COLLOQUIUM MATHEMATICUM

VOL. LX/LXI — DÉDIÉ À M. ANTONI ZYGMUND — 1990

OPTIMAL CONTROLS FOR DISTRIBUTED PARAMETER SYSTEMS WITH MIXED CONSTRAINTS

BY

JIONGMIN YONG (SHANGHAI)

1. Introduction. Consider a semilinear controlled evolution system in some Banach space X:

$$\dot{x}(t) = Ax(t) + f(t, x(t), u(t)), \qquad t \in [0, T], \tag{1.1}$$

where A generates a C_0-semigroup e^{At} on X and f is a given map. The optimal control problem we are going to consider is the following:

PROBLEM MC. Minimize the functional

$$J(u(\cdot)) = \int_0^T f^0(t, x(t), u(t))\, dt \tag{1.2}$$

over all pairs $(x(\cdot), u(\cdot))$ subject to the state equation (1.1) (in some mild sense), the end points constraint

$$h(x(0), x(T)) = 0, \tag{1.3}$$

and the mixed constraint of the state and control

$$g(t, x(t), u(t)) = 0, \qquad \text{a.e. } t \in [0, T], \tag{1.4}$$

where f^0, g, h are given real-valued functions.

The main feature of our problem is that the maps $f^0(t, x, u)$, $g(t, x, u)$ and $h(x, y)$ will only be assumed to be Lipschitz continuous (*not necessarily* C^1!) in x and (x, y), respectively. We find that this allows our result to cover many interesting problems with state and/or mixed type constraints. To take a glance of it, let us give the following simple example. Let Ω be a convex and closed subset of $X \times X$ and let $Q(t)$ be a family of closed subsets of X with certain measurability in t. We look at the following constraints:

$$(x(0), x(T)) \in \Omega, \tag{1.5}$$

This work was partially supported by the Chinese National Natural Science Foundation under Grant 0188416 and the Chinese State Education Commission Science Foundation Grant No. 8824601.

$$x(t) \in Q(t), \qquad \text{a.e. } t \in [0,T]. \tag{1.6}$$

Set

$$g(t,x,u) = d(x,Q(t)) \equiv \inf\{|x-y| : y \in Q(t)\},$$
$$h(x,y) = d((x,y),\Omega) \equiv \inf\{(|x-\hat{x}|^2 + |y-\hat{y}|^2)^{1/2} : (\hat{x},\hat{y}) \in \Omega\}.$$

Then (1.5)–(1.6) is a special case of (1.3)–(1.4). It is clear that we may cook up many interesting examples similar to the above. In particular, if we let $g(t,x,u) \equiv 0$, then we have the problem with no mixed constraints. Thus, our problem properly contains those given in [12, 13]. If we further take $h(x_0,x_1) \equiv x_0 - \bar{x}_0$, then the problem becomes the one with a given initial state $\bar{x}_0$ and free terminal state. This shows that our problem also essentially covers those given in [1, 2, 14].

In this paper, we will apply the method of [13] together with some ideas of [2, 6, 7, 18] to derive the Pontryagin type maximum principle for any optimal trajectory and control of Problem MC. The key tool is Ekeland's variational principle ([9, 10]) and the concept of Clarke's generalized gradient ([6, 7]).

We refer to [5, 6, 7, 16] for classical results about the Pontryagin Maximum Principle for finite-dimensional cases and to [1–4, 11–14, 17] for infinite-demensional cases. We should note that in [11], Fattorini discussed a general input-output system under the condition that the reachable set of the variational system is finite-codimensional. Essentially, the problem studied was also of no mixed constraint. In [6, 7], Clarke studied finite-dimensional systems with mixed constraints. The approach we use in this paper is different from that of [6, 7]. Also, unlike [6, 7], our final result will not involve measures. It seems to us that by using the idea of [18], for systems governed by parabolic partial differential equations, the term f appearing in the state equation can also be only assumed to be Lipschitz continuous in x. We will work out the details elsewhere.

2. Assumptions and the main result. In this section, we state the control problem, the basic assumptions and the main result. Let X be a Banach space with norm $|\cdot|$ (from the context, there will be no ambiguity with the absolute value, of course), and let X^* be the dual space of X with $\langle\cdot,\cdot\rangle$ being the duality between X and X^*. Let U be a metric space with metric $\bar{\rho}$ and $T > 0$ be a constant. We let

$$\mathcal{U} = \{u(\cdot) : [0,T] \to U, \;\; u(\cdot) \text{ is measurable}\}.$$

Now, let us make the following basic hypotheses:

(H1) The operator $A : \mathcal{D}(A) \subset X \to X$ is the generator of some C_0-semigroup e^{At} on the space X.

(H2) The maps $f : [0,T] \times X \times U \to X$, $f^0 : [0,T] \times X \times U \to \mathbf{R}$, $g : [0,T] \times X \times U \to \mathbf{R}^+ \equiv [0,\infty)$ and $h : X \times X \to \mathbf{R}^+$ satisfy:

(i) f, f^0 and g are strongly measurable in t, continuous in u and Lipschitz continuous in x. Moreover, for some constant $L > 0$,

$$
\begin{aligned}
(2.1)\quad & |f(t,x,u) - f(t,\hat{x},u)|,\ |f^0(t,x,u) - f^0(t,\hat{x},u)|, \\
& |g(t,x,u) - g(t,\hat{x},u)| \le L|x - \hat{x}|, \qquad \forall t \in [0,T], x, \hat{x} \in X, u \in U, \\
(2.2)\quad & |f(t,0,u)|, |f^0(t,0,u)|, |g(t,0,u)| \le L \qquad \forall\, (t,u) \in [0,T] \times U,
\end{aligned}
$$

and f is continuously Fréchet differentiable in x.

(ii) There exists a constant $L > 0$ (for simplicity, we take it the same as in (i)) such that

$$
(2.3)\quad |h(x,y) - h(\hat{x},\hat{y})| \le L(|x - \hat{x}| + |y - \hat{y}|), \qquad \forall\, (x,y), (\hat{x},\hat{y}) \in X \times X.
$$

Remark 2.1. The constant $L > 0$ can be replaced by different $L^1(0,T)$ functions in different places of (2.1)–(2.3).

It is easy to see that under (H1)–(H2), for any $x_0 \in X$ and $u(\cdot) \in \mathcal{U}$, there exists a unique solution of the following integral equation ([7, 11]):

$$
(2.4)\qquad x(t) = e^{At}x_0 + \int_0^t e^{A(t-r)} f(r, x(r), u(r))\, dr, \qquad t \in [0,T].
$$

We call the solution $x(\cdot) \in C([0,T]; X)$ of (2.4) the *response* of our controlled evolution system under control $u(\cdot)$. Sometimes, $x(\cdot)$ is also called the *mild solution* of (1.1).

Now, our optimal control problem can be restated as follows:

PROBLEM MC. Minimize the functional

$$
(2.5)\qquad J(u(\cdot)) = \int_0^T f^0(t, x(t), u(t))\, dt
$$

subject to

$$
(2.6)\qquad \begin{cases}
u(\cdot) \in \mathcal{U}, \\
x(t) = e^{At}x_0 + \int_0^t e^{A(t-r)} f(r, x(r), u(r))\, dr, & t \in [0,T], \\
g(t, x(t), u(t)) = 0, & \text{a.e. } t \in [0,T], \\
h(x(0), x(T)) = 0.
\end{cases}
$$

We assume throughout the paper that there exists a pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ such that (2.6) is satisfied and the functional (2.5) is minimized. Any such pair is called an *optimal pair*.

Before stating the main result of the paper, let us recall the notion of generalized gradient for Lipschitz continuous functionals on Banach spaces (see [7]). Let Z be a Banach space and let $\ell : Z \to \mathbf{R}$ be Lipschitz continuous. Then we define

$$\ell^0(z;v) = \overline{\lim_{z'\to z,\varepsilon\downarrow 0}} \frac{\ell(z'+\varepsilon v)-\ell(z')}{\varepsilon}, \quad \forall\, z,v \in Z, \tag{2.7}$$

$$\partial\ell(z) = \{\zeta \in Z^* \mid \langle \zeta, v\rangle \le \ell^0(z;v),\ \forall\, v \in Z\}. \tag{2.8}$$

For details, see [7]. Next, for any $z \in Z$, we define

$$\mathcal{T}_\ell(z) = \{\hat z \in Z \mid \langle \hat z, \zeta\rangle \le 0,\ \forall\, \zeta \in \partial\ell(z)\}, \tag{2.9}$$

$$\mathcal{N}_\ell(z) = \{\zeta \in Z^* \mid \langle \hat z, \zeta\rangle \le 0,\ \forall\, \hat z \in \mathcal{T}_\ell(z)\}. \tag{2.10}$$

We call $\mathcal{T}_\ell(z)$ and $\mathcal{N}_\ell(z)$ the *generalized tangent* and *normal cone* of the functional $\ell(\cdot)$ at $z \in Z$, respectively. It is clear that they are weakly (weakly*) closed convex cones. If $\ell(\cdot)$ is Fréchet differentiable, they are actually the usual tangent space and the normal cone of the level set $\{\hat z \in Z \mid \ell(\hat z) = \ell(z)\}$ at z.

Now, assume $(\bar x(\cdot), \bar u(\cdot))$ is an optimal solution to Problem MC. We set

$$F(t) = f_x(t, \bar x(t), \bar u(t)), \qquad t \in [0,T],$$

and define the evolution operator $G(\cdot,\cdot) : \{(t,s) \in [0,T]\times[0,T] \mid s \le t\} \to \mathcal{L}(X)$ by

$$G(t,s)y = e^{A(t-s)}y + \int_s^t e^{A(t-r)}F(r)G(r,s)y\,dr, \qquad 0 \le s \le t \le T. \tag{2.11}$$

Then we define

$$\mathcal{R} = \Big\{\xi \in X \mid \xi = \int_0^T G(T,r)[f(r,\bar x(r),u(r)) - f(r,\bar x(r),\bar u(r))]\,dr, \quad u(\cdot) \in \mathcal{U}\Big\}, \tag{2.12}$$

$$\mathcal{Q} = \{\eta \in X \mid \eta = z_1 - G(T,0)z_0,\ (z_0,z_1) \in \mathcal{T}_h(\bar x(0), \bar x(T))\}. \tag{2.13}$$

Next, we introduce the Hamiltonian

$$H(t,x,u,\psi,\psi^0,\varphi^0) = \langle \psi, f(t,x,u)\rangle + \psi^0 f^0(t,x,u) + \varphi^0 g(t,x,u), \quad \forall\,(t,x,u,\psi,\psi^0,\varphi^0) \in [0,T]\times X \times U \times X^* \times \mathbf{R}\times\mathbf{R}, \tag{2.14}$$

where $\langle\cdot,\cdot\rangle$ is the duality between X and X^*.

Now, let us give one more hypothesis:

(H3) There exists a $\delta \in (0,1]$ and a neighborhood $\mathcal{O}$ of $(\bar x(0), \bar x(T))$ such that

$$|a_0|^2_{X^*} + |a_1|^2_{X^*} \ge \delta^2, \quad \forall\,(a_0,a_1) \in \partial h(x_0,x_1), (x_0,x_1) \in \mathcal{O}, \text{ with } h(x_0,x_1) > 0. \tag{2.15}$$

It is clear that if h is continuously Fréchet differentiable and $Dh(\bar{x}(0), \bar{x}(T))$, the Fréchet derivative of h at $(\bar{x}(0), \bar{x}(T))$, is bijective, then (H3) holds. Also, by [13], we know that if

$$h(x_0, x_1) = d_\Omega(x_0, x_1) \equiv \inf_{(y_0, y_1)\in\Omega} \{|y_0 - x_0|^2 + |y_1 - x_1|^2\}^{1/2}, \qquad \forall (x_0, x_1) \in X \times X,$$

with Ω being a subset of $X \times X$, then (H3) holds.

Now, let us state the main result of this paper.

THEOREM 2.4 (Maximum Principle). *Let* (H1)–(H3) *hold. Let* $(\bar{x}(\cdot), \bar{u}(\cdot))$ *be an optimal pair for Problem* MC. *Let* $\mathcal{R}$ *and* $\mathcal{Q}$ *be defined as in* (2.12)–(2.13) *and suppose* $\mathcal{R} - \mathcal{Q}$ *is finite-codimensional in* X ([11–13]). *Then there exists a triplet* $(\psi(\cdot), \psi^0, \varphi^0) \in C([0,T]; X^*) \times \mathbf{R} \times \mathbf{R}$ *such that*

$$(\psi(\cdot), \psi^0, \varphi^0) \neq 0, \tag{2.16}$$

$$\psi^0, \varphi^0 \leq 0, \tag{2.17}$$

$$\psi(t) = e^{A^*(T-t)}\psi(T) + \int_t^T e^{A^*(r-t)}\zeta(r)\, dr, \quad t \in [0,T], \tag{2.18}$$

$$\zeta(r) \in \partial_x H(r, \bar{x}(r), \bar{u}(r), \psi(r), \psi^0, \varphi^0), \quad a.e.\ r \in [0,T],$$

$$H(t, \bar{x}(t), \bar{u}(t), \psi(t), \psi^0, \varphi^0) = \max_{u\in U} H(t, \bar{x}(t), u, \psi(t), \psi^0, \varphi^0), \quad a.e.\ t \in [0,T], \tag{2.19}$$

$$(-\psi(0), \psi(T)) \in \mathcal{N}_h(\bar{x}(0), \bar{x}(T)). \tag{2.20}$$

We see that, roughly speaking, $\psi(\cdot)$ satisfies the evolution equation

$$\dot{\psi}(t) \in -A^*\psi(t) - \partial_x H(t, \bar{x}(t), \bar{u}(t), \psi(t), \psi^0, \varphi^0), \quad t \in [0,T], \tag{2.21}$$

which is the "adjoint system" of (2.1) along the optimal pair $(\bar{x}(\cdot), \bar{u}(\cdot))$, while (2.20) is the transversality condition.

3. Proof of the main result. First of all, for any $(x_0, u(\cdot)) \in X \times \mathcal{U}$, we denote the unique solution of (2.4) by $x(\cdot\,; x_0, u(\cdot))$ and define

$$x^0(t; x_0, u(\cdot)) = \int_0^t f^0(r, x(r; x_0, u(\cdot)), u(\cdot))\, dr, \qquad t \in [0,T], \tag{3.1}$$

$$y(t; x_0, u(\cdot)) = \int_0^t g(r, x(r; x_0, u(\cdot)), u(\cdot))\, dr, \qquad t \in [0,T]. \tag{3.2}$$

We note that since g is nonnegative, the mixed constraint in (2.6) is equivalent to

$$y(T;x_0,u(\cdot))=0. \tag{3.3}$$

Next, we introduce a metric on $X\times\mathcal{U}$ as follows:

$$\begin{aligned}\hat{d}((x_0,u(\cdot)),(\tilde{x}_0,\tilde{u}(\cdot)))&\\ =[|x_0-\tilde{x}_0|^2&+\operatorname{meas}\{t\in[0,T]:u(t)\neq\hat{u}(t)\}^2]^{1/2},\\ &\forall\,(x_0,u(\cdot)),(\tilde{x}_0,\tilde{u}(\cdot))\in X\times\mathcal{U}.\end{aligned} \tag{3.4}$$

Similar to [11], we know that $(X\times\mathcal{U},\hat{d})$ is a complete metric space (see also [13]). Now, we let $(\bar{x}(\cdot),\bar{u}(\cdot))$ be an optimal pair of Problem MC with $\bar{x}(0)=x_0$ and let $\bar{x}^0(\cdot)$ and $\bar{y}(\cdot)$ be the corresponding functions defined by (3.1) and (3.2). Then, for any $\varepsilon\geq 0$, we define

$$\begin{aligned}F_\varepsilon(x_0,u(\cdot))=\{h(x_0,x(T;x_0,u(\cdot)))^2&+d_{\Omega(\varepsilon)}(x^0(T;x_0,u(\cdot)))^2\\ &+y(T;x_0,u(\cdot))^2\}^{1/2},\quad\forall\,(x_0,u(\cdot))\in X\times\mathcal{U},\end{aligned} \tag{3.5}$$

where $\Omega(\varepsilon)=(-\infty,\bar{x}^0(T)-\varepsilon]$. It is immediate to show that the functional F_ε is continuous and everywhere defined on $(X\times\mathcal{U},\hat{d})$. Moreover, we see that

$$F_\varepsilon(x_0,u(\cdot))>0,\qquad\forall\,(x_0,u(\cdot))\in X\times\mathcal{U}, \tag{3.6}$$

$$F_\varepsilon(\bar{x}_0,\bar{u}(\cdot))=\varepsilon\leq\inf_{X\times\mathcal{U}}F_\varepsilon(x_0,u(\cdot))+\varepsilon. \tag{3.7}$$

Thus, by Ekeland's variational principle ([9, 10], see also [13]), we can find a pair $(x_0^\varepsilon,u^\varepsilon(\cdot))\in X\times\mathcal{U}$ such that

$$\hat{d}((x_0^\varepsilon,u^\varepsilon(\cdot)),(\bar{x}_0,\bar{u}(\cdot)))\leq\sqrt{\varepsilon}, \tag{3.8}$$

$$F_\varepsilon(x_0^\varepsilon,u^\varepsilon(\cdot))\leq F_\varepsilon(\bar{x}_0,\bar{u}(\cdot)), \tag{3.9}$$

$$\begin{aligned}F_\varepsilon(x_0,u(\cdot))\geq F_\varepsilon(x_0^\varepsilon,u^\varepsilon(\cdot))-\sqrt{\varepsilon}\hat{d}((x_0^\varepsilon,u^\varepsilon(\cdot)),(x_0,u(\cdot))),&\\ \forall\,(x_0,u(\cdot))\in X\times\mathcal{U}.&\end{aligned} \tag{3.10}$$

We set

$$\begin{aligned}x^\varepsilon(\cdot)&=x(\cdot;x_0^\varepsilon,u^\varepsilon(\cdot)),\\ x^{0,\varepsilon}(\cdot)&=x^0(\cdot;x_0^\varepsilon,u^\varepsilon(\cdot)),\\ y^\varepsilon(\cdot)&=y(\cdot;x_0^\varepsilon,u^\varepsilon(\cdot)).\end{aligned}$$

Next, let $(x_0,u(\cdot))\in X\times\mathcal{U}$ be fixed. Then, as in [12, 13], we have the following result.

PROPOSITION 3.1. *For any $\rho\in(0,1]$, there exists a measurable set $E_\rho\subset[0,T]$ such that*

$$\operatorname{meas}E_\rho=\rho T, \tag{3.11}$$

$$(3.12)\qquad \rho\int_0^t e^{A(t-r)}\Delta f_\varepsilon(r)\,dr=\int_{[0,t]\cap E_\rho} e^{A(t-r)}\Delta f_\varepsilon(r)\,dr+o(\rho),$$

$$(3.13)\qquad \rho\int_0^t \Delta f_\varepsilon^0(r)\,dr=\int_{[0,t]\cap E_\rho}\Delta f_\varepsilon^0(r)\,dr+o(\rho),$$

$$(3.14)\qquad \rho\int_0^t \Delta g_\varepsilon(r)\,dr=\int_{[0,t]\cap E_\rho}\Delta g_\varepsilon(r)\,dr+o(\rho),$$

uniformly in $t\in[0,T]$, *where*

$$(3.15)\qquad \begin{cases}\Delta f_\varepsilon(r)=f(r,x^\varepsilon(r),u(r))-f(r,x^\varepsilon(r),u^\varepsilon(r)), & r\in[0,T],\\ \Delta f_\varepsilon^0(r)=f^0(r,x^\varepsilon(r),u(r))-f^0(r,x^\varepsilon(r),u^\varepsilon(r)), & r\in[0,T],\\ \Delta g_\varepsilon(r)=g(r,x^\varepsilon(r),u(r))-g(r,x^\varepsilon(r),u^\varepsilon(r)), & r\in[0,T].\end{cases}$$

Now, we let

$$(3.16)\qquad x_{0,\rho}^\varepsilon=x_0^\varepsilon+\rho x_0,$$

$$(3.17)\qquad u_\rho^\varepsilon(t)=\begin{cases}u^\varepsilon(t), & t\in[0,T]\setminus E_\rho,\\ u(t), & t\in E_\rho.\end{cases}$$

Correspondingly, we let

$$\begin{aligned}x_\rho^\varepsilon(\cdot)&=x(\cdot;x_{0,\rho}^\varepsilon,u_\rho^\varepsilon(\cdot)),\\ x_\rho^{0,\varepsilon}(\cdot)&=x^0(\cdot;x_{0,\rho}^\varepsilon,u_\rho^\varepsilon(\cdot)),\\ y_\rho^\varepsilon(\cdot)&=y(\cdot;x_{0,\rho}^\varepsilon,u_\rho^\varepsilon(\cdot)).\end{aligned}$$

Then, from (3.10), (3.11) and (3.17), we have

$$(3.18)\qquad -\sqrt{\varepsilon}(|x_0|^2+T^2)^{1/2}\rho\le F_\varepsilon(x_{0,\rho}^\varepsilon,u_\rho^\varepsilon(\cdot))-F_\varepsilon(x_0^\varepsilon,u^\varepsilon(\cdot)).$$

Also, from (3.12) and (3.16)–(3.17), we see that

$$(3.19)\qquad x_\rho^\varepsilon(t)=x^\varepsilon(t)+\rho\xi_\varepsilon(t)+o(\rho),\qquad \text{uniformly in } t\in[0,T],$$

with

$$(3.20)\qquad \xi_\varepsilon(t)=e^{At}x_0+\int_0^t e^{A(t-r)}f_x(r,x^\varepsilon(r),u^\varepsilon(r))\xi_\varepsilon(r)\,dr + \int_0^t e^{A(t-r)}\Delta f_\varepsilon(r)\,dr,\quad t\in[0,T].$$

Hence, noting that h is nonnegative, one has

$$(3.21)\qquad \varlimsup_{\rho\downarrow 0}\frac{1}{\rho}[h(x_{0,\rho}^\varepsilon,x_\rho^\varepsilon(T))^2-h(x_0^\varepsilon,x^\varepsilon(T))^2]$$
$$\le 2h(x_0^\varepsilon,x^\varepsilon(T))$$

$$\times \varlimsup_{(x_0',x_1')\to(x_0^\varepsilon,x^\varepsilon(T)),\rho\downarrow 0} \frac{h(x_0'+\rho x_0, x_1'+\rho\xi_\varepsilon(T)) - h(x_0',x_1')}{\rho}$$

$$= 2h(x_0^\varepsilon, x^\varepsilon(T))h^0((x_0^\varepsilon, x^\varepsilon(T));(x_0,\xi_\varepsilon(T)))$$

$$= 2h(x_0^\varepsilon, x^\varepsilon(T))\max\{\langle a_0, x_0\rangle + \langle a_1, \xi_\varepsilon(T)\rangle : (a_0,a_1)\in\partial h(x_0^\varepsilon, x^\varepsilon(T))\}$$

$$= 2h(x_0^\varepsilon, x^\varepsilon(T))[\langle a_0^\varepsilon, x_0\rangle + \langle a_1^\varepsilon, \xi_\varepsilon(T)\rangle],$$

where

$$(a_0^\varepsilon, a_1^\varepsilon)\in\partial h(x_0^\varepsilon, x^\varepsilon(T)). \tag{3.22}$$

Next, we have

$$\begin{aligned} x_\rho^{0,\varepsilon}(T) - x^{0,\varepsilon}(T) &= \int_0^T [f^0(r, x_\rho^\varepsilon(r), u_\rho^\varepsilon(r)) - f^0(r, x^\varepsilon(r), u^\varepsilon(r))]\,dr \\ &= \int_0^T [f^0(r, x_\rho^\varepsilon(r), u^\varepsilon(r)) - f^0(r, x^\varepsilon(r), u^\varepsilon(r))]\,dr \\ &\quad + \rho\int_0^T \Delta f_\varepsilon^0(r)\,dr + o(\rho), \end{aligned} \tag{3.23}$$

and similarly,

$$\begin{aligned} y_\rho^\varepsilon(T) - y^\varepsilon(T) &= \int_0^T [g(r, x_\rho^\varepsilon(r), u^\varepsilon(r)) - g(r, x^\varepsilon(r), u^\varepsilon(r))]\,dr \\ &\quad + \rho\int_0^T \Delta g_\varepsilon(r)\,dr + o(\rho). \end{aligned} \tag{3.24}$$

Thus, in the case

$$d_{\Omega(\varepsilon)}(x_\rho^{0,\varepsilon}(T)) > 0, \qquad \text{for some sequence } \rho\downarrow 0, \tag{3.25}$$

we have

$$\begin{aligned} &\varlimsup_{\rho\downarrow 0}\left\{\frac{d_{\Omega(\varepsilon)}(x_\rho^{0,\varepsilon}(T))^2 - d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))^2}{\rho} + \frac{y_\rho^\varepsilon(T)^2 - y^\varepsilon(T)^2}{\rho}\right\} \\ &= \varlimsup_{\rho\downarrow 0}\left\{2d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))\frac{x_\rho^{0,\varepsilon}(T) - x^{0,\varepsilon}(T)}{\rho}\right. \\ &\quad \left. + 2y^\varepsilon(T)\frac{y_\rho^\varepsilon(T) - y^\varepsilon(T)}{\rho}\right\} \end{aligned} \tag{3.26}$$

$$
\begin{aligned}
&= 2\varlimsup_{\rho\downarrow 0}\Big\{d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T)) \\
&\quad\times\int_0^T \frac{f^0(r,x_\rho^\varepsilon(r),u^\varepsilon(r)) - f^0(r,x^\varepsilon(r),u^\varepsilon(r))}{\rho}\,dr \\
&\quad + y^\varepsilon(T)\int_0^T \frac{g(r,x_\rho^\varepsilon(r),u^\varepsilon(r)) - g(r,x^\varepsilon(r),u^\varepsilon(r))}{\rho}\,dr\Big\} \\
&\quad + 2\Big\{d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))\int_0^T \Delta f_\varepsilon^0(r)\,dr + y^\varepsilon(T)\int_0^T \Delta g_\varepsilon(r)\,dr\Big\} \\
&\le 2\int_0^T \varlimsup_{\rho\downarrow 0}\frac{1}{\rho}\{[d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))f^0(r,x_\rho^\varepsilon(r),u^\varepsilon(r)) \\
&\quad + y^\varepsilon(T)g(r,x_\rho^\varepsilon(r),u^\varepsilon(r))] - [d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))f^0(r,x^\varepsilon(r),u^\varepsilon(r)) \\
&\quad + y^\varepsilon(T)g(r,x^\varepsilon(r),u^\varepsilon(r))]\}\,dr \\
&\quad + 2\int_0^T \{d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))\Delta f_\varepsilon^0(r) + y^\varepsilon(T)\Delta g_\varepsilon(r)\}\,dr \\
&\le 2\int_0^T \{\langle b^\varepsilon(r),\xi_\varepsilon(r)\rangle + d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))\Delta f_\varepsilon^0(r) + y^\varepsilon(T)\Delta g_\varepsilon(r)\}\,dr,
\end{aligned}
$$

where

$$
\begin{aligned}
(3.27)\qquad b^\varepsilon(r) &\in \partial_x[d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))f^0(r,x^\varepsilon(r),u^\varepsilon(r)) \\
&\quad + y^\varepsilon(T)g(r,x^\varepsilon(r),u^\varepsilon(r))] \\
&\subset d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))\partial_x f^0(r,x^\varepsilon(r),u^\varepsilon(r)) \\
&\quad + y^\varepsilon(T)\partial_x g(r,x^\varepsilon(r),u^\varepsilon(r)) \qquad \text{a.e. } r\in[0,T].
\end{aligned}
$$

In case (3.25) fails, the above calculation is still valid since the terms containing $d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))$ simply vanish. Hence, we see from (3.18), (3.21) and (3.26) that

$$
\begin{aligned}
(3.28)\qquad & -\sqrt{\varepsilon}(|x_0|^2 + T^2)^{1/2} \\
&\le \frac{1}{F_\varepsilon(x_0^\varepsilon,u^\varepsilon(\cdot))}\{h(x_0^\varepsilon,x^\varepsilon(T))[\langle a_0^\varepsilon,x_0\rangle + \langle a_1^\varepsilon,\xi_\varepsilon(T)\rangle] \\
&\quad + \int_0^T [\langle b^\varepsilon(r),\xi_\varepsilon(r)\rangle + d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))\Delta f_\varepsilon^0(r) + y^\varepsilon(T)\Delta g_\varepsilon(r)]\,dr\} \\
&= \langle\psi_0^\varepsilon,x_0\rangle + \langle\psi_1^\varepsilon,\xi_\varepsilon(T)\rangle + \psi^{0,\varepsilon} \\
&\quad\times\frac{\int_0^T[\langle b^\varepsilon(r),\xi_\varepsilon(r)\rangle + d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))\Delta f_\varepsilon^0(r) + y^\varepsilon(T)\Delta g_\varepsilon(r)]\,dr}{(d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))^2 + y^\varepsilon(T)^2)^{1/2}},
\end{aligned}
$$

where

$$\psi_0^\varepsilon = \frac{h(x_0^\varepsilon, x^\varepsilon(T))a_0^\varepsilon}{F_\varepsilon(x_0^\varepsilon, u^\varepsilon(\cdot))}, \tag{3.29}$$

$$\psi_1^\varepsilon = \frac{h(x_0^\varepsilon, x^\varepsilon(T))a_1^\varepsilon}{F_\varepsilon(x_0^\varepsilon, u^\varepsilon(\cdot))}, \tag{3.30}$$

$$\psi^{0,\varepsilon} = \frac{(d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))^2 + y^\varepsilon(T)^2)^{1/2}}{F_\varepsilon(x_0^\varepsilon, u^\varepsilon(\cdot))} \tag{3.31}$$

and the convention $\frac{0}{0} = 0$ is taken. By our assumption, we know that

$$\begin{aligned} &|\psi_0^\varepsilon|_{X^*}^2 + |\psi_1^\varepsilon|_{X^*}^2 + (\psi^{0,\varepsilon})^2 \\ &= \frac{h(x_0^\varepsilon, x^\varepsilon(T))^2(|a_0^\varepsilon|_{X^*}^2 + |a_1^\varepsilon|_{X^*}^2) + d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))^2 + y^\varepsilon(T)^2}{F_\varepsilon(x_0^\varepsilon, u^\varepsilon(\cdot))^2} \geq \delta^2, \\ &\qquad \forall \varepsilon > 0. \end{aligned} \tag{3.32}$$

Now, define

$$\varphi^\varepsilon \equiv \int_0^T \{\langle \gamma^\varepsilon(r), \xi_\varepsilon(r)\rangle + \sigma_0^\varepsilon \Delta f_\varepsilon^0(r) + \sigma_1^\varepsilon \Delta g_\varepsilon(r)\}\, dr, \tag{3.33}$$

with

$$\sigma_0^\varepsilon = \frac{d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))}{[d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))^2 + y^\varepsilon(T)^2]^{1/2}} \geq 0, \tag{3.34}$$

$$\sigma_1^\varepsilon = \frac{y^\varepsilon(T)}{[d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))^2 + y^\varepsilon(T)^2]^{1/2}} \geq 0, \tag{3.35}$$

$$\begin{aligned} \gamma^\varepsilon(r) &= \frac{b^\varepsilon(r)}{[d_{\Omega(\varepsilon)}(x^{0,\varepsilon}(T))^2 + y^\varepsilon(T)^2]^{1/2}} \\ &\in \partial_x[\sigma_0^\varepsilon f^0(r, x^\varepsilon(r), u^\varepsilon(r)) + \sigma_1^\varepsilon g(r, x^\varepsilon(r), u^\varepsilon(r))]. \end{aligned} \tag{3.36}$$

We note that

$$(\sigma_0^\varepsilon)^2 + (\sigma_1^\varepsilon)^2 = 1, \qquad |\gamma^\varepsilon(r)| \leq L, \quad \text{a.e. } r \in [0, T]. \tag{3.37}$$

Thus, for some subsequences, one has

$$\sigma_0^\varepsilon \to \sigma_0, \qquad \sigma_1^\varepsilon \to \sigma_1, \qquad \gamma^\varepsilon(\cdot) \xrightarrow{*} \gamma(\cdot). \tag{3.38}$$

Also, we know that (from (3.9) and (3.20))

$$\xi_\varepsilon(t) \to \xi(t), \qquad \text{uniformly in } t \in [0, T], \tag{3.39}$$

with

$$\xi(t) = e^{At}x_0 + \int_0^t e^{A(t-r)} f_x(r, \bar{x}(r), \bar{u}(r))\xi(r)\, dr \tag{3.40}$$

$$+ \int_0^t e^{A(t-r)} \Delta \bar{f}(r)\, dr, \qquad t \in [0,T],$$

with

(3.41) $$\Delta \bar{f}(r) = f(r, \bar{x}(r), u(r)) - f(r, \bar{x}(r), \bar{u}(r)), \qquad \text{a.e. } r \in [0,T],$$

and

(3.42) $$\begin{cases} \Delta f_\varepsilon^0(r) \to \Delta \bar{f}^0(r) \equiv f^0(r, \bar{x}(r), u(r)) - f^0(r, \bar{x}(r), \bar{u}(r)), \\ \hfill \text{a.e. } r \in [0,T], \\ \Delta g_\varepsilon(r) \to \Delta \bar{g}(r) \equiv g(r, \bar{x}(r), u(r)) - g(r, \bar{x}(r), \bar{u}(r)), \\ \hfill \text{a.e. } r \in [0,T]. \end{cases}$$

Then we have

(3.43) $$\varphi^\varepsilon \to \int_0^T \{\langle \gamma(r), \xi(r)\rangle + \sigma_0 \Delta \bar{f}^0(r) + \sigma_1 \Delta \bar{g}_\varepsilon(r)\}\, dr \equiv \varphi$$

with (note the upper semicontinuity of the generalized gradient)

(3.44) $$\sigma_0^2 + \sigma_1^2 = 1,$$

(3.45) $$\gamma(r) \in \partial_x[\sigma_0 f^0(r, \bar{x}(r), \bar{u}(r)) + \sigma_1 g(r, \bar{x}(r), \bar{u}(r))], \qquad \text{a.e. } r \in [0,T].$$

On the other hand, by (3.22) and the upper semicontinuity of $\partial h(\cdot,\cdot)$, we have

(3.46) $$\langle \psi_0^\varepsilon, z_0\rangle + \langle \psi_1^\varepsilon, z_1\rangle \le O(\varepsilon), \qquad \forall\, (z_0, z_1) \in T_h(\bar{x}_0, \bar{x}(T)),$$

with the $O(\varepsilon)$ being uniform for (z_0, z_1) in bounded sets. Thus, we obtain

(3.47) $$\langle \psi_0^\varepsilon, x_0 - z_0\rangle + \langle \psi_1^\varepsilon, \xi(T) - z_1\rangle + \psi^{0,\varepsilon}\varphi^\varepsilon \ge -\theta_\varepsilon \to 0 \quad (\varepsilon \to 0).$$

Then, by our assumption and as in [11, 13], we can find some subsequence of $(\psi_0^\varepsilon, \psi_1^\varepsilon, \psi^{0,\varepsilon})$ (still denoted by the same symbol) such that

(3.48) $$\psi_0^\varepsilon \xrightarrow{*} \bar{\psi}_0, \qquad \psi_1^\varepsilon \xrightarrow{*} \bar{\psi}_1, \qquad \psi^{0,\varepsilon} \to \bar{\psi}^0,$$

with

(3.49) $$|\bar{\psi}_0|_{X^*}^2 + |\bar{\psi}_1|_{X^*}^2 + (\bar{\psi}^0)^2 > 0.$$

Then we obtain

(3.50) $$\begin{aligned} 0 &\le \langle \bar{\psi}_0, x_0\rangle + \langle \bar{\psi}_1, \xi(T)\rangle \bar{\psi}^0 \varphi \\ &= \langle \bar{\psi}_0, x_0\rangle + \Big\langle \bar{\psi}_1, G(T,0)x_0 + \int_0^T G(T,r)\Delta \bar{f}(r)\, dr \Big\rangle \\ &\quad + \bar{\psi}^0 \int_0^T \{\langle \gamma(r), \xi(r)\rangle + \sigma_0 \Delta \bar{f}^0(r) + \sigma_1 \Delta \bar{g}_\varepsilon(r)\}\, dr. \end{aligned}$$

We set

$$\psi(t) = e^{A^*(T-t)}\psi(T) + \int_t^T e^{A^*(r-t)} f_x(r,\bar{x}(r),\bar{u}(r))^*\psi(r)\,dr + \psi^0 \int_t^T e^{A^*(r-t)}\gamma(r)\,dr, \quad t \in [0,T], \tag{3.51}$$

with

$$\psi(T) = -\bar{\psi}_1, \qquad \psi^0 = -\bar{\psi}^0 \le 0. \tag{3.52}$$

Then we have

$$\psi(t) = G(T,t)^*\psi(T) + \psi_0 \int_t^T G(r,t)^*\gamma(r)\,dr, \qquad t \in [0,T]. \tag{3.53}$$

Hence, from (3.50), by some straightforward calculation (see also [13]), we obtain

$$\begin{aligned} 0 &\ge \langle \psi(T), \xi(T)\rangle - \langle \bar{\psi}_0, x_0\rangle + \psi^0\varphi \\ &= \int_0^T \{\langle \psi(r), \Delta\bar{f}(r)\rangle + \psi^0[\sigma_0\Delta\bar{f}^0(r) + \sigma_1\Delta\bar{g}(r)]\}\,dr \\ &\quad + \langle \psi(0) - \bar{\psi}_0, x_0\rangle. \end{aligned} \tag{3.54}$$

Then we see that

$$\psi(0) = \bar{\psi}_0, \tag{3.55}$$

$$\int_0^T \{\langle \psi(r), \Delta\bar{f}(r)\rangle + \psi^0[\sigma_0\Delta\bar{f}^0(r) + \sigma_1\Delta\bar{g}(r)]\}\,dr \le 0. \tag{3.56}$$

Now, set

$$\begin{cases} \hat{\psi}^0 = \psi^0\sigma_0, \\ \hat{\varphi}^0 = \psi^0\sigma_1, \\ \hat{\gamma}(r) = \psi^0\gamma(r), \quad r \in [0,T]. \end{cases} \tag{3.57}$$

Then we see that

$$(\psi(\cdot), \hat{\psi}^0, \hat{\varphi}^0) \ne 0, \tag{3.58}$$

$$\hat{\psi}^0, \hat{\varphi}^0 \le 0, \tag{3.59}$$

$$\hat{\gamma}(r) \in \partial_x[\psi^0 f^0(r,\bar{x}(r),\bar{u}(r)) + \varphi^0 g(r,\bar{x}(r),\bar{u}(r))], \quad \text{a.e. } r \in [0,T]. \tag{3.60}$$

Thus (3.51) reads

$$\psi(t) = e^{A^*(T-t)}\psi(T) + \int_t^T e^{A^*(r-t)}\gamma(r)\,dr, \quad \forall t \in [0,T], \tag{3.61}$$

$$\gamma(r) \in \partial_x H(r, \bar{x}(r), \bar{u}(r), \psi(r), \hat{\psi}^0, \hat{\varphi}^0), \quad \text{a.e. } r \in [0,T].$$

Dropping the $\hat{}$, we obtain (2.16)–(2.19). Noting the relation (3.55) and (3.46), we see that

$$\langle -\psi(0), z_0\rangle + \langle \psi(T), z_1\rangle \le 0, \qquad \forall (z_0, z_1) \in \mathcal{T}_h(\bar{x}_0, \bar{x}(T)). \tag{3.62}$$

Hence the transversality condition (2.20) follows.

Acknowledgement. The author would like to thank Professor X. Li of Fudan University for suggestive discussions.

REFERENCES

[1] N. U. Ahmed and K. L. Teo, *Optimal Control of Distributed Parameter Systems*, North-Holland, New York 1981.

[2] A. V. Balakrishnan, *Applied Functional Analysis*, Springer, New York 1976.

[3] H. T. Banks and G. Kent, *Control of functional equations to target sets in function space*, SIAM J. Control Optim. 10 (1972), 567–593.

[4] V. Barbu, *Optimal Control of Variational Inequalities*, Pitman, Boston 1984.

[5] L. D. Berkovitz, *Optimal Control Theory*, Springer, New York 1974.

[6] F. H. Clarke, *The maximum principle with minimum hypotheses*, SIAM J. Control Optim. 14 (1976), 1078–1091.

[7] — , *Optimization and Nonsmooth Analysis*, Wiley, New York 1983.

[8] R. F. Curtain and A. J. Pritchard, *Infinite Dimensional Linear Systems Theory*, Lecture Notes in Control and Inform. Sci. 8, Springer, New York 1978.

[9] I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. 47 (1974), 324–353.

[10] — , *Nonconvex minimization problems*, Bull. Amer. Math. Soc. (N.S.) 1 (1979), 443–474.

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

[12] X. Li and Y. Yao, *Maximum principle of distributed parameter systems with time lags*, Distributed Parameter Systems, Lecture Notes in Control and Inform. Sci. 75, Springer, New York 1985, 410–427.

[13] X. Li and J. Yong, *Necessary conditions of optimal control for distributed parameter systems*, SIAM J. Control Optim., to appear.

[14] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, New York 1971.

[15] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, New York 1983.

[16] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, *Mathematical Theory of Optimal Processes*, Wiley, New York 1962.

[17] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York 1972.

[18] J. Yong, *An optimal control problem for a nonsmooth semilinear distributed parameter system*, submitted.

DEPARTMENT OF MATHEMATICS
FUDAN UNIVERSITY
SHANGHAI 200433, CHINA

Reçu par la Rédaction le 30.10.1989