_{L^2(0,T;\mathcal{D}(A))} + \left(\int_0^\delta \left| A e^{At} \int_0^{i\delta} e^{A(i\delta-\tau)} h(\tau) d\tau \right|^2 dt \right)^{1/2} \\
 &\leq \left| \int_0^{i\delta} e^{A(\cdot-\tau)} h(\tau) d\tau \right|_{L^2(0,T;\mathcal{D}(A))} + \left| \int_0^{i\delta} e^{A(\cdot-\tau)} h(\tau) d\tau \right|_{C([0,T];X)} \equiv \hat{I}. \tag{4.22}
 \end{aligned}$$

Hence

$$|\zeta(\cdot)|_{L^2(i\delta, (i+1)\delta; \mathcal{D}(A))} \leq \frac{1}{1-\omega_0} [|\zeta(i\delta)|_X + (M_0 + 1)K |\zeta(\cdot)|_{L^2(i\delta-r, i\delta; \mathcal{D}(A))} + \hat{I}]. \tag{4.23}$$

On the other hand,

$$\begin{aligned}
 |\zeta(t)| &\leq |\zeta(i\delta)| + \int_{i\delta}^t |A\zeta(\tau)| d\tau + \int_{i\delta}^t |G(\tau)\zeta_\tau| d\tau + \int_{i\delta}^t |h(\tau)| d\tau \\
 &\leq |\zeta(i\delta)|_X + \sqrt{\delta} \{ |\zeta(\cdot)|_{L^2(i\delta, (i+1)\delta; \mathcal{D}(A))} + |G(\cdot)\zeta_\cdot|_{L^2(i\delta, (i+1)\delta; H)} \} \\
 &\quad + \int_{i\delta}^{(i+1)\delta} |h(\tau)| d\tau \\
 &\leq C [|\zeta(i\delta)|_X + |\zeta(\cdot)|_{L^2(i\delta-r, i\delta; \mathcal{D}(A))} + \hat{I}] + \int_{i\delta}^{(i+1)\delta} |h(\tau)| d\tau. \tag{4.24}
 \end{aligned}$$

Also, for $t \in [i\delta, (i + 1)\delta]$, we have

$$\begin{aligned}
 &\left(\int_0^\delta |A e^{As} \zeta(t)|^2 ds \right)^{1/2} \\
 &\leq \left(\int_0^\delta |A e^{A(s+t-i\delta)} \zeta(i\delta)|^2 ds \right)^{1/2} + \left(\int_0^\delta \left| A \int_{i\delta}^t e^{A(s+t-\tau)} G(\tau)\zeta_\tau d\tau \right|^2 ds \right)^{1/2} \\
 &\quad + \left(\int_0^\delta \left| A e^{As} \int_{i\delta}^t e^{A(t-\tau)} h(\tau) d\tau \right|^2 ds \right)^{1/2} \\
 &\leq \hat{M} \left(\int_0^\delta |A e^{As} \zeta(i\delta)|^2 ds \right)^{1/2} + \left(\int_t^{t+\delta} \left| A \int_{i\delta}^t e^{A(s-\tau)} G(\tau)\zeta_\tau d\tau \right|^2 ds \right)^{1/2}
 \end{aligned}$$

$$\begin{aligned}
 & + \left(\int_0^\delta \left| A e^{As} \int_0^t e^{A(t-\tau)} h(\tau) d\tau \right|^2 ds \right)^{1/2} \\
 & + \left(\int_0^\delta \left| A e^{As} \int_0^{i\delta} e^{A(t-\tau)} h(\tau) d\tau \right|^2 ds \right)^{1/2} \\
 \leq & \hat{M} |\zeta(i\delta)|_X + \left(\int_{i\delta}^{(i+1)\delta} \left| A \int_{i\delta}^s e^{A(s-\tau)} G(\tau) \zeta_\tau \chi_{[i\delta, t]}(\tau) d\tau \right|^2 ds \right)^{1/2} \\
 & + \left| \int_0^\cdot e^{A(\cdot-\tau)} h(\tau) d\tau \right|_{C([0, T]; X)} \\
 & + \left(\int_0^\delta \left| e^{A(t-i\delta)} A e^{As} \int_0^{i\delta} e^{A(i\delta-\tau)} h(\tau) d\tau \right|^2 ds \right)^{1/2} \\
 \leq & \hat{M} |\zeta(i\delta)|_X + (M_0 + 1) \left(\int_{i\delta}^{(i+1)\delta} |G(\tau) \zeta_\tau|^2 d\tau \right)^{1/2} \\
 & + C \left| \int_0^\cdot e^{A(\cdot-\tau)} h(\tau) d\tau \right|_{C([0, T]; X)} \\
 \leq & C [|\zeta(i\delta)|_X + |\zeta(\cdot)|_{L^2(i\delta-r, i\delta; \mathcal{D}(A))} + \hat{I}]. \tag{4.25}
 \end{aligned}$$

This leads to

$$|\zeta(\cdot)|_{C[i\delta, (i+1)\delta]; X} \leq C \left[|\zeta(i\delta)|_X + |\zeta(\cdot)|_{L^2(i\delta-r, i\delta; \mathcal{D}(A))} + \hat{I} \right]. \tag{4.26}$$

Combining (4.23) and (4.26), we obtain (4.18) by induction.

5. Proof of Theorem 3.1

In this section, we give a proof of Theorem 3.1. Let $\bar{u}(\cdot) \in \mathcal{U}_{ad}$ be the given optimal control and $\bar{x}(\cdot)$ be the corresponding optimal trajectory. We denote

$$\begin{aligned}
 d_Q(x, z) & = \inf_{(\hat{x}, \hat{z}) \in \mathcal{Q}} \left\{ |x - \hat{x}|_X^2 + |z - \hat{z}|_{L^2(-r, 0; \mathcal{D}(A))}^2 \right\}^{1/2}, \tag{5.1} \\
 & \forall (x, z) \in X \times L^2(-r, 0; \mathcal{D}(A)),
 \end{aligned}$$

$$\bar{d}(u(\cdot), \hat{u}(\cdot)) = \text{meas}\{t \in [0, T] \mid u(t) \neq \hat{u}(t)\}, \quad \forall u(\cdot), \hat{u}(\cdot) \in \mathcal{U}. \tag{5.2}$$

We know that $(\mathcal{U}, \bar{d}(\cdot, \cdot))$ is a complete metric space ([11, 12]). For any $u(\cdot) \in \mathcal{U}$, we denote

$$x^0(u(\cdot)) = J(u(\cdot)). \tag{5.3}$$

Next, for any $\varepsilon > 0$ and any $u(\cdot) \in \mathcal{U}$ (instead of \mathcal{U}_{ad} !), we define

$$J_\varepsilon(u(\cdot)) = \left\{ \left[(x^0(u(\cdot)) - x^0(\bar{u}(\cdot)) + \varepsilon)^+ \right]^2 + d_Q(x(T; u(\cdot)), x_T(\cdot; u(\cdot)))^2 \right\}^{1/2}. \tag{5.4}$$

Clearly $J_\varepsilon(\cdot) : (\mathcal{U}, \bar{d}) \rightarrow \mathbb{R}$ is continuous. Also, we have

$$\begin{cases} J_\varepsilon(u(\cdot)) > 0, & \forall u(\cdot) \in \mathcal{U}, \\ J_\varepsilon(\bar{u}(\cdot)) = \varepsilon \leq \inf_{u(\cdot) \in \mathcal{U}} J_\varepsilon(u(\cdot)) + \varepsilon. \end{cases} \tag{5.5}$$

Thus, by Ekeland’s variational principle ([12]), we have $u^\varepsilon(\cdot) \in \mathcal{U}$, such that

$$\begin{cases} \bar{d}(u^\varepsilon(\cdot), \bar{u}(\cdot)) \leq \sqrt{\varepsilon}, \\ J_\varepsilon(u^\varepsilon(\cdot)) \leq J_\varepsilon(\bar{u}(\cdot)), \\ -\sqrt{\varepsilon} \bar{d}(u^\varepsilon(\cdot), \hat{u}(\cdot)) \leq J_\varepsilon(\hat{u}(\cdot)) - J_\varepsilon(u^\varepsilon(\cdot)), \quad \forall \hat{u}(\cdot) \in \mathcal{U}. \end{cases} \tag{5.6}$$

We let

$$x^\varepsilon(\cdot) = x(\cdot; u^\varepsilon(\cdot)), \quad x^{0,\varepsilon} = x^0(u^\varepsilon(\cdot)). \tag{5.7}$$

We now take $u(\cdot) \in \mathcal{U}$ fixed. Define

$$\begin{cases} \Delta F^\varepsilon(\cdot) \equiv F(\cdot, (x^\varepsilon), u(\cdot)) - F(\cdot, (x^\varepsilon), u^\varepsilon(\cdot)) \in L^2(0, T; H), \\ \Delta F^{0,\varepsilon}(\cdot) \equiv F^0(\cdot, (x^\varepsilon), u(\cdot)) - F^0(\cdot, (x^\varepsilon), u^\varepsilon(\cdot)) \in L^2(0, T; \mathbb{R}). \end{cases} \tag{5.8}$$

By Lemma 4.1, together with a technique used in [11, 12] (which has been used in the proof of Lemma 4.1.), we may find a measurable set $E_\rho \subseteq [0, T]$ with $|E_\rho| = \rho T$, such that as $\rho \rightarrow 0$,

$$\begin{cases} \left| \int_0^\cdot e^{A(\cdot-\tau)} \left(\frac{1}{\rho} \chi_{E_\rho}(\tau) - 1 \right) \Delta F^\varepsilon(\tau) d\tau \right|_{L^2(0,T;\mathcal{D}(A)) \cap C([0,T];X)} = o(1), \\ \left| \int_0^\cdot \left(\frac{1}{\rho} \chi_{E_\rho}(\tau) - 1 \right) \Delta F^{0,\varepsilon}(\tau) d\tau \right|_{C([0,T];\mathbb{R})} = o(1). \end{cases} \tag{5.9}$$

We now define

$$u_\rho^\varepsilon(\cdot) = u(\cdot) \chi_{E_\rho}(\cdot) + u^\varepsilon(\cdot) \chi_{[0,T] \setminus E_\rho}(\cdot) \tag{5.10}$$

and set

$$x_\rho^\varepsilon(\cdot) = x(\cdot; u_\rho^\varepsilon(\cdot)), \quad x_\rho^{0,\varepsilon} = x^0(u_\rho^\varepsilon(\cdot)). \tag{5.11}$$

Let $\xi^\varepsilon(\cdot) \in L^2(0, T; \mathcal{D}(A)) \cap W^{1,2}([0, T]; H)$ be the unique strong solution of

$$\begin{cases} \dot{\xi}^\varepsilon(t) = A\xi^\varepsilon(t) + (DF)(t, (x^\varepsilon)_t, u^\varepsilon(t))(\xi^\varepsilon)_t + \Delta F^\varepsilon(t), & \text{a.e. } t \in (0, T], \text{ in } H, \\ \xi^\varepsilon(t) = 0, & t \in [-r, 0]. \end{cases} \tag{5.12}$$

Next, define

$$w_\rho^\varepsilon(\cdot) = (1/\rho)(x_\rho^\varepsilon(\cdot) - x^\varepsilon(\cdot)) - \xi^\varepsilon(\cdot). \tag{5.13}$$

We claim that

$$|w_\rho^\varepsilon(\cdot)|_{L^2(0,T;\mathcal{D}(A)) \cap C([0,T];X)} = o(1), \quad \text{as } \rho \rightarrow 0. \tag{5.14}$$

In fact, $w_\rho^\varepsilon(\cdot)$ satisfies

$$\begin{cases} \dot{w}_\rho^\varepsilon(t) = Aw_\rho^\varepsilon(t) + (DF)(t, (x^\varepsilon)_t, u^\varepsilon(t))(w_\rho^\varepsilon)_t \\ \quad + \frac{1}{\rho} \Delta F^\varepsilon(t) \chi_{E_\rho}(t) - \Delta F^\varepsilon(t) + h_\rho^\varepsilon(t), \quad \text{a.e. } t \in (0, T], \text{ in } H, \\ w_\rho^\varepsilon(t) = 0, \quad t \in [-r, 0], \end{cases} \tag{5.15}$$

where (denoting $(DF)_\rho^\varepsilon(t, x_t) = (DF)(t, x_t, u^\varepsilon(t))$)

$$\begin{aligned} h_\rho^\varepsilon(t) = & \int_0^1 \{ (DF)_\rho^\varepsilon(t, (x^\varepsilon)_t + \sigma[(x_\rho^\varepsilon)_t - (x^\varepsilon)_t]) \\ & - (DF)(t, (x^\varepsilon)_t, u^\varepsilon(t)) \} d\sigma \frac{1}{\rho} [(x_\rho^\varepsilon)_t - (x^\varepsilon)_t]. \end{aligned} \tag{5.16}$$

Direct computation shows that

$$\lim_{\rho \rightarrow 0} |x_\rho^\varepsilon(\cdot) - x^\varepsilon(\cdot)|_{L^2(0,T;\mathcal{D}(A)) \cap C([0,T];X)} = 0. \tag{5.17}$$

Thus, by (H2), we have

$$\begin{aligned} \delta_\rho(t) \equiv & \int_0^1 \| (DF)_\rho^\varepsilon(t, (x^\varepsilon)_t + \sigma[(x_\rho^\varepsilon)_t - (x^\varepsilon)_t]) \\ & - (DF)(t, (x^\varepsilon)_t, u^\varepsilon(t)) \|_{\mathcal{L}(L^2(-r,0;\mathcal{D}(A));H)} d\sigma \rightarrow 0, \quad \text{as } \rho \rightarrow 0. \end{aligned} \tag{5.18}$$

Consequently,

$$|h_\rho^\varepsilon(\cdot)|_{L^2(0,T;H)}^2 \leq \int_0^T \delta_\rho(t)^2 |(w_\rho^\varepsilon)_t + (\xi^\varepsilon)_t|_{L^2(-r,0;\mathcal{D}(A))}^2 dt. \tag{5.19}$$

Now, by Lemma 4.3, we obtain (5.14). As a consequence, we have

$$\xi^{0,\varepsilon} \equiv \lim_{\rho \rightarrow 0} \frac{x_\rho^{0,\varepsilon} - x^{0,\varepsilon}}{\rho} = \int_0^T (DF)^{0,\varepsilon}(t)(\xi^\varepsilon)_t dt + \int_0^T \Delta F^{0,\varepsilon}(t) dt, \tag{5.20}$$

where (see (3.3))

$$(DF)^{0,\varepsilon}(t)(\xi^\varepsilon)_t = (DF^0)(t, x_t^\varepsilon, u^\varepsilon(t))(\xi^\varepsilon)_t.$$

Hence, by taking $\hat{u}(\cdot) = u_\rho^\varepsilon(\cdot)$ in (5.6), we get

$$-\sqrt{\varepsilon}T \leq [J_\varepsilon(u_\rho^\varepsilon(\cdot)) - J_\varepsilon(u^\varepsilon(\cdot))]/\rho. \tag{5.21}$$

Letting $\rho \rightarrow 0$, we obtain

$$-\sqrt{\varepsilon}T \leq \psi^{0,\varepsilon}\xi^{0,\varepsilon} + \langle \psi^\varepsilon, \xi^\varepsilon(T) \rangle_{X',X} + \langle \varphi^\varepsilon, (\xi^\varepsilon)_T \rangle_{L^2(-r,0;\mathcal{D}(A)'),L^2(-r,0;\mathcal{D}(A))}, \tag{5.22}$$

where

$$\begin{cases} \psi^{0,\varepsilon} = \frac{(x^0(u^\varepsilon(\cdot)) - x^0(\bar{u}(\cdot)) + \varepsilon)^+}{J_\varepsilon(u^\varepsilon(\cdot))}, \\ \begin{pmatrix} \psi^\varepsilon \\ \varphi^\varepsilon(\cdot) \end{pmatrix} = \frac{d_Q(x^\varepsilon(T), (x^\varepsilon)_T) \nabla d_Q(x^\varepsilon(T), (x^\varepsilon)_T)}{J_\varepsilon(u^\varepsilon(\cdot))}. \end{cases} \tag{5.23}$$

Here we should note that the space $X' \times L^2(-r, 0; \mathcal{D}(A)')$ is strictly convex (since it is the dual of the Hilbert space $X \times L^2(-r, 0; \mathcal{D}(A))$). Thus, the above second equality makes sense. On the other hand, by the convexity of Q , we see that

$$\langle \psi^\varepsilon, q_0 - x^\varepsilon(T) \rangle_{X',X} + \langle \varphi^\varepsilon(\cdot), q_1(\cdot) - (x^\varepsilon)_T \rangle_{L^2(-r,0;\mathcal{D}(A)'),L^2(-r,0;\mathcal{D}(A))} \leq 0, \tag{5.24}$$

$$\forall (q_0, q_1(\cdot)) \in Q.$$

Thus, combining (5.22)–(5.24) and the finite codimensionality of $\mathcal{R} - Q$ in $X \times L^2(-r, 0; \mathcal{D}(A))$, in a manner similar to that used in [11, 12], for some sequence $\varepsilon \rightarrow 0$, we have

$$(\psi^{0,\varepsilon}, \psi^\varepsilon, \varphi^\varepsilon(\cdot)) \xrightarrow{w} (\bar{\psi}^0, \bar{\psi}, \bar{\varphi}(\cdot)) \neq 0. \tag{5.25}$$

Taking limits in (5.22), one has

$$0 \leq \bar{\psi}^0 \xi^0 + \langle \bar{\psi}, \xi(T) \rangle_{X',X} + \langle \bar{\varphi}, \xi_T \rangle_{L^2(-r,0;\mathcal{D}(A)'),L^2(-r,0;\mathcal{D}(A))}, \tag{5.26}$$

where $\xi^0 \in \mathbb{R}$ and $\xi(\cdot) \in L^2(0, T; \mathcal{D}(A)) \cap W^{1,2}(0, T; H)$ satisfy

$$\begin{cases} \dot{\xi}(t) = A\xi(t) + (DF)(t, \bar{x}_t, \bar{u}(t))\xi_t + \Delta F(t), & \text{a.e. } t \in [0, T], \text{ in } H, \\ \xi(t) = 0, & t \in [-r, 0], \end{cases} \tag{5.27}$$

and

$$\xi^0 = \int_0^T [(DF^0)(t, \bar{x}_t, \bar{u}(t))\xi_t + \Delta F^0(t)] dt, \tag{5.28}$$

where

$$\begin{cases} \Delta F(t) = F(t, \bar{x}_t, u(t)) - F(t, \bar{x}_t, \bar{u}(t)), & t \in [0, T], \\ \Delta F^0(t) = F^0(t, \bar{x}_t, u(t)) - F^0(t, \bar{x}_t, \bar{u}(t)). \end{cases} \tag{5.29}$$

We now define a linear functional $\mathcal{F} : L^2(0, T; H) \rightarrow \mathbb{R}$ as follows: for any $\Delta F(\cdot) \in L^2(0, T; H)$, we let $\xi(\cdot) \in L^2(0, T; \mathcal{D}(A)) \cap W^{1,2}(0, T; H)$ be the solution of (5.27). Then let

$$\begin{aligned} \mathcal{F}(\Delta F(\cdot)) &= - \int_0^T \bar{\psi}^0(DF^0)(t, \bar{x}_t, \bar{u}(t))\xi_t dt - \langle \bar{\psi}, \xi(T) \rangle_{X', X} \\ &\quad - \langle \bar{\varphi}, \xi_T \rangle_{L^2(-r, 0; \mathcal{D}(A)'), L^2(-r, 0; \mathcal{D}(A))}. \end{aligned} \tag{5.30}$$

It is immediate that \mathcal{F} is bounded. Thus, by the Riesz representation theorem, we can find a $\psi(\cdot) \in L^2(0, T; H)$, such that

$$\mathcal{F}(\Delta F(\cdot)) = \langle \psi(\cdot), \Delta F(\cdot) \rangle_{L^2(0, T; H)}, \quad \forall \Delta F(\cdot) \in L^2(0, T; H). \tag{5.31}$$

This means that

$$\begin{aligned} \int_0^T \langle \psi(t), \dot{\xi}(t) \rangle_H dt &= \int_0^T [\langle \psi(t), A\xi(t) + (DF)(t, \bar{x}_t, \bar{u}(t))\xi_t \rangle_H \\ &\quad - \bar{\psi}^0(DF^0)(t, \bar{x}_t, \bar{u}(t))\xi_t] dt - \langle \bar{\psi}, \xi(T) \rangle_{X', X} \\ &\quad - \int_{-r}^0 \langle \bar{\varphi}(\theta), \xi(T + \theta) \rangle_{\mathcal{D}(A)', \mathcal{D}(A)} d\theta, \\ \forall \xi(\cdot) &\in L^2(-r, T; \mathcal{D}(A)) \cap W^{1,2}(0, T; H). \end{aligned} \tag{5.32}$$

We now set

$$\psi^0 = -\bar{\psi}^0, \quad \psi(T) = -\bar{\psi}, \quad \varphi(\cdot) = -\bar{\varphi}(\cdot). \tag{5.33}$$

Clearly $\psi(\cdot) \in W^{1,2}([0, T]; \mathcal{D}(A)')$. Then, by direct computation, we see that $\psi(\cdot)$ satisfies the adjoint system (3.11). Furthermore,

$$\begin{aligned} 0 &\geq \int_0^T \{ \psi^0 [F^0(t, \bar{x}_t, u(t)) - F^0(t, \bar{x}_t, \bar{u}(t))] + \langle \psi(t), F(t, \bar{x}_t, u(t)) \\ &\quad - F(t, \bar{x}_t, \bar{u}(t)) \rangle_H \} dt. \end{aligned} \tag{5.34}$$

The above holds for all $u(\cdot) \in \mathcal{U}$. Thus the maximum condition (3.9) follows. Finally, the transversality condition (3.14) follows from (5.24). Hence the proof of Theorem 3.1 is complete.

Acknowledgements

The second author was supported in part by the NSFC under Grant 79790130, the National Distinguished Youth Science Foundation of China under Grant 19725106, the Chinese State Education Commission Science Foundation under Grant 2000024605, and the Cheung Kong Scholars Programme.

References

- [1] N. U. Ahmed and K. L. Teo, "On the optimal controls of a class of systems governed by second order parabolic partial delay-differential equations with first boundary conditions", *Ann. Mat. Pura. Appl.* **122** (1979) 61–82.
- [2] A. Ardito and P. Ricciardi, "Existence and regularity for linear delay partial differential equations", *Nonlinear Anal.* **4** (1980) 411–414.
- [3] H. T. Banks, "Necessary conditions for control problems with variable time lags", *SIAM J. Control* **6** (1968) 9–47.
- [4] H. T. Banks, "Variational problems involving functional differential equations", *SIAM J. Control* **7** (1969) 1–17.
- [5] K. H. Beckurts and K. Wirtz, *Neutron physics* (Springer, Berlin, 1964).
- [6] G. Di Blasio, "The linear-quadratic optimal control problem for delay equations", *Rend. Accad. Naz. Lincei* **71** (1981) 156–161.
- [7] G. Di Blasio, K. Kunisch and E. Sinestrari, " L^2 -regularity for parabolic partial integrodifferential equations with delay in the highest-order derivatives", *J. Math. Anal. Appl.* **102** (1984) 38–57.
- [8] H. O. Fattorini and H. Frankowska, "Necessary conditions for infinite-dimensional control problems", *Math. Control Signals Sys.* **4** (1991) 41–67.
- [9] M. E. Gurtin and A. C. Pipkin, "A general theory of heat conduction with finite wave speeds", *Arch. Rat. Mech. Anal.* **31** (1968) 113–126.
- [10] X. Li and Y. Yao, "Maximum principle of distributed parameter systems with time lags", in *Distributed parameter systems*, Lecture Notes in Control and Inform. Sci. 75, (Springer, 1985) 410–427.
- [11] X. Li and J. Yong, "Necessary conditions of optimal control for distributed parameter systems", *SIAM J. Control Optim.* **29** (1991) 895–908.
- [12] X. Li and J. Yong, *Optimal control theory for infinite dimensional systems* (Birkhäuser, Boston, 1995).
- [13] S. Nababan and E. S. Noussair, "On the existence of optimal controls of systems governed by quasilinear parabolic partial delay-differential equations", *Internat. T. Systems Sci.* **17** (1986) 1245–1259.
- [14] S. Nababan and K. L. Teo, "Necessary conditions for optimal controls for systems governed by parabolic partial delay-differential equations in divergence form with first boundary conditions", *J. Optim. Theory Appl.* **36** (1982) 565–613.
- [15] S. Nakagiri, "Optimal control of linear retarded systems in Banach spaces", *J. Math. Anal. Appl.* **120** (1986) 169–210.
- [16] J. W. Nunziato, "On heat conduction in materials with memory", *Quart. Appl. Math.* **29** (1971) 187–204.
- [17] R. E. Showalter and N. J. Walkington, "A hyperbolic Stefan problem", *Quart. Appl. Math.* **45** (1987) 769–781.
- [18] H. Tanabe, *Functional analytic methods for partial differential equations* (Marcel Dekker, New York, 1997).
- [19] P. K. C. Wang, "Optimal control of parabolic system with boundary conditions involving time delays", *SIAM J. Control* **13** (1975) 274–293.
- [20] J. Yong and L. Pan, "Quasi-linear parabolic partial differential equations with delays in the highest-order spatial derivatives", *J. Austral. Math. Soc., Ser. A* **54** (1993) 174–203.
- [21] Y. You and E. B. Lee, "Quadratic optimization for infinite-dimensional linear differential difference type system", *SIAM J. Control Optim.* **28** (1990) 265–293.