OPTIMAL CONTROL PROBLEMS FOR DISTRIBUTED PARAMETER SYSTEMS GOVERNED BY SEMILINEAR PARABOLIC EQUATIONS IN L^1 AND L^∞ SPACES.*

H. O. Fattorini

University of California, Department of Mathematics
Los Angeles, California 90024, USA

Abstract. We consider the infinite dimensional nonlinear programming problem of minimizing a real valued function $f_0(u)$ defined in a metric space V subject to the constraint $f(u) \in Y$, where $f(u)$ is defined in V and takes values in a Banach space E and Y is a subset of E. We use an extension of a theorem of Kuhn – Tucker type due to Frankowska to obtain Pontryagin's maximum principle for distributed parameter systems described by semilinear parabolic equations in spaces of bounded measurable functions and spaces of regular measures, under general assumptions on the nonlinear term.

1. Introduction. Optimal control problems for semilinear distributed parameter systems were considered in [FA2] as particular cases of optimization problems for input-output maps and in [FF1], [FF2] as infinite dimensional nonlinear programming problems. To motivate what follows we explain briefly this treatment.

Consider a semilinear distributed parameter system modelled as a differential equation in a Banach space E,

$$y'(t) = Ay(t) + f(t, y(t), u(t)), \tag{1.1}$$

$$y(0) = y^o, \tag{1.2}$$

where the operator A is the infinitesimal generator of a strongly continuous semigroup $S(t)$, $t \geq 0$. The initial condition y^o is fixed. The control $u(t)$ takes values in a second Banach space F and satisfies a constraint

$$u(t) \in U, \tag{1.3}$$

---

* This work was supported by the National Science Foundation under grant DMS-8701877

where the set U is a subset of F. The cost functional of the problem is

$$y_0(t, u) = \int_0^t f_0(s, y(s, u), u(s))ds \tag{1.4}$$

where $y(t, u)$ denotes the solution of (1.1)–(1.2) corresponding to the control u (precise conditions on the functions f and f_0 will be given later). The control problem is the usual one of minimizing the cost functional $y_0(t, u)$ among all u such that the trajectory $y(t, u)$ satisfies the target condition

$$y(t, u) \in Y \tag{1.5}$$

(Y a given subset of E). The endtime $\mathbf{t}$ of the control interval $0 \leq t \leq \mathbf{t}$ may be free or fixed.

For the fixed endtime control problem, we consider the functions

$$f(u) = y(\mathbf{t}, u), \qquad f_0(u) = y_0(\mathbf{t}, u) \tag{1.6}$$

(not to be confused with the functions $f(t, y, u)$ and $f_0(t, y, u)$ in (1.1) and (1.4)) and we define $V = V(0, \mathbf{t}; U)$ as the space of all admissible controls (controls subject to measurability conditions to be precised later and satisfying (1.3)). The control problem is a particular case of the following

Abstract nonlinear programming problem. Let V be a metric space, E a Banach space, Y a subset of E, $f : V \to E$, $f_0 : V \to \mathbf{R}$ = {real numbers}. Characterize the solutions of

$$\text{minimize } f_0(u) \tag{1.7}$$

$$\text{subject to } f(u) \in Y \tag{1.8}$$

The requirement that V be just a metric space (instead of, say, a Banach space) is due to the fact that the most convenient topology for $V(0, \mathbf{t}; U)$ is that given by the metric

$$d(u, v) = \text{meas}\,\{t;\, u(t) \neq v(t)\}. \tag{1.9}$$

introduced by Ekeland [E] in his proof of Pontryagin's maximum principle. A similar formulation can be used for the free time problem taking V as a space of pairs $(\mathbf{t}, u)$ endowed with some convenient metric. For the time optimal problem a different model is more useful. Here we use the fact that if $\mathbf{t}$ is the optimal time and $\mathbf{u}$ is the optimal control then $y(\mathbf{t}, \mathbf{u}) \in Y$ whereas if $\{t_n\}$ is a sequence with $t_n < t$ we have $y(t_n, u) \notin Y$ for any admissible control; moreover, by continuity of trajectories, in particular of the optimal trajectory, we have $y(t_n, u) \to y(t, u) \in Y$ if $t_n \to t$. Thus, considering the functions

(1.10) $$f_n(u) = y(t_n, u)$$

the time optimal problem is a particular case of the

Abstract time optimal problem. Let $\{V_n\}$ be a sequence of metric spaces, E a Banach space, Y a subset of E, $f_n: V_n \to E$ such that

(1.11) $$f_n(V_n) \cap Y = \emptyset .$$

Characterize the sequences $\{u_n\}$ such that

(1.12) $$\mathrm{dist}(f(u_n), Y) \to 0 .$$

A sequence $\{u_n\}$, $u_n \in V_n$ that satisfies (1.12) will be called an **optimal sequence** for the abstract time optimal problem.

Both optimal problems were studied in [FF1] and [FF2] in the Hilbert space setting, the results applied to optimal problems for distributed parameter systems, and in a less general formulation in [FA2], where also applications to boundary control systems are considered. There are several motivations to extend the results to the general Banach space setting. The Banach space generalization allows us to handle for instance target conditions of pointwise type $|y(t, u)(x) - y(x)| \leq C$. More importantly, it allows us to include in the abstract treatment problems with state constraints (see [FF3], which cannot be fitted in the Hilbert space setting. Problems with state constraints will not be treated in this paper.

Recently, Frankowska [FR2] [FR2] has generalized the results on the nonlinear programming problem (1.7)–(1.8) to the Banach space case. We state in this paper a generalization of this result under weaker conditions and apply it to parabolic distributed parameter systems. The results in this paper are relevant not only to solutions of (1.7)–(1.8) but to certain approximate or suboptimal solutions, which allows applications to relaxed solutions of optimal control problems [FA7]. Complete proofs of the results in this paper will appear elsewhere ([FA6]).

2. The Kuhn – Tucker theorem. We denote by $B(x, r)$ the ball of center x and radius $r \geq 0$ in an arbitrary metric space. Let $u \in V$ and let $\{D_n\}$ be a sequence of subsets of E. We denote by $\liminf_{n \to \infty} D_n$ the set of all y such that $\lim_{n \to \infty} \mathrm{dist}(y, D_n) = 0$.

Let g be a function from V into E, $u \in V$. The vector $\xi \in E$ is a *variation* of g at u if there exists a sequence $\{h_k\} \subset R_+ = \{$positive real numbers$\}$ with $h_k \to 0$ and a sequence $\{u_k\} \subset V$ with $d(u_k, u) \le h_k$ and such that

$$\frac{g(u_k) - g(u)}{h_k} \to \xi \quad \text{as} \quad k \to \infty .$$

The definition extends in an obvious way to the case where g is defined only in a domain $D(g)$ of V. We denote by $\partial g(u)$ (the *variation set* of g at u) the set of all such ξ.

We consider the abstract nonlinear programming problem (1.7)–(1.8) for a function f defined in a domain $D(f)$ and a function f_0 defined in a domain $D(f_0)$. We denote by $(f_0, f)\colon V \to R \times E$ the function $(f_0, f)(u) = (f_0(u), f(u))$, whose domain $D((f_0, f)) = D(f_0) \cap D(f)$ we assume nonempty. The hypotheses are the following:

(a) The space V is **complete.**

(b) The target set Y is **closed.**

(c) For each $y \in E$ and every $\mu \in R$ the functions

$\Phi(u, y) = |f(u) - y|$ if $u \in D(f)$, $\Phi(u, y) = +\infty$ if $u \notin D(f)$

$\Phi_0(u, y) = |f_0(u) - \mu|$ if $u \in D(f_0)$, $\Phi_0(u, y) = +\infty$ if $u \notin D(f_0)$

are **lower semicontinuous.**

The result below applies not only to solutions of (1.7)–(1.8) but also to certain types of approximate solutions. A sequence $\{u^n\} \subset V$ is called an approximate or suboptimal solution of (1.7)–(1.8) if

(2.1) $$\limsup_{n \to \infty} f_0(u^n) \le m, \quad \lim_{n \to \infty} \mathrm{dist}(f(u^n), Y) = 0,$$

where m is the minimum of (1.7) subject to (1.8). A sequence $\{y^n\} \subset Y$ will be said to be associated with the suboptimal solution $\{u^n\}$ if and only if

(2.2) $$|f(u^n) - y^n| = \varepsilon_n \to 0.$$

Theorem 2.1 Let $\{u^n\} \subset V$ be a suboptimal solution of (1.7)–(1.8) and let $\{y^n\}$ be a sequence associated with $\{u^n\}$. Then there exists a sequence $\{\delta_n\} \subset R_+$, $\delta_n \to 0$, a sequence $\{u^n\} \subset V$ and a sequence $\{y^n\} \subset Y$ with

$$(2.3)\qquad d(u^n, u^n) + |y^n - y^n| \le \delta_n$$

and such that: for every sequence $\{D_n\}$ (D_n a convex subset of $\partial(f_o, f)$) and every $\rho > 0$ there exists a sequence $\{\mu_n\} \subset \mathbf{R}$ and a sequence $\{z_n\} \subset E^*$ with

$$(2.4)\qquad \max(|\mu_n|, |z_n|) = 1,\ \mu_n \ge 0,$$

$$(2.5)\qquad \mu_n\eta^n + \langle z_n, \xi^n - w^n\rangle \ge -\delta_n(1 + \rho)$$

for $(\eta^n, \xi^n) \in \partial(f_o, f)(u^n)$ and $w^n \in C_Y(y^n) \cap B(0, \rho)$. ($C_Y(y)$ the tangent cone to Y at y). Moreover, for every limit point (μ, z) of $\{(\mu_n, z_n)\}$ in the weak (E^*, E) – topology of E^* we have

$$(2.6)\qquad \mu\eta + \langle z, \xi\rangle \ge 0$$

for every $(\eta, \xi) \in \liminf_{n\to\infty} D_n$. Finally,

$$(2.7)\qquad z \in (\liminf_{n\to\infty} C_Y(y^n))^- ,$$

where $^-$ indicates negative polar.

Under the hypotheses of Theorem 2.1, the multiplier (μ, z) may be zero. The following condition ([FR1], [FR2]) prevents this.

Theorem 2. 2 Let $\{D_n\}$ be the sequence of convex sets in Theorem 2.1 and assume there exists $\rho > 0$ and a compact set Q such that the intersection of all the sets

$$(2.8)\qquad \overline{\Pi(D_n) - C_Y(y^n) \cap B(0, \rho) + Q}$$

contains an interior point, where Π is the canonical projection of $\mathbf{R} \times E$ into E. and the bar indicates closure. Then (μ, z) in (2.6) is not zero.

If the space E has Gâteaux differentiable norm off the origin we may take $D_n = \partial(f_o, f)(u^n)$ or its closed convex hull in Theorems 2.1 and 2.2. If the norm is Fréchet differentiable off the origin and $f(u^n) \to y$ then the vector z belongs to the normal cone $N_Y(y) \subset E^*$.

The Kuhn – Tucker condition (2.6) for solutions of the nonlinear programming problem (1.7)–(1.8), as well as the condition on nontriviality of the multiplier were proved in [FR1] under the assumptions that the norm of E is Gâteaux differentiable off zero and that f is continuous, and in [FR2] for a general Banach space and f, f_0 Lipschitz continuous. The generalized version presented here and proved in [FA6], which applies as well to suboptimal solutions is a direct descendant of its Hilbert space version [FF1] [FF2] and the method of proof is similar. An ancestor of this Hilbert space version (where the setup and the hypotheses are much less general) was proved in [FA2]; the case where E is finite dimensional is closely related with the results of [EK1]. We note that allowing the maps f, f_0 to be defined only in subsets of V is decisive in the treatment of the point target problem ($Y = \{y\}$) for distributed parameter systems and boundary control systems.

The treatment of the abstract time optimal problem is similar. The result corresponding to Theorem 2.2 is

Theorem 2.3 Let $\{u^n\}$ be an optimal sequence for the time optimal problem and let $\{y^n\} \subset Y$ be a sequence associated with $\{u^n\}$ (that is, satisfying (2.3)). Then there exists a sequence $\{u^n\}$, $u^n \in V_n$ and a sequence $\{y^n\} \subset Y$ with

$$d(u^n, u^n) + |y^n - y^n| \le \varepsilon_n^{1/2} \tag{2.9}$$

and such that: for every sequence $\{D_n\}$ (D_n a convex subset of $\partial(f)$) and for every $\rho > 0$ there exists a sequence $\{z_n\} \subset E^*$ with

$$|z_n| = 1, \tag{2.10}$$

$$\langle z_n, \xi^n - w^n \rangle \ge -\varepsilon_n^{1/2}(1 + \rho) \tag{2.11}$$

for $\xi^n \in \partial(f)(u^n)$ and $w^n \in C_Y(y^n) \cap B(0, \rho)$. Moreover, for every limit point z of $\{z_n\}$ in the weak (E^*, E) – topology of E^* we have

$$\langle z, \xi \rangle \ge 0 \tag{2.12}$$

for every $(\eta, \xi) \in \liminf_{n \to \infty} D_n$. Finally, we have

$$z \in (\liminf_{n \to \infty} C_Y(y^n))^- . \tag{2.13}$$

Theorem 2.4 Let $\{D_n\}$ be the sequence of convex sets in Theorem 2.3 and assume there exists $\rho > 0$ and a compact set Q such that the intersection of all the sets

$$\overline{D_n - C_Y(y^n) \cap B(0, \rho) + Q} \tag{2.14}$$

contains an interior point. Then the multiplier z in Theorem 2.3 is not zero.

If the space E has Gâteaux differentiable norm off the origin we may take $D_n = \partial(f)(u^n)$ or its closed convex hull in Theorems 2.3 and 2.4. If the norm is Fréchet differentiable and $f(u^n) \to y$ then z belongs to the normal cone $N_Y(y)$.

3. Distributed parameter systems described by elliptic differential equations. Let Ω be a bounded domain of class $C^{(2)}$ with boundary Γ in m-dimensional Euclidean space R^m, and let A be a uniformly elliptic partial differential operator of class $C^{(2)}$,

$$Ay = \sum_{j=1}^{m} \sum_{k=1}^{m} \frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial y}{\partial x_k}\right) + \sum_{j=1}^{m} b_j(x)\frac{\partial y}{\partial x_j} + c(x)y$$

with a boundary condition β on Γ. This boundary condition may be either of Dirichlet type or of variational type $Dy = \gamma(x)y$ (D the conormal derivative). The operator A and the boundary condition β generate a strongly continuous semigroup $S(t, A, \beta)$ in the space $C(K)$ of continuous functions in K = closure of Ω, the space $C(K)$ endowed with the supremum norm (for the Dirichlet boundary condition the space $C(K)$ is replaced by its subspace $C_0(K)$ consisting of all functions vanishing on Γ).

The control system is described by the semilinear initial value problem in the space $E = C(K)$,

$$y'(t) = A(\beta)y(t) + f(t, y(t), u(t)), \tag{3.1}$$

$$y(0) = y^o, \tag{3.2}$$

where $A(\beta)$ is the infinitesimal generator of $S(t, A, \beta)$. There are various reasons to consider the equation (3.1) in a the space $C = C(K)$ rather than, say, in a space $L^p(\Omega)$. One is physical, to wit, the "natural" norm in which temperatures are measured in heat propagation processes is the supremum norm. Othe reason is purely mathematical: in $C(K)$ we may dispense with growth conditions on the nonlinear term which would be unavoidable in a space $L^p(\Omega)$. On the other hand, all the results may be directly applied (in some cases with simplifications) to spaces $L^p(\Omega)$ with $1 < p < \infty$, for instance, to the Navier –Stokes equations in $L^2(\Omega)$. Another way to treat the $L^p(\Omega)$ case is to go from $C(K)$ results to $L^p(\Omega)$ results via Sobolev imbeddings (see [FA3]).

Controls $u(t)$ take values in a closed, bounded subset U (called the *control set*) of a Banach space F and are either strongly or weakly measurable (as precised below). The space $V(0, t; U)$ of all (admissible) controls in an interval

$0 \le t \le \bar{t}$ will be equipped with the distance (1.9). In the assumptions below, E_α is the domain of the fractional power $(-A(\beta))^\alpha$ with its graph norm $|y|_\alpha$ and the function $f(t, y, u)$ is defined in $[0, T] \times E_\alpha \times U$ for some $\alpha, 0 \le \alpha < 1$. We denote by $C(0, \bar{t}; E_\alpha)$ the space of all continuous E_α–valued functions defined in the interval $0 \le t \le \bar{t}$ with its usual supremum norm.

(a) The function $t \to f(t, y(t), u(t))$ is $(L^1(\Omega), L^\infty(\Omega))$ – weakly measurable for each $y \in C(0, \bar{t}; E_\alpha)$ and each admissible control $u(t)$. For every $C > 0$ there exists a constant $K = K(C)$ such that

$$|f(t, y, u)| \le K \quad (0 \le t \le T,\ y \in E_\alpha, |y|_\alpha \le C, u \in U) \tag{3.3}$$

(b) $f(t, y, u)$ has a Fréchet derivative $\partial_y f(t, y, u)$ with respect to y. The function $t \to \partial_y f(t, y(t), u(t))z$ is $(L^1(\Omega), L^\infty(\Omega))$ – weakly measurable for each $y \in C(0, \bar{t}; E_\alpha)$, each admissible control $u(t)$ and every z in $C(K)$. For every $C > 0$ there exists a constant $L = L(C)$ such that

$$|\partial_y f(t, y, u)| \le L \quad (0 \le t \le T,\ y \in E_\alpha, |y|_\alpha \le C, u \in U) \tag{3.4}$$

We note that the requirement of weak $(L^1(\Omega), L^\infty(\Omega))$ measurability (rather than strong measurability) is motivated by the need to include such natural control terms as $f(t, y, u) = f(t, y) + u$. Here, the control functions are elements of the space $L^\infty(\Omega \times [0, T])$; these controls, considered as $L^\infty(\Omega)$ – valued functions are not strongly measurable but only $(L^1(\Omega), L^\infty(\Omega))$ – weakly measurable functions (we note that the present definition of weakly measurable is different from the standard one, where $L^1(\Omega)$ would be replaced by the dual of $L^\infty(\Omega)$). The treatment of (3.1)–(3.2) under these measurability assumptions is slightly nonstandard (see [FA2] for the linear case, [FA7] for the general case) but local existence and uniqueness of $C(K)$–valued solutions $y(t) = y(t, u)$ is proved by successive approximations in the usual way. In general, these solutions do not extend to the entire interval $0 \le t \le T$, but this causes no essential difficulty since we work in a neighborhood of an optimal control, assumed to exist. All that is needed is stability of the global existence property in the distance of the space $V(0, \bar{t}; U)$, which is covered by the following result.

Lemma 3.1 Let $\bar{u}(t) \in V(0, \bar{t}; U)$ be a control whose trajectory $y(t, \bar{u})$ exists in $0 \le t \le \bar{t}$. Then there exists $\delta > 0$ such that if $u(t) \in V(0, \bar{t}; U)$ and $d(u, \bar{u}) \le \delta$ then the trajectory $y(t) = y(t, u)$ also exists in $0 \le t \le \bar{t}$. Moreover, the function

$$u \to y(\bar{t}, u)$$

from $V(0, \bar{t}; U)$ into $C(0, \bar{t}; E)$ is Hölder continuous with exponent $1-\alpha$ in the ball $B(u, \delta)$.

4. Construction of convex cones of (limits of) variations. Construction of variations will be based on the "multiple spike perturbations" classical in control theory [P]. To define these perturbations, we consider a p-dimensional vector $s = (s_1, \ldots s_p)$ such that $0 < s_1 < \ldots < s_p < \bar{t}$, a p-dimensional nonnegative vector $\alpha = (\alpha_1, \ldots \alpha_p)$, $\alpha_j \geq 0$, and a p- dimensional vector $v = (v_1, \ldots v_p)$ of elements of the control set U. Given a control $u \in V(0, \bar{t}; U)$ and $h > 0$, define

$$
(4.1)\qquad
\begin{aligned}
&u_{s,\alpha,h,v}(t) = v_j \quad (s_j - \alpha_j h \leq t \leq s_j,\ j = 1,2,\ldots, p)\\
&u_{s,\alpha,h,v}(t) = u(t) \quad \text{elsewhere}
\end{aligned}
$$

for h so small that the spikes do not overlap.

Lemma 4.1 There exists a set e of full measure in $0 \leq t \leq \bar{t}$ such that, for every $s_j \in e$ we have

$$
(4.2)\qquad \lim_{h\to 0+} \frac{y(t, u_{s,\alpha,h,v}) - y(t,u)}{h} = \xi(t, s, \alpha, u, v) = \sum_{j=1}^{p} \alpha_j \xi(t, s_j, u, v_j),
$$

in $0 \leq t \leq \bar{t}$, where $\xi(t, s, u, v)$ is defined by

$$
(4.3)\qquad
\begin{aligned}
&\xi(t, s, u, v) = 0 \quad (t < s)\\
&\xi(t, s, u, v) = S(t, s, u)\{f(t, y(s, u), v) - f(t, y(s, u), u(s))\} \quad (t \geq s),
\end{aligned}
$$

$S(t, s, u)$ the solution operator of the linearized equation

$$
(4.4)\qquad z'(s) = (A(\beta) + \partial_y f(t, y(s, u), u(s)))z(s) \quad (0 \leq s \leq \bar{t}).
$$

Convergence in (4.2) is uniform outside of the intervals $(s_j + \delta, s_j + \delta)$ for any $\delta > 0$; moreover, $h^{-1}|y(t, u_{s,\alpha,h,v}) - y(t, u)|$ is bounded by a family of functions with equicontinuous integrals.

The assumptions on the kernel $f_0(t, y, u)$ of the cost functional (1.4) are

(a) $f_0(t, y, u)$ is measurable for each $y \in C(0, \bar{t}; E_\alpha)$ and each $u \in V(0, \bar{t}; E)$. For every $C > 0$ there exists a constant $K = K(C)$ such that

$$
(4.5)\qquad |f_0(t, y, u)| \leq K \quad (0 \leq t \leq T,\ y \in E_\alpha, |y|_\alpha \leq C, u \in U)
$$

(b) $f_0(t, y, u)$ has a Fréchet derivative $\partial_y f_0(t, y, u)$ with respect to y. The function $t \to \partial_y f_0(t, y(t), u(t))$ is measurable for each $y \in C(0, t; E_\alpha)$ and each $u \in V(0, t; E)$. For every $C > 0$ there exists a constant $L = L(C)$ such that

$$|\partial_y f_0(t, y, u)| \leq L \quad (0 \leq t \leq T,\ y \in E_\alpha, |y|_\alpha \leq C, u \in U)$$

Lemma 4.2 There exists a set e of full measure in $0 \leq t \leq t$ such that, for every $s_j \in e$ we have

$$\text{(4.6)} \qquad \lim_{h \to 0+} \frac{y_0(t, u_{s, \alpha, h, v}) - y_0(t, u)}{h} = \xi_0(t, s, \alpha, u, v) = \sum_{j=1}^{p} \alpha_j \xi_0(t, s_j, u, v_j),$$

where

$$\text{(4.7)} \qquad \begin{aligned} \xi_0(t, s, u, v) &= 0 \quad (t < s) \\ \xi_0(t, s, u, v) &= f_0(s, y(s, u), v) - f_0(s, y(s, u), u(s)) + \end{aligned}$$

$$+ \int_s^t \langle \partial_y f_0(\sigma, y(\sigma, u), u(\sigma)), \xi(\sigma, s, u, v) \rangle \, d\sigma \qquad (t \geq s)$$

The sets D_n required in Theorem 2.3 will consist of all elements of the form

$$\text{(4.8)} \qquad \xi(t, s, \alpha, u^n, v)) \in E$$

with arbitrary α and v and s with $s_j \in e$, where e is the set in Lemma 4.1, while the sets D_n in Theorem 2.2 consists of all the elements of the form

$$\text{(4.9)} \qquad (\xi_0(t, s, \alpha, u^n, v), \xi(t, s, \alpha, u^n, v)) \in R \times E$$

with arbitrary α and v and s with $s_j \in e_n$, where e_n is the set in Lemma 4.1 and Lemma 4.2 corresponding to u^n (we may assume they are the same). Obviously, the D_n are convex. To figure out the elements of $\liminf_{n \to \infty} D_n$ needed in Theorems 2.1 and 2.2 we use the following result:

Lemma 4.3 Let $\{t_n\}$ be a sequence with $t_n < t$, $\{u^n\}$ a sequence in $V(0, t; U)$. Assume that

$$\Sigma (t - t_n) < \infty, \qquad \Sigma d_n(u^n, u) < \infty,$$

where d_n is the distance (1.9) in the space $V(0, t_n; U)$. Then there exists a subset e of $\cap e_n$ of full measure in $0 \leq t \leq t$ such that, for $s \in e$ we have

$$\xi(t_n, s, u^n, v) \to \xi(t, s, u, v), \quad \xi_o(t_n, s, u^n, v) \to \xi_o(t, s, u, v)\,.$$

The proof of Lemma 4.3 is essentially the same as that of the corresponding Hilbert space result in [FF2]. It follows from Lemma 4.3 that elements of the form (4.8) (resp. 4.9) with $u^n = u$ and $s_j \in e$ belong to $\liminf_{n \to \infty} D_n$ and can thus be used in (2.12) (resp. in (2.6)).

5. The maximum principle Let $u(t)$ be an optimal control for the control problem with cost functional (1.4). We apply Theorem 2.1 with functions f, f_o defined by (1.6) to the sequence $\{u^n\} = \{u\}$, which obviously is an approximate solution (with $\varepsilon_n = 0$) of the nonlinear programming problem (1.7)–(1.8). Theorem 2.1 produces a multiplier $(\mu, z) \in R \times E$ satisfying inequality (2.6) for the convex cones of (limits of) variations constructed in the previous section. Manipulations similar to the ones in [FA2] then produce the maximum (or, rather, minimum) principle

$$\mu f_o(s, y(s, u), u(s)) + \langle z(s), f(s, y(s, u), u(s))\rangle = \tag{5.1}$$

$$= \min_{v \in U} \{\mu f_o(s, y(s, u), v) + \langle z(s), f(s, y(s, u), v)\rangle\}$$

for s almost everywhere in the control interval $0 \le t \le \bar{t}$, where $z(s)$ is the solution of the adjoint backwards initial value, or final value problem:

$$z'(s) = -(A(\beta)^* + \partial_y f(t, y(s, u), u(s))^*)z(s) - \mu\partial_y f_o(t, y(s, u), u(s)) \quad (0 \le s \le \bar{t})\,. \tag{5.2}$$

$$z(\bar{t}) = z\,. \tag{5.3}$$

There are some technicalities in this final value problem. Note first that the adjoint $A(\beta)^*$ is an operator in the dual of the space $C(\Omega)$, which can be identified with the space $\Sigma(K)$ of all regular Borel measures defined in K, the closure of Ω. In this space, $A(\beta)^*$ is not a semigroup generator, since it is not even densely defined. The treatment of (5.2) parallels closely that of (3.1), the role of the space $C(\Omega)$ played by $L^1(\Omega)$ and the role of $L^\infty(\Omega)$ played by $\Sigma(K)$; in particular, solutions of the backwards equation (5.2) belong to $L^1(\Omega)$ for $t < \bar{t}$, thus both sides of (5.1) make sense.

The key question is whether nontriviality of the multiplier (μ, z) (and thus of the maximum principle (5.1)) can be insured. By virtue of (4.2) and of a limiting argument, the closure of the set $\Pi(D_n) \subset E$ will contain all elements of the form

$$\int_0^t S(t, s, u)\{f(s, y(s, u), v(s)) - f(s, y(s, u), u(s))\}ds\,,$$

for every $v \in V(0, t; U)$, that is, the reachable space of the linearized system

$$z'(s) = (A(\beta) + \partial_y f(t, y(s, u), u(s)))z(s) + g(t) \quad (0 \le s \le t) \tag{5.4}$$

$$z(0) = 0 \tag{5.5}$$

where the class of admissible controls consists of all g of the form

$$g(t) = f(t, y(t, u), v(s)) - f(t, y(t, u), u(t)).$$

However, due to the smoothing properties of parabolic equations, this reachable set is very "thin" in the space $C(\Omega)$ or, for that matter, in any space $L^p(\Omega)$; typically, it will be contained in the domain of some fractional power of the infinitesimal generator $A(\beta)$ (see [FA1] for the linear case). Thus, one must rely on $C_Y(y^n) \cap B(0, \rho)$ to cause the sets in (2.8) to contain a common open set. A situation where this happens is that where the target set is a convex set with nonempty interior, (or more generally a set satisfying an open cone condition) or a $C^{(1)}$ manifold of finite codimension.

The treatment of the time optimal problem is similar: the maximum principle is (5.1) with $\mu = 0$.

6. Final remarks. In view of the observations in the previous section, the point target case $Y = \{y\}$ is not amenable to the treatment. It can be studied in other ways, roughly speaking working in the domain of $A(\beta)$ rather than in the whole space E. This has been done in [FA1] in the linear case and in [FA2] in the semilinear case (see also the references in [FA2]). Some of the results can be extended to the present setting, but the treatment of the time optimal problem only extends in spaces $L^p(\Omega)$ for $1 < p < \infty$, since the proof of the maximum principle depends on L^p estimations on the derivative $y'(t)$ of the solution of the abstract differential equation $y'(t) = Ay(t) + f(t)$ (A the infinitesimal generator of an analytic semigroup) in terms of the L^p norm of $f(t)$. These estimations depend in turn on results for vector valued singular integrals [BO], [BU], [DV], which require conditions verified in L^p spaces for $1 < p < \infty$ but not in such spaces as $C(K)$. The same observation holds for results concerning invariance of the Hamiltonian proved in [FA5] in Hilbert spaces.

The methods in this paper can be used with minor modifications for the treatment of the equation (3.1) in the space $L^1(\Omega)$; as for the equation (5.2), the role of $C(\Omega)$ is played by $L^1(\Omega)$ and the role of $L^\infty(\Omega)$ is played by $\Sigma(K)$. The nonlinear term may take values in $\Sigma(K)$ but the solutions take values in $L^1(\Omega)$. This sort of setting is natural when modelling diffusion processes and allows for control terms such as as the "travelling delta" $u(t)\delta(x - x(t))$.

References

[AE] J-P. AUBIN and I. EKELAND, **Applied Nonlinear Analysis,** Wiley- Interscience, New York (1984)

[BO] J. BOURGAIN, Some remarks on Banach spaces in which martingale difference sequences are unconditional, Ark. Mat. 21 (1983) 163-168

[BU] D. L. BURKHOLDER, A geometric characterization of Banach spaces that implies the existence of certain singular integrals of Banach–space valued functions, *Conference on Harmonic Analysis in honor of A. Zygmund,* Wadsworth (1983) 270–286

[C] F. CLARKE, **Optimization and Nonsmooth Analysis,** Wiley - Interscience, New York (1983)

[DV] G. DORE and A. VENNI, On the closedness of the sum of two closed operators, Math. Zeitschrift 196 (1987) 189-201

[E] I. EKELAND, Nonconvex minimization problems, Bull. Amer. Math. Soc. 1 (NS) (1979) 443-474

[FA1] H. O. FATTORINI, The time optimal control problem in Banach spaces, Appl. Math. Optimization 1 (1974/75) 163-188

[FA2] H. O. FATTORINI, A unified theory of necessary conditions for nonlinear nonconvex control systems, Applied Math. Optim. 15 (1987) 141-185

[FA3] H. O. FATTORINI, Optimal control of nonlinear systems: convergence of suboptimal controls, I, Lecture Notes in Pure and Applied Mathematics vol. 108, Marcel Dekker, New York (1987) 159-199

[FA4] H. O. FATTORINI, Optimal control of nonlinear systems: convergence of suboptimal controls, II, Springer Lecture Notes in Control and Information Sciences vol. 97, Berlin (1987) 230-246

[FA5] H. O. FATTORINI, Constancy of the Hamiltonian in infinite dimensional systems, to appear in Proceedings of *4th. International Conference on Control of Distributed Parameter Systems,* Vorau, July 1988

[FA6] H. O. FATTORINI, Optimal control problems in Banach spaces, to appear.

[FA7] H. O. FATTORINI, Existence and the maximum principle for relaxed solutions of control problems in infinite dimensional spaces, to appear.

[FF1] H. O. FATTORINI and H. FRANKOWSKA, Necessary conditions for infinite dimensional control problems, Proceeedings of *8th. International Conference on Analysis and Optimization of Systems,* Antibes-Juan Les Pins, June 1988

[FF2] H. O. FATTORINI and H. FRANKOWSKA, Necessary conditions for infinite dimensional control problems, to appear in Mathematics of Control, Signals and Systems

[FF3] H. O. FATTORINI and H. FRANKOWSKA, Explicit convergence estimates for suboptimal controls I, II, to appear.

[FF4] H. O. FATTORINI AND H. FRANKOWSKA, Infinite dimensional control problems with state constraints, to appear in *Proceedings of IFIP-IIASA Conference on Modelling and Inverse Problems of Control for Distributed Parameter Systems,* Laxenburg, July 1989

[FR1] H. FRANKOWSKA, A general multiplier rule for infinite dimensional optimization problems with constraints, to appear.

[FR2] H. FRANKOWSKA, Some inverse mapping theorems, to appear.

[P] L. S. PONTRYAGIN, V. G. BOLTYANSKII, R. V. GAMKRELIDZE and E. F. MISCHENKO, **The Mathematical Theory of Optimal Processes** (Russian), Goztekhizdat, Moscow (1961)