J. Austral. Math. Soc. Ser. B 36(1994), 194–212

THE MAXIMUM PRINCIPLE FOR A TYPE OF HEREDITARY SEMILINEAR DIFFERENTIAL EQUATION

FEIYUE HE[1]

(Received 24 June 1992; revised 2 February 1993)

Abstract

An optimal control problem governed by a class of delay semilinear differential equations is studied. The existence of an optimal control is proven, and the maximum principle and approximating schemes are found. As applications, three examples are discussed.

1. Introduction

Let H, V and F be three Hilbert spaces with $V \subseteq H \subseteq V'$ algebraically and topologically, where V' is the dual space of V. Let A, B_1 and B_2 be the three operators defined by $A : V \to H$ and $B_1, B_2 : F \to H$ respectively. Let T and σ be two given positive numbers. For any function $f(t)$, we denote by f_σ the function defined by $f_\sigma(t) = f(t - \sigma)$.

Now we introduce the control set U to be $L^2([0, T], F)$ and for any $u \in U$, we consider the semilinear differential equation:

$$\begin{aligned} y_t(t) + Ay(t) + \beta(y(t), y_\sigma(t)) &= B_1u(t) + B_2u_\sigma(t) \text{ for } t \in [0, T] \\ y(s) &= y_0(s) \in H \end{aligned} \tag{1}$$

in H, where $y_t(t) = \frac{d}{dt}y(t)$ and β is going to be defined later. σ denotes the length of the time delay. The state at time t of the delay equation (1) is given by the function segment y_t. Therefore, $y(s) = y_0(s)$ for $s \in [-\sigma, 0]$ is a condition for the early state y_0.

We define the cost functional by

$$J(u, y) = \int_0^T g(t, y(t))dt + \int_\sigma^T h(u(t))dt + \phi_0(y(T)) \tag{2}$$

[1]Department of Mathematics, Davis and Elkins College, Elkins, WV 26241.

for $B_2 \neq 0$. If $B_2 = 0$ then we define the cost functional by

$$J(u, y) = \int_0^T g(t, y(t))dt + \int_0^T h(u(t))dt + \phi_0(y(T)), \tag{3}$$

where y is a solution of (1) corresponding to the control u, and g, h, ϕ_0 are functions defined by $g : [0, T] \times H \to R$, $h : F \to \overline{R}$, $\phi_0 : H \to R$, respectively.

We consider the following optimization problem:

$$Minimize \;\; J(u, y) \;\; \text{on all} \;\; u \in U \tag{4}$$

This paper contains five sections as follows. In Section 2, the assumptions are formulated and the existence theorem for the state equation (1) will be given. In Section 3, we will establish the existence of an optimal control for the optimization problem (4). We will discuss an approximating control process and therefore obtain the maximum principle for the optimum control problem of (1) combined with (4) in Section 4. Finally, in Section 5, as applications, three examples have been studied: optimal control problems governed by parabolic equations, semilinear hyperbolic equations and *Navier-Stokes* equations respectively. For those three examples, we have been able to improve on previous results. The optimal control problem governed by the same kind of semilinear equations without delays has been discussed in [8]. Delay optimal control problems have been studied in a series of papers: Banks and Jacobs [1], Barbu [3], Barbu and Precupanu [4], Colonius [5] and Fattorini [7].

2. Some assumptions and existence of the state equation

The following hypotheses will be in effect throughout the rest of the paper:

H1: The operator A is selfadjoint, continuous and positive:

$$(Ay, y)_H \geq 0 \;\; \text{for any} \;\; y \in D(A)$$

and $0 \in D(A)$, where $(\cdot, \cdot)_H$ denotes the scalar product in H and $D(A)$ is the domain of the operator A. Moreover the semigroup $S(t) = \exp(-tA)$ generated by $-A$ is compact for $t > 0$.

H1 implies that the spectrum $\sigma \subset [0, \infty)$ and $S(t)$ is an analytic semigroup. Under this hypothesis, the fractional power A^α is bounded for $\alpha < 0$. For $\alpha \geq 0$, we denote by H_α the space $D(A^\alpha)$, this space equipped with its natural inner product $(y, z)_\alpha = (A^\alpha y, A^\alpha z)_H$. The inner product corresponds to the norm $\|y\| = \|A^\alpha y\|_H$. For $\alpha < 0$, H^α is the closure of H under the norm $\|\cdot\|_\alpha$; for details see [2] or [8].

The nonlinear term $\beta(\cdot, \cdot)$ is defined in $H_{\frac{1}{2}} \times H_{\frac{1}{2}}$ and satisfies the following hypothesis ($L(X, Y)$ denotes the space of all bounded linear operators from the Banach space X into the Banach space Y endowed with its uniform operator norm):

H2: There exists $\theta, 0 \leq \theta < \frac{1}{2}$, such that the map $\beta(\cdot, f) : H_{\frac{1}{2}} \to H_{-\theta}$ is continuous and locally bounded for any $f \in H_{\frac{1}{2}}$. Furthermore β has *Frechet* derivatives for each of the variables $D_1\beta(y, z)$ and $D_2\beta(y, z)$, which are continuous and locally bounded (as an element of $L(H_{\frac{1}{2}} \times H_{\frac{1}{2}}, H_{-\theta})$-valued function).

The control $u(t)$ takes a value in F. The linear operators B_1 and B_2 satisfy:

H3: $B_i \in L(F, H)$ for $i = 1, 2$.

We will assume that control u belongs to $L^2([-\sigma, T]; F)$ (if $B_2 = 0$, then $u \in L^2([0, T], F)$). As usual, we define a solution of the delay problem (1) in an interval $0 \leq t \leq T$ as a $H_{\frac{1}{2}}$-valued function $y(t)$ which is continuous in the norm of $H_{\frac{1}{2}}$ and satisfies:

$$\begin{aligned} y(t) \quad = \quad & S(t)y_0(0) - \textstyle\int_0^t A^{\theta} s(t-r)A^{-\theta}\beta(y(r), y_\sigma(r))dr \\ & + \textstyle\int_0^t S(t-r)(B_1u(r) + b_2u_\sigma(r))dr \qquad \text{for } t \in [0, T]. \end{aligned} \tag{5}$$

LEMMA 1. *If H_1 to H_3 hold and if $y_0 \in C((-\sigma, 0], H_{\frac{1}{2}})$, then for any $u(\cdot) \in U$, (1) has an unique solution y in the space $C([O, T]; H_{\frac{1}{2}}) \cap L^2([0, T]; D(A))$ and $y_t \in L^2([0, T]; H)$.*

PROOF. First we prove this lemma holds for $T \leq \sigma$. In this case, (5) can be written as:

$$\begin{aligned} y_1(t) = S(t)y_0(0) - \int_0^t A^{\theta} S(t-r)A^{-\theta}\beta(y_1(r), y_0(r-\sigma))dr \\ + \int_0^t S(t-r)(B_1u(r) + B_2(u_\sigma(r))dr. \end{aligned}$$

Now with hypotheses H_1, H_2 and $N(y) = \beta(y, y_0(r-\sigma))$, all hypotheses of Theorem 1 of [10] are satisfied. Hence there exists a unique solution y_1 of (5) such that $y_1 \in C\left([0, T]; H_{\frac{1}{2}}\right) \cap L^2([0, T]; D(A))$ and $y_t \in L^2([0, T]; H))$. For $T > \sigma$, by the above result, there exists a unique solution $y_1(t)$ in $[0, \sigma]$. We extend $y_1(t)$ into the interval $[-\sigma, 0]$ by defining $y_1(t) = y_0(t)$ for $t \in [-\sigma, 0]$. Then using the same argument as above, with y_0 and the interval $[-\sigma, 0]$ replaced by y_1 and the interval $[-\sigma, \sigma]$ respectively, we imply that the corresponding equation to (5),

$$\begin{aligned} y_2(t) = S(t-\sigma)y_1(\sigma) - \int_\sigma^t A^{\sigma} S(t-r)A^{-\sigma}\beta(y_2(r), y_1(r-\sigma))dr \\ + \int_\sigma^t S(t-r)(B_1u(r) + b_2u_\sigma(r))dr, \end{aligned} \tag{6}$$

has a unique solution $y_2(t) \in C([\sigma, 3\sigma]; H_{\frac{1}{2}}) \cap L^2([\sigma, 3\sigma]; D(A))$ and $y_t \in L^2([\sigma, 3\sigma]; H)$. Moreover, by the definition of y_1, we have

https://doi.org/10.1017/S0334270000010353 Published online by Cambridge University Press

$$
\begin{aligned}
S(t-\sigma)y_1(\sigma)& \\
&= S(t-\sigma)\Big[S(\sigma)y_0(0) - \int_0^\sigma S(\sigma-r)A^{-\theta}\beta(y_1(r), y_0(r-\sigma))dr \\
&\qquad + \int_0^\sigma S(\sigma-r)(B_1u(r) - B_2u_\sigma(r))dr\Big] \\
&= S(t)y_0(0) - \int_0^\sigma A^\theta S(t-r)A^{-\theta}\beta(y_1(r), y_0(r-\sigma))dr \\
&\qquad + \int_0^\sigma S(t-r)(B_1u(r) + B_2u_\sigma(r))dr. \qquad (7)
\end{aligned}
$$

In (6), substitute the right hand side of (7) for $S(t-\sigma)y_1(\sigma)$; we obtain

$$
\begin{aligned}
y_2(t) = S(t)y_0(0) &- \int_0^t A^\theta S(t-r)A^{-\theta}\beta(y_2(r), y_0(r-\sigma))dr \\
&+ \int_0^t S(t-r)(B_1(r) + B_2u_\sigma(r))dr,
\end{aligned}
$$

that is, y_2 is the unique solution of (5) in $[0, 3\sigma]$.

Using the same argument as the proof of the existence of y_2, step by step, we can finally conclude this lemma for any $T > 0$.

3. The existence of an optimal control

Let W be a convex closed subset of F. We denote by $M([-\sigma, T]; W)$ the space of all strongly measurable functions $u(\cdot)$ such that $u(t) \in W$ *a.e.* in $[-\sigma, T]$ and $u \in U$. In order to state and prove the existence theorem for the optimal control problem (4), we need make some assumptions for the functions g, h, ϕ_0:

H4: The function $h : U \to \overline{R}$ is convex, lower semicontinuous (*l.s.c.*) and there exist $C_1, C_2 \in R$ such that

$$h(u) \geq C_1|u|_W^2 - C_2 \qquad \forall u \in M([-\sigma, T]; W).$$

H5: The function $g : [0, T] \times H \to R^+$ is measurable in t, and for every $r > 0$ there exists an $L_r > 0$ independent of t such that $g(t, 0) \in L^\infty(0, T)$ and

$$|g(t, y) - g(t, z)| + |\phi_0(y) - \phi_0(z)| \leq L_r|y - z|_2$$

for all $t \in [0, T]$, $|y|_2 + |z|_2 \leq r$.

A sequence $\{u_n(\cdot)\} \subset M([-\sigma, T]; W)$ of some controls is called a minimizing sequence if

$$J(u_n, y_n) \to \inf_{u \in M} J(u, y) = J^*.$$

https://doi.org/10.1017/S0334270000010353 Published online by Cambridge University Press

Here y_n is the solution of (1) corresponding to u_n for each n. We have the following existence theorem:

THEOREM 2. *Assume that the assumptions* H1 *to* H5 *hold, W is convex and there exists a minimizing sequence $\{u_n(\cdot)\} \subset M([-\sigma, T]; W)$ such that the corresponding $\{y_n(t)\}$ is uniformly bounded in the norm of $C([0, T]; H_{\frac{1}{2}})$. Then there exists an optimal control $\hat{u}(\cdot)$ as the weak L^2-limit of a subsequence of $\{u_n(\cdot)\}$.*

REMARK. Furthermore, if we assume that β satisfies

H6: $|\langle\beta(y, \zeta), y\rangle| \leq C\|y\|_H^2$, for any $y \in H_{\frac{1}{2}}$,

(here $\zeta \in H_{\frac{1}{2}}$ and C depends on ζ), then the seqence $\{y_n\}$ is uniformly bounded in the norm of $C([0, T]; H_{\frac{1}{2}})$. In fact, taking duality pairing with y_n on both sides of (1) with u replaced by u_n (denote the corresponding y by y_n), then in the interval $[0, \sigma]$, we get

$$\frac{1}{2}\frac{d}{dt}\|y_n\|_H^2 + \|A^{\frac{1}{2}}y_n\|_H^2 + \langle\beta(y_n, y_\sigma), y_n\rangle = (B_1u_n + B_2u_{n\sigma}, y_n)_H.$$

Then by using H1, H3, H6 and Gronwall's inequality, we can easily imply that $\{y_n\}$ is uniformly bounded in the norm of $C([0, \sigma]; H_{\frac{1}{2}})$. Hence the remark follows step by step.

PROOF OF THE THEOREM. First the existence of $\{y_n\}$ has been proven by Lemma 1. The uniformly bounded assumptions of $\{y_n\}$ and H5 imply that $J^* < \infty$. Now in the interval $[0, \sigma]$, by the definition of solutions, we have

$$y_n(t) = S(t)y_0(0) - \int_0^t A^\theta S(t-r)A^{-\theta}\beta(y_n(r), y_0(r-\sigma))dr + \int_0^t S(t-r)(B_1u_n(r) + B_2u_{n\sigma}(r))dr.$$

Using the same arguments as in the proof of Theorem 4.11 of [8], we conclude that there exist two subsequences $\{y_n\}$ and $\{u_n\}$ such that

$$u_n \to \hat{u} \text{ weakly in } L^2([-\sigma, T]; W)$$
$$y_n \to \hat{y} \text{ in } C([0, \sigma], H).$$

Similarly, in the interval $[\sigma, 2\sigma]$, we have

$$y_n(t) = S(t)y_n(\sigma) - \int_\sigma^t A^\theta S(t-r)A^{-\sigma}\beta(y_n(r), y_{n\sigma}(r))dr + \int_\sigma^t S(t-r)(B_1u_n(r) + B_2u_n(r-\sigma))dr$$

or

$$
\begin{aligned}
A^{\frac{1}{2}}y_n(t) = A^{\frac{1}{2}}S(t)y_n(\sigma) &- \int_\sigma^t A^{\theta+\frac{1}{2}}S(t-r)A^{-\theta}\beta(y_n(r), \hat{y}_{n\sigma}(r))dr \\
&+ \int_\sigma^t A^{\theta+\frac{1}{2}}S(t-r)A^{-\theta}[\beta(y_n(r), \hat{y}_{n\sigma}(r) - \beta(y_n(r), y_{n\sigma}(r))]dr \\
&+ \int_\sigma^t A^{\frac{1}{2}}S(t-r)(B_1u_n(r) + B_2u_{n\sigma}(r))dr. \qquad (8)
\end{aligned}
$$

We apply Proposition 4.1(II) of [8] with $\alpha = 0 + \frac{1}{2}$ for the first integral. For the second integral, we use H2 and the fact: $y_n(r-\sigma) \to \hat{y}(r-\sigma)$ in $C([\sigma, 2\sigma]; H)$, in addition to the uniformly bounded property of $\{y_n\}$. For the third integral, we use Proposition 4.1(I) of [8] for $p = 2$ and $\alpha = \frac{1}{2}$. Finally, for $A^{\frac{1}{2}}S(t)y_n$, we use Lemma 2.1 of [8] (or see [11]), $A^{\frac{1}{2}}y_n(t)$ is convergent to an element $z(t)$ in $L^p([\sigma, 2\sigma]; H)$ for any $p < \infty$ and we may assume (again choosing a subsequence) that it is convergent almost everywhere. Moreover, by the assumption that $\{A^{\frac{1}{2}}y_n(t)\}$ is bounded, so is $z(\cdot)$. Taking limits in (8) and defining $\tilde{y} = A^{\frac{1}{2}}z(t)$, we obtain

$$
\begin{aligned}
A^{\frac{1}{2}}\tilde{y} \;=\; A^{\frac{1}{2}}S(t-\sigma)\hat{y}(\sigma) &- \int_\sigma^t A^{\theta+\frac{1}{2}}S(t-r)A^{-\sigma}\beta(y_n(r), \hat{y}_\sigma(r))dr \\
&+ \int_\sigma^t A^{\frac{1}{2}}S(t-r)(B_1\hat{u}(r) + B_2\hat{u}_\sigma(r))dr.
\end{aligned} \qquad (9)
$$

So $A^{\frac{1}{2}}\tilde{y}(t)$ is continuous and $\tilde{y}$ is the solution corresponding to $u \in [\sigma, 2\sigma]$, that is $\tilde{y}(t) = \hat{y}(t)$ for all $t \in [\sigma, 2\sigma]$. Hence, when we left multiply both sides (9) by $A^{-\frac{1}{2}}$, we get

$$
\begin{aligned}
y_n(t) = S(t-\sigma)\hat{y}(\sigma) &- \int_\sigma^t A^{\theta}S(t-r)A^{-\theta}\beta(y_n(r), \hat{y}_\sigma(r))dr \\
&+ \int_\sigma^t S(t-r)(B_1u_n(r) + B_2u_{n\sigma}(r))dr.
\end{aligned}
$$

In Proposition 4.1 of [8] with $\alpha = \theta$ and $\alpha = 0$ respectively and selecting still another subsequence, we deduce that $\{y_n\}$ is convergent in $C([\sigma, 2\sigma], H)$, that is, we obtain

$$y_n \to \hat{y} \text{ in } C([0, 2\sigma], H).$$

Then we use similar arguments iteratively in the interval $[i\sigma, (i+1)\sigma]$, for $i = 2, 3, \ldots$, step by step; we conclude that there exist subsequences u_n, y_n such that

$$u_n \to \hat{u} \text{ weakly in } L^2([-\sigma, T]; W),$$

$$y_n \to \hat{y}_n \text{ in } C([0, T]; H).$$

Therefore, by using standard arguments, we can readily imply that the theorem holds.

4. The approximating control problem and the maximum principle

Let $(\hat{y}, \hat{u}) \in C([0, T]; H_{\frac{1}{2}}) \times M([-\sigma, T]; W)$ be any optimal pair for (4). For every $\epsilon > 0$, we consider the following approximating optimal control problem:

$$\textit{Minimize} \quad J^{\epsilon}(y, u), \tag{10}$$

where $J^{\epsilon}(y, u)$ is defined by

$$J^{\epsilon}(y, u) = \int_0^T g^{\epsilon}(t, y)dt + \int_{-\sigma}^T \left(h(u) + \frac{1}{2}|u - \hat{u}|\right) dt + \phi_0^{\epsilon}(y(T)).$$

on all $y \in C([0, T]; H_{\frac{1}{2}})$, $u \in M([-\sigma, T]; W)$ subject to (1) and where $g^{\epsilon} : [0, T] \to H$ and $\phi_0^{\epsilon} : H \to R$ are defined as

$$g^{\epsilon}(t, y(t)) = \int_{R^n} g(t, P_n y - \epsilon \Lambda_n \tau)\rho_n(\tau)d\tau, \tag{11}$$

$$\phi_0^{\epsilon}(y) = \int_{R^n} \phi_0(P_n y - \epsilon \Lambda_n \tau)\rho_n(\tau)d\tau. \tag{12}$$

Here $n = [\epsilon^{-1}]$, ρ_n is a mollifier in R^n and $P_n : H \to X_n$ are given by $P_n u = \sum_1^n u_i e_i$ if $u = \sum_1^{\infty} u_i e_i$ $\Lambda_n \tau = \sum_1^n \tau_i e_i$ and $\{e_i\}_1^{\infty}$ is an orthonormal basis in H.

Using standard arguments, we easily prove the existence results for (10). Moreover, we have the following lemma.

LEMMA 3. *Let $\{y_{\epsilon}, u_{\epsilon}\}$ be a solution to the optimality problem* (10). *Then*

$$u_{\epsilon} \to \hat{u} \ \textit{strongly in} \ L^2([0, T]; W),$$

$$y_{\epsilon} \to \hat{y} \ \textit{strongly in} \ L^2([0, T]; V) \cap C([0, T]; H).$$

PROOF. By the assumption H4, we have that $\{u_{\epsilon}\}$ is bounded in $L^2([-\sigma, T]; W)$. Therefore there exists $u^* \in L^2([-\sigma, T]; W)$ such that

$$u_{\epsilon} \to u^* \ \text{weakly in} \ L^2([-\sigma, T]; W).$$

Then by the same arguments as in the proof of Theorem 2, we readily obtain that

$$y_{\epsilon} \to y^* \ \text{in} \ C([0, T]; H).$$

Since the functional $u \to \int_{-\sigma}^T h(u(t))dt$ is weakly lower semicontinuous on $L^2([0, T]; W)$, we have

$$\liminf_{\epsilon \to 0} J^{\epsilon}(y_{\epsilon}, u_{\epsilon}) \geq J^{\epsilon}(y^*, u^*) \geq J(y^*, u^*). \tag{13}$$

https://doi.org/10.1017/S0334270000010353 Published online by Cambridge University Press

On the other hand, since $u_\epsilon \to u^*$, again as the proof of Theorem 2, we have

$$\hat{y}_\epsilon \to \hat{y} \text{ in } C([0,T];H),$$

and by H5 and Proposition 2.15 of [2],

$$g^\epsilon(t,\hat{y}_\epsilon(t)) \to g(t,\hat{y}(t)) \qquad \forall t \in [0,T].$$

Then by the Lebesgue dominated convergence theorem, we obtain

$$\lim_{\epsilon\to 0}\int_0^T g^\epsilon(t,\hat{y}_\epsilon(t))dt = \int_0^T g(t,\hat{y}(t))dt.$$

Similarly

$$\lim_{\epsilon\to 0}\phi_0^*(\hat{y}_\epsilon(T)) = \phi_0(\hat{y}(T))$$

whence

$$\limsup J^\epsilon(y_\epsilon,u_\epsilon) \le J(\hat{y},\hat{u})$$

and by (13), we get:

$$\lim_{\epsilon\to 0}\int_{-\sigma}^T |u_\epsilon - \hat{u}|^2 dt = 0.$$

Hence $y^* = \hat{y}$ and $u^* = \hat{u}$, that is, the proof is complete.

Now we consider the linear backward problem:

$$\begin{aligned}(p_\epsilon)_t = Ap_\epsilon + D_1\beta(y_\epsilon,y_{\epsilon\sigma})^*p_\epsilon + D_2\beta(y_\epsilon(t+\sigma),y_\epsilon(t))^*p_\epsilon(t+\sigma)&\\ + Dg^\epsilon(t,y_\epsilon), \qquad \forall t\in[T-\sigma,T]&\end{aligned} \tag{14}$$

$$(p_\epsilon)_t = Ap_\epsilon + D_1\beta(y_\epsilon,y_{\epsilon\sigma})^*p_\epsilon + D_2g^\epsilon(t,y_\epsilon), \qquad \forall t\in[T-\sigma,T], \tag{15}$$

$$p_\epsilon(T) = -D\phi_0^\epsilon(y_\epsilon(T)), \tag{16}$$

$$p_\epsilon(s) = 0, \qquad \forall s\in[T,T+\sigma], \tag{17}$$

where $D_1\beta(y_\epsilon,y_{\epsilon\sigma})^*$ and $D_2\beta(y_\epsilon(t+\sigma),y_\epsilon)^*$ are the dual operators of $D_1\beta(y_\epsilon,y_{\epsilon\sigma})$ and $D_2\beta(y_\epsilon(t+\epsilon),y_\epsilon)$, respectively.

First we state the following Lemma, which can be found in, say, [8].

LEMMA 4. *Let $\gamma(\cdot)\in L^2([\tau,T];H)$ and A be a given linear bounded operator from H_α to $H_{-\theta}$. Then the system*

$$\begin{aligned} z_t(t) + (A+L(t))z(t) &= \gamma(t) \qquad \text{for } \tau\le t\le T,\\ z(\tau) &= z_0\end{aligned} \tag{18}$$

possesses a unique solution $z(\cdot)$ in $\tau\le t\le T$. This solution has the following properties:

https://doi.org/10.1017/S0334270000010353 Published online by Cambridge University Press

(i) $z(\cdot)$ *is an* H_α*-valued continuous function in* $\tau \le t \le T$ *and satisfies*

$$\|z(t)\|_\alpha \le C_1 \left((t-\tau)^{-\alpha}\|z_0\|_H + \|\gamma\|_{L^2([\tau,T];H)}\right) \text{ for } t \in (\tau, T]; \tag{19}$$

(ii) $z(\cdot) \in C^2([\tau, T]; H)$ *and*

$$\|z(t)\|_H \le C_2 \left(\|z_0\|_H + \|\gamma\|_{L^2([\tau,T];H)}\right) \text{ for } t \in (\tau, T]. \tag{20}$$

Moreover, $z(\cdot) \in L^2([0, T]; H_\alpha)$ *and*

$$\|z(\cdot)\|_{L^2([\tau,T];H_\alpha)} \le C_3 \left(\|z_0\|_H + \|\gamma\|_{L^2([\tau,T];H)}\right) \text{ for } t \in (\tau, T].$$

From this lemma, we easily conclude that (15)–(16) has a unique solution $\overline{p}_\epsilon \in C([T-\sigma, T]; H) \cap L^2([T-\sigma, T]; H_\alpha)$.

We now consider the backward problem (14) with initial value

$$p_\epsilon(T-\sigma) = \overline{p}_\epsilon(T-\sigma). \tag{21}$$

In the interval $[T-2\sigma, T-\sigma]$, if we let

$$\gamma(t) = D_2\beta(y_\epsilon(t+\sigma), y_\epsilon(t))^* \overline{p}_\epsilon(t+\sigma) + D_2 g^\epsilon(t, y_\epsilon),$$

then $\gamma(t) \in L^2([T-2\sigma, T-\sigma]; H)$ and $L(t) = D_1\beta(y_\epsilon(t), y_{\epsilon\sigma}(t))$ again satisfies the hypotheses of Lemma 4. Hence, in $[T-2\sigma, T-\sigma]$, (14) and (21) has a unique solution in $C([T-2\sigma, T-\sigma]; H) \cap L^2([T-2\sigma, T-\sigma]; H_\alpha)$. Therefore, step by step and using the same argument, we finally conclude that the system (14)–(17) has a unique solution in $C([0, T]; H) \cap L^2([0, T]; H_\alpha)$.

Simlarly using Lemma 4, we can obtain the following estimates:

$$\|p_\epsilon(t)\|_H \le C \quad \text{for } t \in (0, T] \tag{22}$$

and

$$\|p_\epsilon\|_{L^2([0,T];H_\alpha)} \le C, \tag{23}$$

where C is a constant independent of ϵ. Then multiplying (14) and (15) by p'_ϵ and using (22) and (23), we obtain

$$\left\|\frac{d}{dt} p_\epsilon\right\|^2_{L^2([0,t];H)} \le C. \tag{24}$$

On the other hand, since (y_ϵ, u_ϵ) is an optimal control for (10), we have

$$J^\epsilon(y_\epsilon^{u+\lambda v}, u_\epsilon + \lambda v) \ge J^\epsilon(y_\epsilon, u_\epsilon) \qquad \forall\lambda > 0, \quad v \in L^2([-\sigma, T]; W). \tag{25}$$

Here $y_\epsilon^{u+\lambda v}$ is the solution of (1) corresponding to $u_\epsilon + \lambda v$. Then (25) yields

$$\int_0^T (D_2 g^\epsilon(t, y_\epsilon), z_\epsilon)dt + \int_{-\sigma}^T [(h'(u_\epsilon), v) + \langle u_\epsilon - \hat{u}, v\rangle]dt + (\nabla\phi_0^\epsilon(y(T)), z_\epsilon) \geq 0 \qquad \forall v \in L^2([-\sigma, T]; W), \tag{26}$$

where $z_\epsilon \in C([0, T]; H) \cap L^2([0, T]; H_\alpha)$ is the solution to the linear equation

$$\begin{aligned} z_{\epsilon t} + Az_\epsilon + D_1\beta(y_\epsilon, y_{\epsilon\sigma})^* z_\epsilon + D_2\beta(y_\epsilon, y_{\epsilon\sigma})^* z_{\epsilon\sigma} = B_1 v + B_2 v_\sigma, \qquad t \in [0, T] \\ z_\epsilon = 0, \qquad s \in [-\sigma, 0] \end{aligned} \tag{27}$$

This delay problem can be solved in $C([0, T]; H) \cap L^2([0, T]; H_\sigma)$ by the same methods.

If we multiply (14) and (15) by z_ϵ, integrate and then use (27), we can obtain

$$\int_0^T Dh(u_\epsilon, v)dt + \int_{-\sigma}^T \langle u_\epsilon - \hat{u}, v\rangle dt + \int_0^T (p_\epsilon, B_1 v)dt + \int_{-\sigma}^{T-\sigma} (p_\epsilon, B_2 v_\epsilon \sigma)dt \geq 0,$$

for all $v \in L([0, T]; W)$, where $Dh(u_\epsilon, v)$ is the derivative of h at u_ϵ in the direction v. This yields

$$B^* p_\epsilon(t) = Dh(u(t)) + u_\epsilon(t) - \hat{u}(t), \qquad t \in [0, T], \tag{28}$$

where B^* is the operator defined by $B^* p_\epsilon(t) = B_1^* p_\epsilon(t) + B_2^* p_\epsilon(t + \sigma)$. Equations (14)–(17), (28) together represent the *Euler-Lagrange* optimality condition for (10).

Now we will prove that there exists a subsequence $\{p_{\epsilon_n}\}$ such that

$$p_{\epsilon_n} \to \hat{p} \ \text{ in } \ C([0, T]; H) \cap L^2([0, T]; H_\alpha). \tag{29}$$

In the interval $[T - \sigma, T]$ we have

$$\begin{aligned} p_\epsilon = -S(t - T)D\phi_0^\epsilon(y_\epsilon(T)) &+ \int_{T-\sigma}^t A^\theta S(t - r)A^{-\theta} D_2\beta(y_\epsilon, y_{\epsilon\sigma})^* p_\epsilon(r)dr \\ &+ \int_{T-\sigma}^t S(t - r)D_2 g^\epsilon(t, y_\epsilon(t))dr. \end{aligned}$$

By Proposition 2.15, Ch. 2 and Lemma 1.4, Ch. 5 of [2], we have that

$$\begin{aligned} -D\phi_0^\epsilon(y(T)) \to p(T) \in \partial\phi_0(\hat{y}(T)) \ \text{ weakly in } \ H, \\ D_2 g\epsilon(t, y_\epsilon) \to \xi \in D_2 g(t, \hat{y}(t)) \ \text{ weakly star in } \ L^\infty([0, T]; H). \end{aligned} \tag{30}$$

Then, of course, $D_2 g^\epsilon(t, y_\epsilon) \to \xi$ weakly in L^2. Moreover, we have that $\|D_2\beta(y_\epsilon, y_{\epsilon\sigma})^* p_\epsilon(r)\|_{\frac{1}{2}}$ is uniformly bounded. Therefore, by the same arguments as in the proof of Theorem 2, we conclude that for some subsequence, denoted again by $\{p_\epsilon\}$, we have

$$p_\epsilon \to \hat{p} \text{ in } C([T-\sigma, T]; H) \cap L^2([T-\sigma, T]; H_\alpha). \tag{31}$$

In the interval $[T-2\sigma, T-\sigma]$, we have

$$\begin{aligned} p_\epsilon(t) = S(t-T+\sigma)p_\epsilon(T-\sigma) + \int_{T-2\sigma}^{t} A^\theta S(t-r)A^{-\theta} D_1\beta(y_\epsilon, y_{\epsilon\sigma})p_\epsilon(r)\,dr \\ + \int_{T-2\sigma}^{t} S(t-r)\big[D_2 g^\epsilon(r, y_\epsilon) + D_2\beta(y_\epsilon(r+\sigma), y_\epsilon(r))\big]p_\epsilon(r+\sigma)\,dr. \end{aligned}$$

Then by (31) and the facts $D_1\beta$ and $D_2\beta$ are continuous and bounded, the same argument implies that for some subsequence of $\{p_\epsilon\}$ and some function in $C([T-2\sigma, T-\sigma]; H) \cap L^2([T-2\sigma, T-\sigma]; H_\alpha)$, again denoted by $\{p_\epsilon\}$ and $\hat{p}$, we have

$$p_\epsilon \to \hat{p} \text{ in } C([T-2\sigma, T-\sigma]; H) \cap L^2([T-2\sigma, T-\sigma]; H_\alpha).$$

By the same arguments, we can then prove the above convergence holds in the interval $[T-3\sigma, T-2\sigma]$. Iteratively, we can finally conclude (29). Moreover, by (24), we have that $(p_\epsilon)_t \to \hat{p}$ weakly in $L^2([0, T]; H)$.

Letting $\epsilon \to 0$ in the system (14) to (17), we conclude that there is a $p \in C([0, T]; H) \cap L^2([0, T]; H_\alpha)$ such that p satisfies the following equations:

$$\begin{aligned} -p_t(t) + Ap(t) + D\beta(\hat{y}(t), \hat{y}_\sigma(t))^* p(t) + D_2\beta(\hat{y}(t+\sigma), \hat{y}(t))^* p(t+\sigma) \\ \in D_2 g(t, \hat{y}(t)), \qquad t \in [0, T-\sigma], \end{aligned} \tag{32}$$

$$\begin{aligned} -p_t(t) + Ap(t) + D_1\beta(\hat{y}(t), \hat{y}_\sigma(t))^* p(t) \in D_2 g(t, \hat{y}(t)), \\ t \in [T-\sigma, T], \end{aligned} \tag{33}$$

$$p(T) \in -D\phi_0(\hat{y}(T)), \tag{34}$$

$$p(s) = 0, \qquad s \in [T, T+\sigma]. \tag{35}$$

Moreover, since the map $Dh : W \to W$ is closed, taking $\epsilon \to 0$ in (28) we obtain

$$B^* p(t) \in Dh(\hat{u}(t)), \qquad a.e.\ in\ [0, T]. \tag{36}$$

Therefore, we have proved the following weak form of the maximum principle for (10).

THEOREM 5. *Let $(\hat{y}, \hat{u})$ be any optimal pair for (10). Then there exists a function $p \in W^{1,2}([0, T]; H) \cap L^2([0, T]; H_\alpha)$ satisfying* (33)–(36).

5. Some examples

1: *Optimal control of delay semilinear parabolic equations*

Let Ω be an open set of R^n and $G = [0, T] \times \Omega$. Here, we shall study (1) and (4) in the special case where $A = -\Delta$, $V = -H_0^1(\Omega)$, $H = L^2(\Omega)$, $V' = H^{-1}(\Omega)$ and $\beta(y, y_\sigma) = \beta(y_\sigma)$ is locally *Lipschitz* continuous and satisfies the hypothesis

H2′:

$$|\beta'(y)| \le C(|\beta(y)| + |y| + 1).$$

In other words, we consider the case where the state equation is given by

$$\begin{aligned} y_t - \Delta y + \beta(y_\sigma) &= B_1 u \text{ in } G, \\ y(s, x) &= Y_0(s, x)(zs, x) \in [-\sigma, T] \times \Omega, \\ y(t, x) &= 0 \text{ on } \Sigma = (0, T) \times \partial\Omega \end{aligned} \tag{37}$$

for $y_0 \in L^2([-\sigma, 0]; H_0^1(\Omega))$, $B_1 \in L(W, L^2(\Omega))$.

THEOREM 6. *The optimal control problem* (4) *with the state system* (37) *has at least one solution* $(\hat{y}, \hat{u}) \in W^{1,2}([0, T]; H) \cap L^2([0, T]; H_0^1(\Omega) \cap H^2) \times L^2([0, T]; W)$. *Moreover, for every such optimal pair, there is a function* $p \in L^2([0, T]; H_0^1(\Omega)) \cap W^{1,2}([0, T]; H)$ *and* $\mu \in L^1(G)$ *such that*

$$p_t(t) + \Delta p(t) - \mu \in D_2 g(t, \hat{y}(t)), \qquad t \in [0, T - \sigma], \tag{38}$$

$$p_t(t) + \Delta p(t) \in D_2 g(t, \hat{y}(t)), \qquad t \in [0, T - \sigma], \tag{39}$$

$$p(T) \in -D\phi_0(\hat{y}(T)), \qquad \text{in } \Omega, \tag{40}$$

$$B^* p(t) \in Dh(\hat{u}(t)), \quad \mu(t) \in \partial\beta(\hat{y}(t)) p(t + \sigma), \qquad a.e. \text{ in } [0, T], \tag{41}$$

$$p(s) = 0, \qquad s \in [T, T + \sigma]. \tag{42}$$

Here $\partial\beta$ *is the subgradient of* β *in the sense of Clarke.*

PROOF. First we shall prove that for every $y_0 \in L^2([-\sigma, 0]; H_0^1(\Omega))$, $\mu \in L^2([0, T]; W)$, (37) has a unique solution in $W^{1,2}([0, T]; H) \cap L^2([0, T]; H_0^1(\Omega) \cap H^2)$. Indeed, in the interval $[0, \sigma]$, (37) can be written as the following:

$$\begin{aligned} &y_t - \Delta = Bu - \beta(y_{0\sigma}), \quad \text{in } [0, T] \times \Omega, \\ &y(0, x) = y_0(0, x), \quad \text{for all } x \in \Omega, \\ &y(t, x) = 0, \qquad (t, x) \in (0, \sigma) \times \partial\Omega. \end{aligned}$$

By [4], we obtain that there exists a unique solution $y \in W^{1,2}([0, \sigma]; V \cap H^2(\Omega))$. Then by using a step by step argument as before, we can easily conclude that (37) has a unique solution y in $W^{1,2}([0,T]; V \cap H^2(\Omega))$. The existence of an optimal pair $(\hat{y}, \hat{u})$ can be proved by standard methods. Here we omit this proof.

To prove the necessity of optimality system, we consider the approximation problem (10) with the state equation

$$\begin{aligned} y_t - \Delta y + \beta^\epsilon(y_\sigma) &= Bu, \qquad \text{in } G, \\ y(s,x) &= y_0(s,x), \qquad s \in [-\sigma, 0], \\ y &= 0, \qquad (t,x) \in \Sigma. \end{aligned} \tag{43}$$

For this smooth problem, we conclude that there exists a function $p_\epsilon \in L^2([0,T]; H_0^1(\Omega)) \cap C([0,T]; H)$ with $p'_\epsilon \in L^2([0,T]; L^2(\Omega))$ (see Theorem 1.9, Ch. IV of [2]) such that

$$(p_\epsilon)_t + \Delta p_\epsilon - D\beta(y_\epsilon)p_\epsilon(t+\sigma) = D_2 g^\epsilon(t, y_\epsilon), \quad (t,x) \in [0, T-\sigma] \times \Omega, \tag{44}$$

$$(p_\epsilon)_t + \Delta p_\epsilon = D_2 g^\epsilon(t, y_\epsilon), \quad (t,x) \in [T-\sigma, T] \times \Omega, \tag{45}$$

$$p_\epsilon(T) = -\partial\phi_0^\epsilon(y_\epsilon(T)), \quad \text{in } \Omega, \tag{46}$$

$$p_\epsilon(s) = 0, \quad (t,x) \in [T, T+\sigma] \times \Omega. \tag{47}$$

Moreover, $p_\epsilon = 0$ on Σ and

$$\tfrac{1}{2}(u_\epsilon - \hat{u}) + B^* p_\epsilon \in Dh(u_\epsilon).$$

For (45)–(46), using the estimates of parabolic equation, we obtain

$$\|p_\epsilon\|_{W^{1,2}([T-\sigma,T]; H^2)} \le C\left(\|\nabla g(t, y_\epsilon)\|_{L^\infty([0,T]; H)} + |\nabla\phi_0|_2\right).$$

Then by assumptions H4 and H5, we conclude that

$$\|p_\epsilon\|_{W^{1,2}([T-\sigma,T]; H^2)} \le C.$$

Moreover, by using the assumption H2′ and the step by step argument, we have that

$$\|p_\epsilon\|_{W^{1,2}([T-\sigma,T]; H^2)} \le C \quad \text{and} \quad |\partial\beta^\epsilon(y_\epsilon(t))p_\epsilon(t+\sigma)| \le C.$$

Hence, we conclude that

$$p_{\epsilon_n}(t) \to p(t), \ \text{ weakly in } H, \tag{48}$$

$$p_{\epsilon_n} \to p, \ \text{weakly star in } L^\infty([0,T]; H) \text{ and strongly in } L^2([0,T]; H), \tag{49}$$

$$\partial\beta^{\epsilon_n}(y_{\epsilon_n})p_{\epsilon_n}(t+\sigma) \to \mu \quad \text{in } L^1(G). \tag{50}$$

Therefore, we can take the limit and conclude Theorem 4.

2: *Optimal control of nonlinear delay hyperbolic equation*

We study (4) in the case when the state equation is nonlinear hyperbolic and where $V = H_0^1(\Omega)$, $V' = H^{-1}(\Omega)$, $H = L^2(\Omega)$, $A = -\Delta$ and $D(A) = A \cap H^2(\Omega)$, that is, we study the state equation

$$\begin{aligned} y_{tt} - \Delta y + \beta(y_\sigma) &= Bu, \qquad \text{in } G, \\ y(s,x) = y_0(s,x) \text{ and } y_t(s,x) &= y_{0t}(s,x), \qquad (s,x) \in [-\sigma, 0] \times \Omega, \\ y &= 0, \qquad \text{on } \Sigma. \end{aligned} \tag{51}$$

Here $y_0 \in L^2([-\sigma, 0]; V)$, $B \in L^2(W, H)$. Equation (51) can be written

$$\begin{aligned} y_t &= z \qquad \text{in } G, \\ z_t - \Delta y + \beta(y_\sigma) &= Bu, \qquad \text{in } G, \\ y(s,x) = y_0(s,x) \text{ and } z(s,x) &= z_0(s,x), \qquad (s,x) \in [-\sigma, 0] \times \Omega, \\ y &= 0, \qquad \text{on } \Sigma. \end{aligned} \tag{52}$$

Now let X denote the product space $V \times H$ endowed with the scalar product

$$\langle (y,z),(y_1,z_1)\rangle = (\Delta^{\frac{1}{2}} y, \Delta^{\frac{1}{2}} y_1) + (z, z_1).$$

Let

$$A = \begin{pmatrix} 0 & -1 \\ -\Delta & 0 \end{pmatrix} \quad \text{and} \quad D(A) = D(-\Delta) \times V;$$

then A is self adjoint, continuous and positive for all (y, z) in X. So (52) can be written as

$$\begin{aligned} \vec{y}_t + A\vec{y} + \vec{\beta}(\vec{y}_\sigma) &= \vec{B}u, \\ \vec{y}(s) &= \vec{y}_1(s) = (y_0(s), y_{0t}(s)), \qquad s \in [-\sigma, 0], \\ \vec{y} &= 0, \quad \text{on } \Sigma, \end{aligned} \tag{53}$$

on X, Where $\vec{y} = (y, z)$, $\vec{\beta} = (0, \beta)$ and $\vec{B}u = (0, Bu)$. We can easily check that (4) with (53) satisfies all hypotheses H1–H5. Moreover, since $\vec{\beta}$ does not depend on the first variable y, other than Theorem 3 in Section 3, we have the following result:

THEOREM 7. *The optimal control problem* (4), *having* (51) *as the state equation with β locally Lipschitz continuous and satisfying hypothesis* H2′, *has at least one solution (y, u) with $y \in C^1([0, T]; H) \cap C([0, Tl; V \cap H^2)$, $\hat{y}_t \in C^1([0, T]; H) \cap C([0, T]; V)$, $\hat{y}_{tt}, \Delta\hat{y} \in L^\infty([0, T]; H)$, and $u \in L^2([0, T]; W)$. Furthermore, for every such optimal pair, there exist $p \in W^{1,2}([0, T]; H)$ and $\mu \in L^1(G)$ such that*

$$p_{tt} - \Delta p + \mu \in D_2 g(t, \hat{y}), \qquad t \in [0, T-\sigma], \tag{54}$$

$$p_{tt} - \Delta p \in D_2 g(t, \hat{y}), \qquad t \in [0, T-\sigma], \tag{55}$$

$$p(T) = 0, \quad p_t(T) \in -D\phi_0(\hat{y}(T)) \text{ and } p = 0 \quad \text{on } \Sigma, \tag{56}$$

$$B^* p + Dh(\hat{u}) = 0 \text{ and } \mu(t) \in \partial\beta(y(t))p(t+\sigma), \qquad \text{a.e. } t \in [0, T]. \tag{57}$$

PROOF. For $u \in L^2([0,T]; W)$, the existence of a solution of (53) can be easily proved by using the existence results for wave equations (see Ch. 2 of [12]). Moreover, by using standard arguments, we can conclude the existence of an optimal pair $(\hat{y}, \hat{u})$.

We now prove the existence of p, by using the approximation method. Similarly as in (14)–(17), we consider the ϵ-*approximating* delay problem:

$$(p_\epsilon)_{tt} - \Delta p_\epsilon + D_2\beta^\epsilon(y_\epsilon)p_\epsilon(t+\sigma) = D_2 g^\epsilon(t, y_\epsilon), \qquad t \in [0, T-\sigma], \tag{58}$$

$$(p_\epsilon)_{tt} - \Delta p_\epsilon = D_2 g^\epsilon(t, y_\epsilon), \qquad t \in [T-\sigma, T], \tag{59}$$

$$(p_\epsilon)_t(T) = -D\phi_0^\epsilon(y_\epsilon(T)) \text{ and } p_\epsilon(T) = 0, \tag{60}$$

$$p_\epsilon(s) = p_{\epsilon t}(s) = 0, \qquad s \in [T, T+\sigma]. \tag{61}$$

$$p_\epsilon = 0, \quad \text{on} \quad \Sigma,$$

where

$$\begin{aligned} (y_\epsilon)_{tt} - \Delta y_\epsilon + \beta_\epsilon(y_{\epsilon\sigma}) &= Bu_\epsilon, \qquad \text{in } G, \\ y &= 0, \qquad \text{on } \Sigma, \\ y_\epsilon(s) = y_0(s) \text{ and } y_{\epsilon t}(s) &= y_{0t}(s), \qquad s \in [-\sigma, 0]. \end{aligned} \tag{62}$$

First we note, as before, that $u_\epsilon \to \hat{u}$ strongly in $L^2([0,T]; W)$. Then we consider (62) in the interval $[0, \sigma]$. Multiplying the first equation of (62) by $(y_\epsilon)_t$ and integrating over $[0, t]$, we get

$$\frac{1}{2}|y_t(t)|_2^2 + \frac{1}{2}|\nabla y_\epsilon(t)|_2^2 - \frac{1}{2}\left[|y_0(0)|_2^2 + |\nabla y_0(0)|_2^2\right]$$
$$\leq \int_0^t |Bu_\epsilon(s) - \beta^\epsilon(y_{\epsilon\sigma}(s))|_2^2\, |y_{\epsilon t}|_2^2\, ds.$$

By Gronwall's lemma, we obtain

$$|y_{\epsilon t}(t)|_2^2 + |\nabla y_\epsilon(t)|_2^2 \leq C, \qquad \forall \epsilon > 0,\ t \in [0, t]. \tag{63}$$

Next we multiply (62) by $\Delta(y_\epsilon)_t$ and integrate over $[0, T]$ to get

$$|\nabla(y_\epsilon)_t|_2^2 + |\Delta y_\epsilon(t)|_2^2 \leq C. \tag{64}$$

Finally, multiply the same equation by $(y_\epsilon)_{tt}$ and integrate over $[0, t]$; we obtain

$$\int_0^t |(y_\epsilon)_{tt}|\, ds \leq C. \tag{65}$$

Hence the Arzela-Ascoli criterion gives

$$\begin{gathered} y_\epsilon \to \hat{y} \ \text{ strongly in } \ C([0,\sigma], V) \ \text{ and } \ \text{weakly in } \ C([0,\sigma], H^2(\Omega)) \\ (y_\epsilon)_t \to \hat{y}_t \ \text{ strongly in } \ C([0,\sigma]; H), \\ (y_\epsilon)_{tt} \to \hat{y}_{tt} \text{weakly in} L^2([0,\sigma] \times \Omega). \end{gathered} \tag{66}$$

https://doi.org/10.1017/S0334270000010353 Published online by Cambridge University Press

Now considering y_ϵ in the interval $[\sigma, 2\sigma]$ and using the same argument as above and the fact that (67) holds in $[0, \sigma]$, we conclude (67) also holds in the interval $[\sigma, 2\sigma]$. If we continue doing that, we can finally conclude that

$$\begin{aligned} y_\epsilon \to \hat{y} \ \text{ strongly in } \ C([0,T], V) \ \text{ and } \ \text{weakly in } \ C([0,T], H^2(\Omega)) \\ (y_\epsilon)_t \to \hat{y}_t \ \text{ strongly in } \ C([0,T]; H), \\ (y_\epsilon)_{tt} \to \hat{y}_{tt} \ \text{ weakly in } \ L^2(G). \end{aligned} \tag{67}$$

Similarly we multiply (58) by $(p_\epsilon)_t$ and integrate over $[t, T]$ to get

$$|\nabla p_\epsilon(t)|_2^2 + |p_{\epsilon t}(t)|_2^2 \leq |D\phi_0^\epsilon(y_\epsilon(T))|_2^2 + \int_t^T |D_2 g^\epsilon(t, y)|_2^2 |p_{\epsilon t}(t)|_2^2 dt.$$

This yields

$$|p_\epsilon(t)|_2^2 + |(p_\epsilon)_t(t)|_2^2 \leq C, \qquad \forall t \in [T-\sigma, T] \ \text{ and } \ \epsilon > 0. \tag{68}$$

By hypotheses H4 and H5, we have

$$\|\nabla g^\epsilon(t, y_\epsilon)\|_{L^\theta([0,T];H)} + |D\phi_0^\epsilon(y_\epsilon(T))|_2^2 \leq C.$$

Finally since $\{D\beta_\epsilon(y_\epsilon)\}$ is bounded and $y_\epsilon \to \hat{y}$ in $C([0,T]; V)$, using (57) and (68) in the interval $[T-2\sigma, T-\sigma]$, we obtain that (68) also holds in the interval $[T-2\sigma, T-\sigma]$. Step by step, we conclude that (68) holds for all $t \in [0, T]$. Hence we have that

$$p_\epsilon \to p \ \text{ strongly in } \ C([0,T]; H) \ \text{ and weakly in } \ L^\infty([0,T]; V).$$

Moreover,

$$p_\epsilon \to p \ \text{ weakly in } \ W^{1,2}([0,T]; H),$$

which completes the proof.

3: *Optimal control of Navier-Stokes equation*

Let $\Omega \subset R^n$, $n = 2, 3$, be an arbitrary bounded open set with a C^2 boundary $\partial\Omega$. Let $(\vec{w}, p)$ satisfy the Navier- Stokes equation:

$$\begin{aligned} \vec{w}_t + (\vec{w} \cdot \nabla)\vec{w} &= -\nabla p + \eta \Delta \vec{w} \qquad \text{in } G, \\ div\, \vec{w} &= 0, \qquad \text{in } G, \\ \vec{w}(t,x) &= v(t)\vec{g}(t) + v_\sigma(t)\vec{g}_2(x), \qquad \text{on } \Sigma, \\ v(s) &= v_0(s), \qquad s \in [-\sigma, 0], \end{aligned} \tag{69}$$

with $\int_{\partial\Omega} \vec{g}_i(x) \nu dS = 0$ for $i = 1, 2$, and $\vec{w}(0, x) = \vec{w}_0(x)$ for $x \in \Omega$. Where η is a constant, $v(\cdot) : [-\sigma, T] \to R$ and $\vec{g}_i(\cdot) : \partial\Omega \to R^n$, $i = 1, 2$.

Our task is to find v, or rather, its time derivative $u = v_t$ such that the cost functional

$$\int_\Omega |\vec{w}(T,x) - \vec{w}^d(x)|\,dx + 2\eta \int_0^T \int_\Omega |\nabla\left(\vec{w}(t,x) - \vec{w}^d(t,x)\right)|^2\,dxdt \\ + \lambda \int_0^T \left(v_t(t) - v_t^d\right)^2 dt \tag{70}$$

attains its minimum value.

In practical terms, $\vec{w} : [0,T] \times \Omega \to R^n$ is a smooth desired velocity field, $v^d : [-\sigma, T] \to R$ is a nominal forcing speed and $\lambda > 0$ is some given number. The second term in the cost functional (70) represents the cost of forcing.

This problem in the non-delay case has been considered by Fattorini and Sritharan. We now reformulate this problem. Consider the space

$$H = \{u : \Omega \to R^n,\ u \in L^2(\Omega),\ \nabla \cdot u = 0\}$$

and

$$V = H \times H_0^1(\Omega).$$

Using the method which has been used in [11], we reduce this problem to a distributed control problem by taking $y = z + w_1 u + w_2 u_\sigma$, where w_1 and w_2 are the Leray-Hopf cut functions corresponding to $\vec{g}_1$ and $\vec{g}_2$ respectively (for detail see [11] or [8]). Then taking the projection on H, we get the system:

$$\begin{aligned} z_t + \eta A z + B(z,z) + v(t)B_1(z) + v_\sigma(t)B_2(z) \\ = f_{11}v(t) + f_{12}v^2(t) + f_{21}v_\sigma(t) + f_{22}v_\sigma^2(t) + f_{33}v_\sigma(t)v(t) \\ + f_{31}v_t(t) + f_{32}v_{\sigma t}(t), \\ z(0,x) = z_0(x) \ \text{ and } \ v(s) = v^d(s), \qquad \text{for } s \in [-\sigma, 0], \end{aligned} \tag{71}$$

where z_0 is determined by w_0, $v(t)$ and Leray-Hopf cut functions. A is Stokes operator which is selfadjoint and (since $\Omega \subset R^n$ is bounded) positive defined. The operator $B(\cdot,\cdot)$ and $B_1(\cdot)$ are standard in Navier-Stokes theory ([6], [9], [11]). Note that $f_{ij} \in H$ for $i, j = 1, 2, 3$. It can be shown that $D(A) = H^2(\Omega) \cap V$, $D(A^{\frac{1}{2}}) = V$ and the linear operator $B_i \in L(D(A^{\frac{1}{2}}), H)$ for $i = 1, 2$.

Let us now set

$$y = \begin{pmatrix} z \\ v \end{pmatrix},\ u = v_t,\ \widetilde{A} = \begin{pmatrix} \eta A & 0 \\ 0 & 1 \end{pmatrix},\ B_1 = \begin{pmatrix} f_{31} \\ 1 \end{pmatrix},\ B_2 = \begin{pmatrix} f_{32} \\ 0 \end{pmatrix},$$

$$\beta_1(y, y_\sigma) = \begin{pmatrix} b(z,z) + vB_1(z) - (f_{11} + v_\sigma f_{33})v - f_{12}v^2 \\ -v \end{pmatrix},$$

and

$$\beta_2(y_\sigma) = \begin{pmatrix} -f_{21}v_\sigma - f_{22}v_\sigma^2 \\ 0 \end{pmatrix}.$$

Then (71) can be written as

$$y_t = \widetilde{A}y + \beta_1(y, y_\sigma) + \beta_2(y_\sigma) = B_1u + B_2u_\sigma, \quad \text{in } G$$
$$y(0) = \begin{pmatrix} z_0 \\ v^d \end{pmatrix} \text{ for } x \in \Omega \qquad \text{and} \qquad v(s) = v^d(s) \text{ for } x \in [-\sigma, 0]. \tag{72}$$

THEOREM 8. *Let $u \in L^2([-\sigma, T]; R)$ and $u(0) \in H$. Then there exists a unique solution $y \in L^2([0, T]; D(A)) \cap C([0, T]; D(A^{\frac{1}{2}}))$ to (69). Moreover, if $u(0) \in D(A^{\frac{1}{2}})$, then $y_t \in L^2([0, T]; H)$.*

PROOF. (72) is equivalent to the following problem:

$$y_t + \widetilde{A}y + \beta_1(y, y_\sigma) = B_1u + B_2u_\sigma - \beta_2(y_\sigma), \quad \text{in } G$$
$$y(0) = \begin{pmatrix} u(0) \\ v^d \end{pmatrix} \text{ for } x \in \Omega \qquad \text{and} \qquad v(s) = v^d(s) \text{ for } x \in [-\sigma, 0]. \tag{73}$$

In the interval $[0, \sigma]$, taking $N(v) = \beta_1(y, v^d)$ and using Theorem 3.3 of [8], we see that there exists a solution $y_1 \in L^2([0, \sigma]; D(A^{\frac{1}{2}})) \cap C([0, \sigma]; H)$ to (73) in $[0, \sigma]$. Moreover, if $\zeta \in D(A^{\frac{1}{2}})$, then $y_1 \in L^2([0, \sigma]; D(A)) \cap C([0, \sigma]; D(A^{\frac{1}{2}}))$ and $y_{1t} \in L^2([0, \sigma]; H)$ (Theorem 3.3 of [8] can be also found in [9]). Then, using the same arguments on (73) in the interval $[(i-1)\sigma, i\sigma]$ iteratively for $i = 2, 3, \ldots,$ and finally in the interval $[T - \sigma, T]$, we complete the proof of this theorem.

Using the above Theorem, we can easily conclude the following theorem

THEOREM 9. *Let $(\hat{y}, \hat{u})$ be an optimal pair of* (70) *with the state equation* (71). *Then there exists a function $p \in W^{1,2}([0, T]; H) \cap L^2([0, T]; H_\alpha)$ satisfying the following equations:*

$$-p_t(t) + \widetilde{A}p(t) + \left(\widetilde{M}(\hat{y}, \hat{y}_\sigma), p(t)\right) + \left(N(\hat{y}(t+\sigma), \hat{y})^1 p(t)\right) \in D_2g(t, \hat{y}(t)), \quad t \in [O, T - \sigma], \tag{74}$$
$$-p_t(t) + \widetilde{A}p(t) + \left(\widetilde{M}(\hat{y}, \hat{y}_\sigma), p(t)\right) \in D_2g(t, \hat{y}(t)), \quad t \in [T - \sigma, T], \tag{75}$$
$$p(T) \in -D\phi_0(\hat{y}(T)) \text{ and } p(s) = 0, \qquad s \in [T, T + \sigma], \tag{76}$$
$$B^*p(t) \in Dh(\hat{u}(t)),$$

where $\widetilde{M}$ and $\widetilde{N}$ are defined as follows $\forall z = (z_1, z_2)$:

$$\left(\widetilde{M}(y, y_\sigma), z\right) = (D_1\beta_1(y, y), z) = \begin{pmatrix} 2B(v, z_1) + uB_1(z_1) - (f_{11} + u_\sigma f_{33})z_2 - 2f_{12}uz_2 \\ -z_2 \end{pmatrix},$$

and

$$\left(\widetilde{N}(y, y_\sigma), z\right) = (D_2\beta_1(y(t+\sigma), y) + D\beta_2(y), z)$$
$$= \begin{pmatrix} -f_{12}z_2 - 2f_{22}u_\sigma z_2 \\ 0 \end{pmatrix},$$

respectively. Moreover, $h(u(t))$, $g(t, y)$ *and* $\phi_0(y(T))$ *are defined as in* [8].

6. Acknowledgement

I would like to thank Prof. V. Barbu, who proposed this problem to me and helped me when I worked on it. I would like to thank Prof. S. Stojanovic for his help as well.

References

[1] H. T. Banks and M. Q. Jacobs, "An approach to optimal control of functional differential equations with function space boundary conditions", *J. Differential Equations* **13** (1973) .

[2] V. Barbu, "Nonlinear analysis and nonlinear optimal control problems", manuscript.

[3] V. Barbu, "Convex control problems for linear differential systems of retarded type", *Ricerche Math.* **XXVI** (1977) 3–26.

[4] V. Barbu and T. H. Precupanu, *Convexity and optimization in Banach spaces*, Editura Academic (Bucuresti/Sijthoff and Noordoff, Groningen, 1978).

[5] F. Colonius, "Necessary optimality conditions for nonlinear hereditary differential systems with function space end constraints", in *Functional Differential Systems and Retarded Topics* (ed. M. Kisielenicz), (The Higher College of Engineering in Zielona Gora, Zielona Gora, Poland, 1980) 62–71.

[6] I. Ekeland, "Nonconvex minimization problems", *Bull. Amer. Soc.* **1** (1979) 443–474.

[7] H. O. Fattorini, "A unified theory of necessary conditions for nonlinear nonconvex control systems", *Appl. Math. Optim.* **15** (1987) 141–185.

[8] H. O. Fattorini and S. S. Sritharan, "Necessary and sufficient conditions for optimal controls in viscous flow problems", PAM Report No: 118, University of Colorado, to appear in *Proceedings of the Royal Society of Edinburgh Series A*.

[9] O. A. Ladyzhenskaya, *The mathematical theory of viscous incompressible flow*, 2nd ed., English translation (Gordon and Breach, New York, 1969).

[10] S. S. Sritharan, "Dyanamic programming of the Navier-Stokes equation", *Systems and Control Letters* **16** (1991) 299–307, North-Holland.

[11] R. Temam, *Navier-Stokes equations*, 3rd ed. (North-Holland Publ. Comp., Amsterdam, 1984).

[12] D. Tiba, "Optimal control of nonsmooth distributed parameter systems", in *Lecture notes in mathematics 1459*, (Springer-Verlag, Berlin, New York, c. 1990).