THE MAXIMUM PRINCIPLE FOR LINEAR INFINITE DIMENSIONAL CONTROL SYSTEMS WITH STATE CONSTRAINTS

H.O. Fattorini
Department of Mathematics
University of California
Los Angeles, California 90024

Abstract. We prove a version of Pontryagin’s maximum principle for linear infinite dimensional control systems (including point target conditions and state constraints). This result covers some examples for which no nonlinear theory is available at present.

1. Introduction. Let $E$ be a Banach space, $A$ the infinitesimal generator of a strongly continuous semigroup $S(t), t \geq 0$. In recent years, optimal control problems for semilinear control systems $y' = Ay + f(t,y,u(t))$ have been the object of a number of papers; these include the treatment of control constraints $u(t) \in U$, state constraints $y(t) \in M$ and target conditions $y(t) \in Y$ at the end of the trajectory. See [22], [23] for results, bibliography, and information on other ways to approach these problems that do not use semigroup theory. A common feature of all theories is that point target conditions $Y = \{y\}$ are difficult to treat except in exceptional cases.

We present in this paper a linear-convex theory (for the system $y' = Ay + Bu(t)$ with $A$ the infinitesimal generator of a compact analytic semigroup) that can handle some interesting problems with state constraints for abstract parabolic systems with point targets; nonlinear perturbations of these problems are not amenable at present to nonlinear theories (see Remark 10.2 for an explanation of this last point). The theory is expounded in detail for time optimal problems, but it can handle as well other cost functionals (see Section 9). Applications to various distributed parameter systems are presented in Sections 6 and 8.

2. Preliminaries and existence results. The system is

$$y'(t) = Ay(t) + Bu(t), \quad y(0) = \zeta,$$

with $A$ the infinitesimal generator of a compact, analytic semigroup $S(\cdot)$ in a Banach space $E$ and $B \in (F,E) = \{\text{space of all linear bounded operators from } F \text{ into } E\}$. In order to deal with parabolic equations in the most interesting cases - $L^1$ spaces and spaces of continuous functions - we use the theory of Phillips or $\odot$-adjoints. Recall that if $S(\cdot)$ is an arbitrary strongly continuous semigroup in a Banach space $E$ its adjoint semigroup $S(\cdot)^*$ may not be strongly continuous in $E$ if $E$ is not reflexive. This motivates [26] the introduction of the Phillips adjoint semigroup $S^\odot(t)$, the restriction of $S(t)^*$ to the subspace $E^\odot \subseteq E^*$ whose elements make

---

This research was partly supported by the National Science Foundation under grant DMS-9221819

Keywords: distributed parameter systems, maximum principle, optimal controls, state constraints.
\( S(t)^* y \) continuous in \( t \geq 0 \). The space \( E^\circ \) coincides with the closure of \( D(A^*) \) in \( E^* \) and \( A^* \), the infinitesimal generator of \( S^\circ(\cdot) \) is the restriction of \( A^* \) to \( E^\circ \) with domain \( D(A^\circ) = \{ y^* \in D(A^*); A^* y \in E^\circ \} \) (see [26, Chapter XIV] or [18, Chapter 2] for complete details). The space \( E^\circ \) is determining in the sense that the norm \( \|y^*\| = \sup_{z \in E^\circ} |\langle z, y \rangle| \) is equivalent to the original norm of \( E \); in particular \( \langle z, y \rangle = 0 \) for all \( z \in E^\circ \) implies \( y = 0 \) (\( \langle \cdot, \cdot \rangle \) is the duality pairing of \( E \) and \( E^* \)).

When \( S(\cdot) \) is continuous in the norm of \( (E, E) \) for \( t > 0 \) \( S(t)^* \) is continuous in the norm of \((E^*, E^*)\) for \( t > 0 \). Moreover

\[
S(t)^* E^* \subseteq E^\circ, \quad S(t)^*|_{E^\circ} = S^\circ(t) \quad (t > 0).
\]

Application of the \( \circ \)-adjoint theory to \( S^\circ(t) \) produces the space \( E^{\circ\circ} = (E^\circ)^\circ \), the semigroup \( S^{\circ\circ}(t) = (S^\circ)^\circ(t) \) and its infinitesimal generator \( A^{\circ\circ} = (A^\circ)^\circ \).

The space \( E \) is \( \circ \)-reflexive if

\[
E^{\circ\circ} = E
\]

and this property implies

\[
S^{\circ\circ}(t) = S(t), \quad A^{\circ\circ} = A.
\]

Application of (2.2) to \( S^\circ(t) \) and \( A^\circ \) produces

\[
S^\circ(t)^* (E^\circ)^* \subseteq E, \quad S^\circ(t)^*|_{E} = S(t) \quad (t > 0).
\]

For existence reasons we assume

\[
F = X^*
\]

with \( X \) a separable Banach space. \( B \) is a linear bounded operator from \( X^* \) into \( E \) and

(a) \( B^*: E^\circ \to X \),

(b) \( E \) is \( \circ \)-reflexive and \( E, E^\circ \) are separable.

The state and control constraint for (2.1) are

\[
y(t) \in M \subseteq E, \quad u(t) \in U \subseteq F = X^*
\]

respectively. Existence results below are based on

**Theorem 2.1.** \( L^1(0, T; X)^* = L^\infty_w(0, T; X^*) \).

The space \( L^\infty_w(0, T; X^*) \) consists of all \( X^* \)-valued functions \( u(\cdot) \) such that \( \langle u(\cdot), x \rangle \) is measurable and a.e. bounded for all \( x \in X \) ("a. e." depending on \( x \)). If \( X \) is separable, elements of \( L^\infty_w(0, T; X^*) \) have measurable norm (although they may not be strongly measurable) and the norm of \( L^\infty_w(0, T; X^*) \) is the essential supremum norm.

The duality pairing of \( L^1(0, T; E) \) and \( L^\infty_w(0, T; X^*) \) is \( \langle u(\cdot), f(\cdot) \rangle = \int \langle u(t), f(t) \rangle dt \) (the proof of Theorem 2.1 follows from the Dunford-Pettis theorem and is essentially contained in [11]; see [21] for other references). If \( X^* \) itself is separable, then

\[
L^\infty_w(0, T; X^*) = L^\infty(0, T; X^*).
\]

Given an arbitrary Banach space, we denote below by \( C(0, T; U) \) the space of all continuous \( E \)-valued functions \( f(\cdot) \) defined in \( 0 \leq t \leq T \) equipped with the supremum norm and by \( \Sigma(0, T; E^*) \) the space of all \( E^* \)-valued countably additive measures \( \mu \) defined in the field \( \Phi \) generated by all subintervals \([a, b]\), \((a, b), [a, b), (a, b)\). The space is equipped with the total variation norm \( \|\mu\| = \sup \Sigma|\mu(e_j)| \), supremum taken over all finite collections of pairwise disjoint elements of \( \Phi \). With this norm, \( \Sigma(0, T; E^*) \) becomes a Banach space and
Theorem 2.2. The space $\Sigma(0, T; E^*)$ is isometrically isomorphic to $C(0, T; E)^*$. The duality pairing is given by $\langle \mu, f \rangle = \int (\mu(dt), f(t))dt$.

For a proof and references see [21, §4]. The integral is defined as (a particular case of) the Bartle integral [3]. For general information on vector valued measures and integrals, see [10].

Given $\zeta \in E$ and $u(\cdot) \in L^\infty_{\text{w}}(0, T; X^*)$ the solution of (2.1) is

$$y(t) = S(t)\zeta + \int_0^t S^\circ(t - \tau)^*Bu(\tau)d\tau. \quad (2.9)$$

If $\bar{z} \in E^\circ$ then $\langle \bar{z}, Bu(\tau) \rangle = \langle B^*\bar{z}, u(\tau) \rangle$, so that $Bu(\cdot)$ is a $E^\circ$-weakly measurable function; then the integrand is a strongly measurable $E$-valued function ([16, Lemma 6.1] or [20, Lemma 4.1]). The integral in (2.9) is a Lebesgue-Bochner integral and defines a continuous $E$-valued function in $0 \leq t \leq T$, so that trajectories $y(t, \zeta, u)$ are continuous in $0 \leq t \leq T$.

The admissible control space $C_{\text{ad}}(0, \bar{t}; U)$ is the subset of $L^\infty_{\text{w}}(0, T; X^*)$ whose elements satisfy the control constraint (2.7) a.e. The reachable set $R(\bar{t}, \zeta; U)$ consists of the instantaneous values $y(\bar{t}, \zeta; u)$ of trajectories for all $u(\cdot) \in C_{\text{ad}}(0, T; U)$. We denote by $R(\bar{t}, \zeta; U) \subseteq C(0, \bar{t}; E) \times E$ the set of all elements $(y(\cdot, \zeta, u), y(\bar{t}, \zeta, u)) \in C(0, \bar{t}; E) \times E$ for all $u(\cdot) \in C_{\text{ad}}(0, \bar{t}; U)$.

In what follows $K$ is a linear operator with dense domain $D(K)$ and range in $E$. We assume that $K$ has a bounded everywhere defined inverse and that it commutes with $S(t)$, that is, $S(t)E \subseteq D(K)$ and $KS(t) \supseteq S(t)K$; equivalently, $S(t)K^{-1} = K^{-1}S(t)$. If $0 < h < t$ we have $KS(t) = S(t-h)KS(h)$, hence $KS(\cdot)$ is continuous in the norm of $(E, E)$ in $t > 0$. Since $K^{-1}$ is bounded and everywhere defined $K^*(K^{-1})^* = I$; on the other hand, $(K^{-1})^*K^* \subseteq I$, so that, although $K^*$ may not be densely defined, $(K^{-1})^* = (K^*)^{-1}$. Taking adjoints we obtain $S(t)^*K^*(K^{-1})^* = (K^*)^{-1}S(t)^*$, thus $S(t)^*E^* \subseteq D(K^*)$. $K^*S(t)^* \supseteq S(t)^*K^*$ and $K^*S(t)^*$ is continuous in the norm of $(E^*, E^*)$ for $t > 0$.

The space $E^*(K)$ consists of the completion of $E^*$ with respect to the norm $\|K^{-1}\|_E^* \bar{z}$; obviously we have the embedding $E^* \hookrightarrow E^*(K)$ and, almost by definition, $(K^{-1})^* : E^*(K) \rightarrow E^*$ is an isometric isomorphism.

Lemma 2.3. $S(t)^*$ can be extended to $E^*(K)$ and $S(t)^*E^*(K) \subseteq E^\circ$ for $t > 0$.

Proof. The extension is

$$S(t)^*z = K^*S(t)^*(K^{-1})^*z. \quad (2.10)$$

If $t > h > 0$ then $K^*S(t)^*(K^{-1})^*z = S(t-h)^*K^*S(h)^*(K^{-1})^*z$, so the fact that $S(t)^*E^*(K) \subseteq E^\circ$ follows from (2.2).

We need the backwards adjoint equation

$$dz(t) = -A^*z(t)dt - \mu(dt), \quad z(\bar{t}) = z \quad (2.11)$$

in $0 \leq t \leq \bar{t}$ with a forcing term $\mu$ in $\Sigma(0, \bar{t}; E^*)$ and $z \in E^*(K)$. By definition, the solution of (2.11) is

$$z(s) = S(\bar{t} - s)^*z + \int_s^{\bar{t}} S(\sigma - s)^*\mu(d\sigma) = z_h(s) + z_1(s). \quad (2.12)$$
The first term has been given sense in Lemma 2.3. The integral is understood as follows: for each $s$, $z_i(s)$ is the unique element of $E^\ast$ such that

$$
(z_i(s), y) = \int_s^t \langle \mu(d\sigma), S(\sigma - s)y \rangle \quad (y \in E) .
$$

(2.13)

endpoints included in the integral. We have $\|z_i(s)\| \leq C\|\mu\|\Sigma(0,T;E^\ast)$ ($0 \leq s \leq t$).

Moreover,

**Lemma 2.4.** Let $\mu \in \Sigma(0,T;E^\ast)$. Then $z_i(s)$ is a $E$-weakly left continuous $E^\ast$-valued function in $0 \leq s \leq t$.

For the proof (of a more general result) see [22, Theorem 4.3].

We note that the “final condition” $z(t) = z''$ in (2.11) should not be taken seriously, even as a left-sided weak limit; if $\mu$ has mass at $t$, the limit will not equal $z$.

**Lemma 2.5.** $z_i(s)$ is a $E$-weakly left continuous, strongly measurable $E^\ast$-valued function in $0 \leq t \leq T$. There exists a finite or countable sequence $\{s_n\} \subseteq [0,T]$ such that $z(s) \in E^\circ$ except for $s = s_n$.

**Proof.** That $z(\cdot)$ is $E$-weakly left continuous was proved in Lemma 2.4. If we prove that $z(s)$ takes values in $E^\circ$ almost everywhere, then since $E^\circ$ is separable, $z(\cdot)$ is almost separably valued, thus strongly measurable ([25]). Since the measure $\mu$ is countably additive in $\Phi$, it cannot have more than a countable number of atoms $s_j$ (points where $\mu(s_j) \neq 0$). Assuming $s$ is not an atom, write

$$
z_i(s) = \int_{[s,s+\varepsilon]} S(\sigma - s)^\ast \mu(d\sigma)
+ \int_{[s+\varepsilon,t]} S(\sigma - s)^\ast \mu(d\sigma) = z_{0,i,\varepsilon}(s) + z_{i,\varepsilon}(s).
$$

We have

$$
\langle S(h)^\ast z_{i,\varepsilon}(s) - z_{i,\varepsilon}(s), y \rangle = \langle z_{i,\varepsilon}(s), S(h)y - y \rangle
= \int_{s+\varepsilon}^t \langle \mu(d\sigma), (S(\sigma - s + h) - S(\sigma - s))y \rangle,
$$

so that

$$
|\langle S(h)^\ast z_{i}(s) - z_{i}(s), y \rangle|
\leq \|y\|_E \|\mu\|_{\Sigma(0,t;E)} \max_{s+\varepsilon \leq \sigma \leq t} \|S(\sigma - s + h) - S(\sigma - s)\|_{(E,E)} .
$$

(2.14)

In view of the uniform continuity of $S(t)$ for $t > 0$, the max on the right side tends to zero as $h \to 0$, so that $S(\cdot)^\ast z_{i,\varepsilon}(s)$ is continuous at $t = 0$, proving that $z_{i,\varepsilon}(s) \in E^\circ$.

On the other hand, estimating in a similar fashion,

$$
|\langle z_{0,i,\varepsilon}(s), y \rangle| \leq \|y\|_E \int_{[s,s+\varepsilon]} \|\mu(d\sigma)\|_{E^\ast},
$$

which shows that $\|z_{0,i,\varepsilon}(s)\|_{E^\ast} \to 0$ as $\varepsilon \to 0$; in other words, $\|z_{i,\varepsilon}(s) - z_{i}(s)\|_{E^\ast} \to 0$.

Since $z_{i,\varepsilon}(s) \in E^\circ$ and $E^\circ$ is closed, $z_i(s) \in E^\circ$ and the proof is finished.
Lemma 2.6. Let $g(\cdot)$ be a $E^\otimes$-weakly measurable $(E^\otimes)^*$-valued essentially bounded function and let $\mu \in \Sigma(0, T; E^*)$. Then
\[
\int_0^t \langle \mu(dt), \int_0^t S^\otimes(t-s)^*g(s)ds \rangle = \int_0^t \left( \int_s^t S(t-s)^*\mu(dt), g(s) \right) ds. \tag{2.15}
\]

Proof. If $g(\cdot)$ is a strongly continuous essentially bounded function with values in $E$ then Lemma 2.6 is [22, Lemma 4.3]. Under the conditions asked on $g(\cdot)$ in Lemma 2.6 $S^\otimes(h)^*g(\cdot)$ is strongly measurable if $h > 0$ ([16, Lemma 3.1] or [20, Lemma 4.1]) so that we have
\[
\int_0^t \langle \mu(dt), S(h) \int_0^t S^\otimes(t-s)^*g(s)ds \rangle = \int_0^t \langle \mu(dt), \int_0^t S(t-s)S^\otimes(h)^*g(s)ds \rangle = \int_0^t \left( \int_s^t S(t-s)^*\mu(dt), S^\otimes(h)^*g(s)ds \right) = \int_0^t \left( S^\otimes(h) \int_s^t S(t-s)^*\mu(dt), g(s) \right) ds.
\]

where in the last equality we have used the fact that $\int S(t-s)^*\mu(dt)$ takes values in $E^\otimes$ except perhaps in a countable set (thus almost everywhere). We take now limits under the integrals as $h \to 0$. On the right we use the dominated convergence theorem. On the left we use the fact that $h(t) = \int S^\otimes(t-s)^*g(s)ds$ is continuous in $0 \leq t \leq \bar{t}$, hence its range is compact, thus $S(h)h(t) \to h(t)$ uniformly in $0 \leq t \leq \bar{t}$.

Lemma 2.7. The operator
\[
Au(t) = \int_0^t S^\otimes(t-\tau)^*Bu(\tau)d\tau \tag{2.16}
\]
from $L^\infty_w(0, T; X^*)$ into $C(0, T; E)$ is compact.

For a proof see [21, Theorem 9.5]. We look at the time optimal problem with target condition
\[
y(t, \zeta, u) \in Y \subseteq E. \tag{2.17}
\]
Recall that, since $X$ is separable $L^1(0, T; X)$ is separable as well, thus the $L^1(0, T; X)$-weak topology of $L^\infty_w(0, T; X^*)$ is given by a metric; we don’t need generalized sequences.

We add to (a) and (b) the assumption (c): The admissible control space $C_{ad}(0, T; U)$ is $L^1(0, T; X)$-weakly compact in $L^\infty_w(0, T; X^*)$.

Hypothesis (c) is satisfied, for instance if $U$ is the unit ball of $X^*$; in this case $C_{ad}(0, T; U)$ is the unit ball of $L^\infty_w(0, T; X^*)$, which is $L^1(0, T; X)$-weakly compact by Alaoglu’s Theorem. It is easy to see that Hypothesis (c) implies automatically that $U$ is convex and closed which, in turn, implies convexity of $R(t, \zeta; U)$. 

THE MAXIMUM PRINCIPLE WITH STATE CONSTRAINTS
Theorem 2.8. Assume that the state constraint set $M$ and the target set $Y$ are closed. Then, if there exists a control driving $\zeta$ to $Y$ for some $t > 0$ and satisfying the state constraints, the time optimal problem has a solution.

The proof is standard. The hypotheses imply that a minimizing sequence $\{u_n(\cdot)\}$, $u_n(\cdot) \in C_{ad}(0, t_n; U)$ exists, where $t_n \geq \bar{t} = \text{optimal time}$. Extend $u_n(\cdot)$ (with the same name) to $[0, T]$ ($T = t_1$) setting $u_n(t) = u = \text{fixed element of } U$ in $t \geq t_n$. Selecting if necessary a subsequence and using Lemma 2.7 we may assume that $y(\cdot)$ can be identified with the space $E$ with the norm $\|y\|_{D(K)} = \|Ky\|$, which makes it a Banach space. The dual $D(K)^*$ can be identified with the space $E^*(K)$, the duality pairing given by

$$\langle z, y \rangle_K = \langle (K^{-1})^*z, Ky \rangle.$$  \hfill (3.1)

In fact, (3.1) defines a bounded linear functional in $D(K)$ for any $z \in E^*(K)$. Conversely, if $y \rightarrow \phi(y)$ is a bounded linear functional in $D(K)$ then $y \rightarrow \phi(K^{-1}y)$ is a bounded linear functional in $E$, hence there exists $w \in E^*$ with $\phi(K^{-1}y) = \langle w, y \rangle$; since $(K^{-1})^* : E^*(K) \rightarrow E^*$ is an isometric isomorphism, $w = (K^{-1})^*z$, and $\phi(y) = \langle (K^{-1})^*z, Ky \rangle$. We see easily that the identification of $D(K)^*$ is an algebraic and metric isomorphism.

The space $E^*_1(K, B) \subseteq E^*(K)$ consists of all $z \in E^*(K)$ such that

$$\int_0^1 \|B^*S(t)^*z\| dt < \infty.$$  \hfill (3.2)

Given a closed convex set $Y$ in a Banach space $E$ and an element $y \in Y$ the tangent cone $T_Y(y)$ to $Y$ at $y$ is

$$T_Y(y) = \bigcup_{\lambda \geq 0} \lambda(Y - y).$$

The normal cone $N_Y(y) \subseteq E^*$ to $Y$ at $y$ is

$$N_Y(y) = \{ z \in E^*; \langle z, y \rangle \leq 0 \quad (y \in T_Y(y)) \}.$$  \hfill  

If $M$ is the state constraint set, $M(\bar{t}) \subseteq C(0, \bar{t}; E)$ consists of all $y(\cdot)$ with $y(t) \in M$ ($0 \leq t \leq \bar{t}$).

The results in this section are on the set target condition

$$y(\bar{t}, \zeta, u) \in Y \subseteq D(K)$$  \hfill (3.3)

(in particular, on the point target condition $y(\bar{t}, \zeta, u) = \bar{y} \in D(K)$). Assumptions (a), (b) and (c) in last section are in force.
Theorem 3.1. Let \( \tilde{u}(\cdot) \) be a time optimal control with optimal time \( \tilde{t} \). Assume that \( U, M \) and \( Y \) are convex and closed with \( \text{Int}(U) \neq \emptyset \) and \( \text{Int}(M) \neq \emptyset \) in \( E \), and that
\[
R(\tilde{t}, \zeta; U) - Y
\]
contains a ball in \( D(K) \) for some \( \tilde{t} < \hat{t} \). Then there exists \( (\mu, z) \in \Sigma(0, \tilde{t}; E^*) \times \mathcal{E}_1^*(K, B), (\mu, z) \neq 0 \) with \( \mu \in N_{M(\tau)}(y(\cdot, \zeta, \tilde{u})), z \in N_Y(y(\tilde{t}, \zeta, \tilde{u})) \) (the latter normal cone defined according to the duality (3.1)) and such that
\[
\langle B^* \tilde{z}(s), \tilde{u}(s) \rangle = \min_{v \in U} \langle B^* \tilde{z}(s), v \rangle
\]
a.e. in \( 0 \leq s \leq \tilde{t} \), \( \tilde{z}(\cdot) \) the solution of (2.11) for \( \mu, z \).

In Theorem 3.1, “\( Y \) is closed” means “\( Y \) is closed in \( E^* \).” The proof follows from a chain of auxiliary results.

Lemma 3.2. Let \( E \) be a Banach space, \( \{ \Delta_n \} \) a sequence of sets in \( E \), \( \{ z_n \} \) a sequence in \( E^* \) such that
\[
0 < c \leq \| z_n \| \leq C, \quad \langle z_n, y \rangle \leq \varepsilon_n \to 0 \quad (y \in \delta_n).
\]
Assume that the \( \Delta_n \) contain a common ball for \( n \) large enough. Then every weakly convergent subsequence of \( \{ z_n \} \) has a nonzero limit.

Lemma 3.3. Let \( E \) be a Banach space, \( Q \subseteq E \) a compact subset, \( \{ z_n \} \subseteq E^* \) a bounded sequence with \( z_n \to 0 \) weakly. Then \( \langle z_n, y \rangle \to 0 \) uniformly for \( y \in Q \).

The proof of Lemma 3.3 follows covering \( Q \) with a finite number of balls of arbitrarily small diameter.

Lemma 3.4. Assume \( R(\tilde{t}, \zeta; U) - Y \) contains a ball in \( D(K) \) for \( \tilde{t} < \hat{t} \). Then, if \( \{ t_n \} \) is a sequence with \( t_n \to \tilde{t} \), the sets \( R(t_n, \zeta; U) - Y \) contain a common ball for \( n \) large enough.

Proof. Let \( u(\cdot) \in C_{ad}(0, t; U) \). Then
\[
\begin{align*}
y(t, \zeta, u) &= S(t)\zeta + \int_0^t S^\circ(t - \tau)^* Bu(\tau) d\tau \\
&= (S(t) - S(t'))\zeta + S(t')\zeta + \int_{t'}^t S^\circ(t' - \tau)^* Bu(\tau + (t' - t)) d\tau.
\end{align*}
\]
Since the function \( \tilde{u}(\tau) = 0 \) (0 \( \leq \tau \leq t' - t \), \( \tilde{u}(\tau) = u(\tau + (t' - t)) \) belongs to \( C_{ad}(0, t'; U) \), we obtain
\[
R(t, \zeta; U) \subseteq (S(t) - S(t'))\zeta + R(t', \zeta; U) \quad (t < t'),
\]
hence
\[ R(t, \zeta; U) - Y \subseteq (S(t) - S(t'))\zeta + (R(t', \zeta; U) - Y) \quad (t < t'), \]
and the result follows from continuity of $KS(t)$ in $t > 0$.

**Proof of Theorem 3.1.** Given $\tilde{t} > 0$, $M_K(\tilde{t}, Y) \subseteq C(0, \tilde{t}; E) \times E$ consists of all elements of the form

\[ (y(\cdot), Ky) \quad (y(\cdot) \in M(\tilde{t}), y \in Y). \]

We denote by $C_{ad,K}(0, \tilde{t}; U)$ the set of all $u(\cdot) \in C_{ad}(0, \tilde{t}; U)$ with $y(\tilde{t}, \zeta, u) \in D(K)$ and by $R_K(\tilde{t}, \zeta; U) \subseteq C(0, \tilde{t}; E)$ the set of all elements of the form

\[ (y(\cdot, \zeta, u), Ky(\tilde{t}, \zeta, u)) \quad (u(\cdot) \in C_{ad,K}(0, \tilde{t}, U)). \]

Select a sequence $\{t_n\} \subseteq [0, \tilde{t}]$, $\tilde{t} = \text{optimal time}$, $t_n \to \tilde{t}$. Due to optimality of $\tilde{t}$, the sets $M_K(t_n, Y)$ and $R_K(t_n, \zeta; U)$ are disjoint. More is true: if $t < \tilde{t}$ the distance from $M_K(\tilde{t}, Y)$ to $R_K(\tilde{t}, \zeta; U)$ is positive. In fact, if this were not the case we could construct a sequence $\{u_m(\cdot)\} \subseteq C_{ad,K}(0, \tilde{t}; U)$ and a sequence $y_m(\cdot) \in M(\tilde{t})$ such that

\[ y(t, \zeta, u_m) - y_m(t) \to 0 \quad (3.10) \]

uniformly in $0 \leq t \leq \tilde{t}$ and

\[ Ky(\tilde{t}, \zeta, u_m) - Ky_m \to 0. \quad (3.11) \]

as $m \to \infty$, where $\{y_m\} \in Y$. Using Alaoglu’s Theorem and Lemma 2.7 we can select a subsequence of $\{u_m\}$ such that $u_m(\cdot) \to \bar{u}(\cdot) L^1(0, t; E)$-weakly in $L^\infty(0, t; X^*)$ and $y(\cdot, \zeta, u_m)$ is uniformly convergent in $0 \leq t \leq \tilde{t}$; in particular $y(\tilde{t}, \zeta, u_m) \to \bar{y} \in E$. Applying $K^{-1}$ to both sides of $(3.11)$ we deduce that $y_m \to \bar{y} \in Y$ as well, so that $y(\tilde{t}, \zeta, \bar{u}) = \bar{y} \in Y$. On the other hand, $(3.10)$ implies $\text{dist}(y(t, \zeta, u_m), M) \to 0$ as $m \to \infty$ in $0 \leq t \leq \tilde{t}$, thus $y(t, \zeta, \bar{u})$ satisfies the state constraint and goes from $\zeta$ to $\bar{y} \in Y$ in time $\bar{t} < \tilde{t}$, which contradicts the definition of $\tilde{t}$ as optimal time.

Since $\text{dist}(M_K(t_n, Y), R_K(t_n, \zeta; U)) > 0$, $M_K(t_n, Y)$ and $R_K(t_n, \zeta; U)$ will be disjoint for $\varepsilon > 0$ sufficiently small, where, for any set $A, A_\varepsilon = \{y; \text{dist}(y, A) < \varepsilon\}$. If $A$ is convex $A_\varepsilon$ is convex with interior points, thus we can apply the separation theorem and construct $(\mu_n, w_n) \in \Sigma(0, t_n; E^*) \times E^*, \|\langle \mu, w_n \rangle \| = 1$ such that

\[ \langle (\mu_n, w_n), (y(\cdot), Ky) \rangle \leq \langle (\mu_n, w_n), (y(\cdot, \zeta, u), Ky(t_n, \zeta, u)) \rangle \quad (3.12) \]

or

\[ \int_0^{t_n} \langle \mu_n(dt), y(t) \rangle + \langle w_n, Ky \rangle \quad (3.13) \]

\[ \leq \int_0^{t_n} \langle \mu_n(dt), y(t, \zeta, u) \rangle + \langle w_n, Ky(t_n, \zeta, u) \rangle \]

for every $y(\cdot) \in M(t_n), y \in Y$ and $u(\cdot) \in C_{ad,K}(0, t_n; U)$. We extend $\mu_n$ to $[0, \tilde{t}]$ setting $\mu_n = 0$ there and, using Alaoglu’s theorem, select a subsequence such that $(\mu_n, w_n) \to (\mu, w) \in \Sigma(0, \tilde{t}; E^*) \times E^* C(0, \tilde{t}; E) \times E$-weakly. Assume $(\mu, w) = 0$. 

H.O. FATTORINI
By Lemma 2.7 the set \( \{y(\cdot, \zeta, u); u \in C_{ad}(0, \bar{t}; U) \} \) is compact in \( C(0, \bar{t}; E) \), thus the integral on the right side of (3.13) tends to zero uniformly in \( u \). Sending left the second term on the right of (3.13) we obtain

\[
\int_{0}^{t_n} \langle \mu_n(dt), y(t) \rangle - \langle w_n, y \rangle \leq \varepsilon_n \to 0 \tag{3.14}
\]

for \( \langle y(\cdot), y \rangle \in \mathbf{M}(\bar{t}) \times KB_K(\tilde{y}, \delta), B_K(\tilde{y}, \delta) \) a ball in \( D(K) \) contained in all \( R(t_n, \zeta; U) - Y \). Hence we obtain a contradiction using Lemma 3.2, and we conclude that \( \langle \mu, w \rangle \neq 0 \).

Inequality (3.13) implies in particular that

\[
|\langle w_n, Ky(t_n, \zeta, u) \rangle| \leq C \tag{3.15}
\]

for all \( u(\cdot) \in C_{ad,K}(0, t_n; B(u_0, \rho)) \) (\( B(u_0, \rho) \) a ball contained in \( U \), thus for all \( u(\cdot) \in C_{ad}(0, t_n - \delta; B(u_0, \rho)) \) (\( \delta > 0 \)), since \( y(t_n, \zeta, u) \subseteq S(\delta)E \subseteq D(K) \) for controls that vanish in \( t_n - \delta \leq t \leq t_n \). Subtracting the constant \( w_0 \) from elements of \( C_{ad}(0, t_n - \delta; B(u_0, \rho)) \) we deduce that

\[
\langle w_n, K \rangle \int_{0}^{t_n - \delta} S^\dagger(t_n - \tau)^* Bu(\tau) d\tau
\]

\[
= \int_{0}^{t_n - \delta} \langle B^* K^* S(t_n - \tau)^* w_n, u(\tau) \rangle d\tau
\]

is bounded for \( u(\cdot) \in C_{ad}(0, t_n - \delta; B(0, \rho)) \) independently of \( \delta > 0 \). We then approximate \( B^* K^* S(t_n - \tau)^* w_n \) in \( [0, t_n - \delta] \) by a piecewise constant function \( h(\cdot) \) and select \( u(\cdot) \in C_{ad}(0, t_n - \delta, B(0, \rho)) \) with \( \langle h(t), u(t) \rangle = \rho \|h(t)\| \). In view of the arbitrariness of \( \delta \), we obtain in this way that

\[
\int_{0}^{1} \| B^* K^* S(t)^* w_n \| dt \leq C
\]

where \( C \) does not depend on \( n = 1, 2, \ldots \). Since \( w_n \rightarrow w \) \( E \)-weakly and \( S(t)^* \) is compact, then \( S(t)^* w_n \rightarrow S(t)^* w \) in the norm of \( E^* \) \cite[Theorem 6, p. 486]{fatou} and it follows from Fatou’s theorem that

\[
\int_{0}^{1} \| B^* K^* S(t)^* w \| dt \leq C \tag{3.17}
\]

We take limits in (3.13) as \( n \to \infty \). Justification is obvious in all terms except the second on the right side. We can obviously take the limit if \( u(\cdot) \) vanishes in \( \bar{t} - \delta \leq t \leq \bar{t} \); then we write the integral term in the form (3.16), use the fact that \( B^* K^* S(t_n - \tau)^* z \in L^1(0, \bar{t}; X) \) and let \( \delta \to 0 \). The result is

\[
\int_{0}^{\bar{t}} \langle \mu(dt), y(t) \rangle + \langle w, Ky \rangle \leq \int_{0}^{\bar{t}} \langle \mu(dt), y(t, \zeta, u) \rangle + \langle w, Ky(\bar{t}, \zeta, u) \rangle
\]
which we use thrice. The first time we take \( y(t) = y(t, \zeta, \bar{u}) \) and \( y = y(\tilde{t}, \zeta, u) \), cross out \( S(\tilde{t} - t) \zeta \) in both sides and use Lemma 2.6 in both integrals. We obtain in this way

\[
\int_{0}^{\tilde{t}} \langle B^* \bar{z}(t), \bar{u}(t) \rangle dt \leq \int_{0}^{\tilde{t}} \langle B^* \bar{z}(t), u(t) \rangle dt
\]  

(3.19)

where \( \bar{z}(t) \) is the solution of (2.11) corresponding to \( \mu \) and \( z \in E^*(K) \) defined by \( w = (K^{-1})^* z \); see (2.10) and the proof of Lemma 2.3. (3.20) implies that \( z \in E^*_1(K, B) \) and, since \( (K^{-1})^* \) is an isometric isomorphism from \( E^*(K) \) into \( E^* \), \( z \neq 0 \) if \( w \neq 0 \), so that \( (\mu, z) \neq 0 \).

If \( u \in X^* \) then \( \langle B^* \bar{z}(t), u \rangle = \langle \bar{z}(t), Bu \rangle \), hence \( \bar{z}(\cdot) \) is \( X^* \)-weakly measurable in \( X \), hence strongly measurable since \( X \) is separable. On the other hand, \( \bar{u}(\cdot) \) is \( X \)-weakly measurable, so that \( \langle B^* \bar{z}(\cdot), \bar{u}(\cdot) \rangle \) is a bounded measurable function. Call \( e \) the set of all Lebesgue points of \( (b^* \bar{z}(\cdot), \bar{u}(\cdot)) \) and of \( B^* \bar{z}(\cdot) \).

Let \( u(\cdot) \in C_{ad}(0, \tilde{t}; U), \ 0 < s < \tilde{t}, \ v \in U \). The spike perturbation \( u_{s,v,h}(\cdot) \) of \( u(\cdot) \) is defined by \( u_{s,v,h}(t) = v \ (s - h \leq t \leq s), \ u_{s,v,h}(t) = u(t) \) elsewhere. Take \( s < e \) and use (3.19) with \( u = u_{s,v,h} \); the result is

\[
\frac{1}{h} \int_{s-h}^{s} \langle B^* \bar{z}(t), \bar{u}(t) \rangle dt \leq \frac{1}{h} \int_{s-h}^{s} \langle B^* \bar{z}(t), v \rangle dt
\]

so that, taking limits, (3.5) results. In the second use of (3.18) we take \( u = \bar{u} \) on both sides, obtaining

\[
\int_{0}^{\tilde{t}} \langle \mu(dt), y(t) - y(t, \zeta, \bar{u}) \rangle \leq 0
\]  

(3.20)

for all \( y(\cdot) \in X(\tilde{t}) \), so that \( \mu \in N_M(\tilde{t}; y(\cdot, \zeta, \bar{u})) \) as claimed. The third time we take \( y(t) = y(t, \zeta, \bar{u}) \) and obtain

\[
\langle w, Ky - Ky(t, \zeta, \bar{u}) \rangle \leq 0
\]

hence

\[
\langle z, y - y(t, \zeta, \bar{u}) \rangle = \langle (K^{-1})^* z, Ky - Ky(t, \zeta, \bar{u}) \rangle
\]

\[
\langle w, Ky - Ky(t, \zeta, \bar{u}) \rangle \leq 0
\]  

(3.21)

for all \( y \in Y \). This shows that \( z \in N_Y(y(\tilde{t}, \zeta, u)) \) and thus completes the proof of Theorem 3.1.

Following finite dimensional usage, we shall call \( \bar{z}(t) \) the costate corresponding to the optimal control \( u(t) \). When \( Y \) is “thin” (say, a point), the assumptions in Theorem 3.1 demand that \( R(\tilde{t}, \zeta; U) \) contain a ball in \( D(K) \). If we define \( R^\infty(\tilde{t}) \) as the subspace of \( E \) consisting of all elements \( y(\tilde{t}, 0, u), u \in L^2_w(0, T; X^*) \) the assumption implies

\[
R^\infty(\tilde{t}) \supseteq D(K).
\]  

(3.22)

We may baptize (3.22) “exact reachability to \( D(K) \) in time \( \tilde{t} \). It implies in particular that \( R^\infty(\tilde{t}) \) is dense in \( E \).

4. The “full control” system. We give this name to

\[
y'(t) = Ay(t) + u(t)
\]  

(4.1)
with $X = E^\ominus$, so that $U \subseteq (E^\ominus)^*$ and $u(\cdot) \in L^\infty_w(0,T;(E^\ominus)^*)$. Hypotheses (a) and (b) (with $B = I$) in §2 and §3 require $E$ and $E^\ominus$ separable. If hypothesis (c) holds Theorem 3.1 applies with $K = A$ (we may always assume via a translation that $A^{-1}$ exists) if $U$, $M$ are convex with nonempty interior. For the point target condition
\[ y(t, \zeta, u) = \bar{y} \in D(A) \] (4.2)
we only have to check that $R(\bar{t}, \zeta; U)$ contains a ball in $D(A)$ for any $\bar{t} > 0$, and we can limit ourselves to $\zeta = 0$. This was done in [16 p. 167] for $U = \text{unit ball}$. In the general case, we take an interior point $u_0$ of $U$ and let $u_0(t) \equiv u_0$, $y_0 = y(t, 0, u_0)$; then $y_0 \in D(A)$ with $A y_0 = (S(\bar{t}) - I) u_0$. If $y \in D(A)$ we have
\[ y = \int_0^{\bar{t}} S(\bar{t} - \tau) \frac{1}{\bar{t}} (\bar{y} - \tau A y) d\tau \] (4.3)
so that $y_0 + y = y(t_n, 0, u)$ for $u = u(\tau) = u_0 + \bar{t}^{-1}(\bar{y} - \tau A y)$. Obviously, if $\|Ay\| \leq \varepsilon$ sufficiently small, $u(\cdot) \in C_{ad}(0; t_n; U)$.

Theorem 4.1. Let $\bar{u}(\cdot)$ be a time optimal control with optimal time $\bar{t}$ with $\bar{y} \in D(A)$. Assume that $U$, $M$ are convex with nonempty interior. Then there exists $(\mu, z) \in \Sigma(0, \bar{t}; E^*) \times E^*_1(A, I), (\mu, z) \neq 0$ with $\mu \in N_{M(\bar{t})}(y(\cdot, \zeta, \bar{u}))$ and
\[ (\bar{z}(s), \bar{u}(s)) = \min_{u \in U} \{\bar{z}(s), v\} \] (4.4)
a.e. in $0 \leq s \leq \bar{t}$, $\bar{z}(\cdot)$ the solution of (2.11).

We note that there exists a version of this result (Theorem 4.1 in [16]) where $S(\cdot)$ is an arbitrary semigroup. See [16] for the exact statement and proof.

When $E$ is reflexive, $X = E^*$. A subclass of the class of reflexive spaces is that of $\zeta$-convex spaces introduced in [7], which are exactly the spaces where the Hilbert transform is a bounded operator from $L^p(-\infty, \infty; E)$ into itself for $1 < p < \infty$ [6]. It was proved in [12] that if $E$ is $\zeta$-convex and $A$ is the infinitesimal generator of an analytic semigroup such that
\[ \|(-A)^{\tau}\| \leq C e^{C|\tau|}, \quad (-\infty < \tau < \infty) \] (4.5)
then the solution $y(\cdot)$ of (4.1) with $u(\cdot) \in L^p(0, T; E)$, $1 < p < \infty$ belongs to $D(A)$ a.e. (and satisfies
\[ \|Ay\|_{L^p(0,T;E)} \leq C\|u\|_{L^p(0,T;E)}. \] (4.6)
This can be used as follows: if $\bar{u}(t)$ is an arbitrary time optimal control in $0 \leq t \leq \bar{t}$ then $\bar{u}(\cdot)$ is as well time optimal in any interval $0 \leq t \leq \bar{t} < \bar{t}$ (this is the well known optimality principle valid for time invariant equations like (4.1)). We can then take an increasing sequence $\{t_n\} \subseteq [0, \bar{t})$ with $t_n \to \bar{t}$ and such that $y_n = y(t_n, \zeta, \bar{u}) \in D(A)$ and apply Theorem 4.1 in each subinterval $0 \leq t \leq t_n$. We obtain a sequence $(\mu_n, z_n) \in \Sigma(0, t; E^*) \times E^*_1(A, I), (\mu_n, z_n) \neq 0$ such that $\mu \in N_{M(t_n)}(y(\cdot, \zeta, \bar{u}))$ and
\[ \langle \bar{z}_n(s), \bar{u}(s) \rangle = \min_{v \in U} \{\bar{z}_n(s), v\}, \] (4.7)
a.e. in $0 \leq t \leq t_n$, $\bar{z}_n(\cdot)$ the solution of (2.11) with $z_n$, $\mu_n$ (it does not seem obvious that one can “take limits as $n \to \infty$” in (4.7) in any useful way).

Hilbert spaces are $\zeta$-convex, as are $L^p(\Omega)$ spaces, $1 < p < \infty$. The spaces $L^1(\Omega)$ and $C(\Omega)$ are not $\zeta$-convex. The Hilbert space particular case of (4.6) was proved much earlier in [9] without Assumption (4.5), and is used in [16] to show a result of the type of (4.7).

5. Nontriviality of the minimum principle and saturation of state constraints. There exists the possibility that the costate $\bar{z}(t)$ in Theorem 3.1 is trivial in the sense that $B^*\bar{z}(t) \equiv 0$; this empties the maximum principle. The fact that $(\mu, z) \neq 0$ does not necessarily prevent this (for instance, take $\mu = -\delta(t - \bar{t})z, z \in E^*, z \neq 0$). We believe that nontriviality of the costate cannot be prevented in general, but it still holds in a number of particular cases. The first result has to do with this, but it also has independent interest.

**Lemma 5.1.** Assume that $M$ is a convex set with nonempty interior, and let $\bar{y}(\cdot) \in M(\bar{t}), \mu \in N_{M(\bar{t})}(\bar{y}(\cdot))$. Then, if

$$e(\bar{y}(\cdot)) = \{t \in [0, \bar{t}]; \bar{y}(t) \in \text{Int}(M)\}$$

we have $\mu = 0$ in $e(\bar{y})$.

For the proof (of a much more general theorem) see [22, Lemma 6.3]. Lemma 5.1 applied to the maximum principle implies that the measure $\mu$ vanishes on the set where the optimal trajectory $y(t, \zeta; u)$ does not saturate the state constraint (saturation means $y(t, \zeta; u)$ hits the boundary of $M$). In particular, $\mu = 0$ if the state constraint is never saturated (or if there is no state constraint; $M = E$). In this case, we obtain the “ordinary” maximum principle

$$\langle B^*S(\bar{t} - t)^*z, \bar{u}(t) \rangle = \min_{v \in U} \langle B^*S(\bar{t} - t)^*z, v \rangle$$

which is (sometimes) guaranteed to be highly nontrivial, as seen below.

**Theorem 5.2.** Let the hypotheses of Theorem 3.1 be satisfied with $Y = \{\bar{y}\}$ and let $z \in E^*(K)$. Then

$$n(\bar{t}, z) = \{t \in [0, \bar{t}]; B^*S(\bar{t} - t)^*z = 0\}$$

is either empty or consists of a sequence $\{t_n\}$; if $\{t_n\}$ is infinite, its only limit point may be $t$.

We prove first

**Lemma 5.3.** Let $S(h)^*z = 0$ for $z \in E^*(K)$. Then $z = 0$.

**Proof.** Let $w = (K^{-1})^*z \in E^*$. Then $K^*S(h)^*w = K^*S(h)^*(K^{-1})^*z = S(h)^*z = 0$. Since $K^*$ is one-to-one, $S(h)^*w = 0$; if $t \geq h$ we have $S(t)^*w = S^\circ(t - h)S(h)^*w = 0$ for $t \geq h$, thus by analyticity for all $t > 0$, and $\langle w, y \rangle = \lim_{t \to 0^+} \langle S(t)^*w, y \rangle = 0$, so that $w = 0$, which implies $z = 0$. 

Lemma 5.4 [15]. \( R^\infty(\bar t) \) is dense in \( E \) if and only if
\[
B^*S(\bar t - t)^*z = 0 \quad \text{in} \quad 0 \leq t \leq \bar t \quad \text{implies} \quad z = 0. \tag{5.4}
\]

Proof. \( R^\infty(\bar t) \) is not dense in \( E \) is and only if there exists \( z \in E, z \neq 0 \) with
\[
\langle z, y(\bar t, 0, u) \rangle = \int_0^\bar t (B^*S(\bar t - t)^*z, u(t))dt.
\]

Proof of Theorem 5.2. We show first that \( n(t, z) \neq [0, \bar t] \). Note first that if \( \bar t < \bar t \)
we have \( B^*S(\bar t - t)^*z = B^*S(\bar t - t)^*z \) for \( 0 \leq t \leq \bar t \), hence
\[
n(\bar t, z) \cap [0, \bar t] = n(\bar t, S(\bar t - t)^*z) \tag{5.5}
\]
with \( S(\bar t - t)^*z \in E^* \). The hypotheses in Theorem 3.1 imply that \( R^\infty(\bar t) \) is dense in \( E \) for some \( \bar t < \bar t \) optimal time, thus \( n(\bar t, z) = [0, \bar t] \) implies \( n(\bar t, S(\bar t - t)^*z) = [0, \bar t] \) implies \( S(\bar t - t)^*z = 0 \), thus \( z = 0 \) by Lemma 5.3. The rest of Theorem 5.2 follows from the fact that \( S(\bar t - t)^*z \) is analytic in \( t < \bar t \).

Result like Theorem 5.2 give meaning to the name switching points of an optimal control. \( \bar u(\cdot) \). To fix ideas, assume that \( F \) is a Hilbert space and that the control set \( U \) is the unit ball of \( F \). Then
\[
\bar u(t) = \frac{B^*S(\bar t - t)^*z}{\|B^*S(\bar t - t)^*z\|} \tag{5.6}
\]
in the intervals \( (t_n, t_{n+1}) \) between points of the sequence \( \{t_n\} \) in Theorem 9.3. The optimal control is then analytic between “switching points” \( t_n \).

The following result connects with the ideas in §4.

Theorem 5.5. Consider the equation (4.1) with \( U = \) unit ball of \( E \) assuming, either, that (i) \( E \) is a Hilbert space, or (ii) \( E \) is a \( \zeta \)-convex Banach space and (4.5) holds. Then if \( \bar u(\cdot) \) is a time optimal control for the point target condition (4.2), the assumptions of Theorem 3.1 are satisfied and \( e(y(\cdot, \zeta, \bar u)) \) is the (open) set (5.1) where the state constraint is not saturated, we have
\[
\|u(t)\| = 1 \quad \text{if} \quad t \in e(y(\cdot, \zeta, \bar u)). \tag{5.7}
\]
Moreover, \( u(\cdot) \) has at most a countable number of switching points in \( e(y(\cdot, \zeta, \bar u)) \) (possibly with infinitely many accumulation points) and is (as) smooth (as the norm) in between.

In fact, let \([t_0, t_1]\) be a component interval of \( e(y(\cdot, \zeta, \bar u)) \). Divide it into a (most) countable union of intervals \([t_n, t_{n+1}]\) with \( t_n < t_{n+1} \) and \( y(t_{n+1}, \zeta, u) \in D(A) \) and apply Theorem 3.1 in each of these subintervals.

In the set target case \( R(\bar t, \zeta; U) - Y \) rather than \( R(\bar t, \zeta, U) \) contains a ball in \( D(K) \) and there is no analogue of Theorem 5.2, except in certain particular cases;
one is equation (4.1). In fact, it follows as in Lemma 5.3 that $A^* S(t - t)^* z \neq 0$ in $0 \leq t \leq t$.

6. Applications to parabolic distributed parameter systems. We consider the uniformly elliptic operator $A_y = \Sigma \Sigma \partial^j (a_{jk}(x) \partial^k y) + \Sigma b_j(x) \partial^j y + c(x)y$ in a domain $\Omega \subseteq \mathbb{R}^m$ with boundary $\Gamma \ (a_{jk} = a_{kj}, \ \Sigma \Sigma a_{kj}(x) \xi_j \xi_k \geq K \|\xi\|^2$ for $x \in \Omega$ and $\xi \in \mathbb{R}^m$, $a_{jk}(x)$ and $b_j$ are continuously differentiable in $\Omega$, $c$ continuous in $\Omega$). The domain $\Omega$ is bounded and of class $C^{(2)}$. The formal adjoint is $A'^y = \Sigma \Sigma \partial^j (a_{jk}(x) \partial^k y) - \Sigma \partial^j (b_j(x)y) + c(x)y$. We denote by $A_p(\beta)$, $1 \leq p < \infty$ the restriction of $A$ in $L^p(\Omega)$ corresponding to a boundary condition $\beta$, either the Dirichlet boundary condition or a variational boundary condition $\partial^\nu y(x) = \gamma(x)y(x)$ ($x \in \Gamma$) with $\gamma(x)$ continuously differentiable on $\Gamma$ and $\partial^\nu$ the conormal derivative $\Sigma \Sigma a_{jk}(x) \eta_j(x) \partial^k$, $\{\eta_j(x)\}$ the outer normal vector on $\Gamma$. The adjoint boundary condition $\beta'$ is $\beta'$ for the Dirichlet boundary condition; for a variational boundary condition, $\beta'$ is $\partial^\nu y(x) = (\gamma(x) + b(x))y(x)$ with $b(x) = \Sigma \partial^j (b_j(x)\eta_j(x))$. We also consider the operator $A_c(\beta)$ determined by $A$ and $\beta$ in the space $C(\Omega)$ of continuous functions in $\Omega$ equipped with the supremum norm; for the Dirichlet boundary condition we use $C_0(\Omega) \subseteq C(\Omega)$ consisting of all $y$ which are zero at the boundary $\Gamma$. Precise characterizations of the domains can be found for instance in [18]. The generation and duality properties of these operators can be summarized as follows. $A_p(\beta)$ generates a compact analytic semigroup $S_p(t; A, \beta)$ in $L^p(\Omega)$, $1 \leq p < \infty$ and $A_c(\beta)$ generates a compact analytic semigroup $S_c(t; A, \beta)$ in $C(\Omega)$ for the Dirichlet boundary condition. If $1 < p < \infty$ then $L^p(\Omega)^* = L^q(\Omega)$ ($1/p + 1/q = 1$) and

$$A_p(\beta)^* = A_p'(\beta'), \ S_p(t; A, \beta)^* = S_q(t; A', \beta'). \quad (6.1)$$

If $p = 1$ then $L^1(\Omega) = L^\infty(\Omega)$ and $L^1(\Omega)^\circ = C(\Omega)^\circ$ ($C_0(\Omega)$ for the Dirichlet boundary condition) and

$$A_1(\beta)^\circ = A_1'(\beta'), \ S_1(t; A, \beta)^\circ = S_1(t; A', \beta'). \quad (6.2)$$

Finally, $C(\Omega)^* = \Sigma(\Omega)$, the space of all finite regular Borel measures in $\Omega$, $C(\Omega)^\circ = L^1(\Omega)$ and

$$A_c(\beta)^\circ = A_c'(\beta'), \ S_c(t; A, \beta) = S_c(t; A', \beta'); \quad (6.3)$$

in the Dirichlet case $C_0(\Omega)^* = \Sigma_0(\Omega) \subseteq \Sigma(\Omega)$ consisting of all $\nu$ with $|\nu|(\Gamma) = 0$, $C_0(\Omega)^\circ = L^1(\Omega)$ and (6.3) holds.

We look below at the equations

$$y'(t) = A_p(\beta)y(t) + u(t), \quad y(0) = \zeta \quad (6.4)$$

in the space $L^p(\Omega)$, $1 \leq p < \infty$, and

$$y'(t) = A_c(\beta)y(t) + u(t), \quad y(0) = \zeta \quad (6.5)$$

in the space $C(\Omega)$ or $C_0(\Omega)$, depending on the boundary condition. Both equations fit into the hypotheses in §3. For $1 < p < \infty$ we take $E = L^p(\Omega)$, $X = L^q(\Omega)$ with $1/p + 1/q = 1$, so that $X^* = L^p(\Omega)$. The control set is $U = B(0, 1)$ — unit ball of $L^p(\Omega)$. 

H. O. FATTORINI
For $p = 1$, $E = L^1(\Omega)$, and we take $X = C(\bar{\Omega})$, so that $X^* = \Sigma(\bar{\Omega})$. The admissible control space is then the unit ball of $L_\infty^0(0, T; \Sigma(\Omega))$, so that we rewrite the equation as

$$
y'(t) = A_1(\beta)y(t) + \nu(t), \quad y(0) = \zeta. \tag{6.6}
$$

For (6.5) we take $E = C(\bar{\Omega})$, $X = C(\bar{\Omega})^\circ = L^1(\Omega)$, so that $X^* = L_\infty^0(\Omega)$; the control set is the unit ball of $L_\infty^0(\Omega)$. In all cases we take a closed, convex state constraint set $M$ with nonempty interior. The target condition is

$$
y(t, \zeta, u) = \bar{y} \tag{6.7}
$$

The conclusions are based on (4.4). Recalling that the duality set $\Theta$ of an element $x \in X$ consists of all $x^* \in X^*$ such that $\|x^*\|^2 = \langle x^*, x \rangle$, (4.4) implies

$$
\|B^*z(t)\|u(t) \in \Theta(B^*z(t)) \tag{6.8}
$$

if $B^*z(t)^*z \neq 0$. This is also true (but trivial) if $B^*z(t)^*z = 0$.

**Example 6.1.** If $1 < p < \infty$ we have $\Theta(y) = \|y\|^{2-p}\langle y(x)\rangle^{2p-2}y(x)$ if $y \neq 0$, thus by Theorem 3.1, if $\bar{u}(t, x)$ is a time optimal control such that the final point of the trajectory belongs to $D(A_p(\beta))$ we have

$$
\bar{u}(t, x) = \frac{|\bar{z}(t, x)|^{p-2}\bar{z}(t, x)}{\|\bar{z}(t, \cdot)|^{p-1}} \tag{6.9}
$$

where $\bar{z}(t, x)$ is the solution of

$$
d\bar{z}(t, x) = -A_1^*(\beta')\bar{z}(t, x) - \nu(dt), \quad \bar{z}(\bar{t}, \cdot) = z, \tag{6.10}
$$

with $z(\cdot) \in E^*_1(A, I) = L^q(\Omega)_1(A_p(\beta), I), \mu \in \Sigma(0, \bar{t}; L^q(\Omega)), \langle \mu, z \rangle \neq 0$. On the other hand, $E = L^p(\Omega)$ is a $\zeta$-convex Banach space and (4.5) is true [2] so Theorem 5.3 applies and we have

$$
\|\bar{u}(t, \cdot)\|_{L^p(\Omega)} = 1 \tag{6.11}
$$

for a totally arbitrary optimal control $\bar{u}(\cdot)$ (no conditions on the target) outside of the set $e(y(\cdot, \zeta, \bar{u}))$ where the state constraint is saturated.

**Example 6.2.** The system is (6.4) with $p = 1$ (or, rather, (6.6)). Given $y(\cdot) \in C(\bar{\Omega})$ the duality set $\Theta(y)$ of $y$ is the set of all measures $\nu \in \Sigma(\bar{\Omega})$ supported in the set

$$
m(y(\cdot)) = \{x \in \bar{\Omega}; |y(x)| = \|y\|_{C(\bar{\Omega})}\}
$$

and such that $y\nu$ is a positive measure and $\|\nu\|_{\Sigma(\Omega)} = \|y\|_{C(\bar{\Omega})}$. It follows that if $\bar{\nu}(\cdot)$ is an optimal control and the target belongs to $D(A_1(\beta))$ then there exists $z \in L_\infty(\Omega)_1(A_1(\beta), I), \mu \in \Sigma(0, \bar{t}; L_\infty(\Omega))$ with $\langle \mu, z \rangle \neq 0$ such that if $\bar{z}(t, x)$ is the solution of

$$
d\bar{z}(t, x) = -A_1^*(\beta')\bar{z}(t, x) - \mu(dt), \quad \bar{z}(\bar{t}, \cdot) = z \tag{6.12}
$$

then the optimal control $\bar{\nu}(t)$ is supported in the set $m(\bar{z}(\cdot, \cdot)), \bar{z}(t, x)\bar{\nu}(t, dx)$ is a positive measure and $\|\bar{\nu}(\cdot)\| = 1$. The space $L^1(\Omega)$ is not a $\zeta$-space, and we know of no results on arbitrary optimal controls.

**Example 6.3.** The system is (6.5). The duality set of $z(\cdot) \in L^1(\Omega)$ consists of all functions $y(\cdot)$ in $L_\infty(\Omega)$ with $y(x) = \|z\|\text{sign } z(x)$ where $z(x) \neq 0$ ($|y(x)| \leq \|z\|$
isomorphism (a) in Lemma 6.4

\[ \Sigma([0, \bar{\Omega}]) \text{ are isometrically isomorphic.} \]
Proof. Isomorphism (a) is obvious. Then so is (b), since \( \Sigma(0, \bar{\Omega}; \Sigma(\Omega)) \) and \( \Sigma([0, \bar{\Omega}] \times \Omega) \) are the duals of isometrically isomorphic spaces. The isomorphism is given as follows: if \( \mu \in \Sigma([0, \bar{\Omega}] \times \Omega) \) we define \( \hat{\mu} \in \Sigma(0, \bar{\Omega}; \Sigma(\Omega)) \) by

\[ \langle \hat{\mu}, f \rangle = \int_{[0, \bar{\Omega}] \times \Omega} f(t, x)\mu(\,dt\,dx) . \]

Conversely, if \( \hat{\mu} \in \Sigma(0, \bar{\Omega}; \Sigma(\Omega)) \) we define an element of \( \Sigma([0, \bar{\Omega}] \times \Omega) \) by

\[ \langle \mu, f \rangle = \int_0^\bar{\Omega} f(t, \cdot)\mu(\,dt) . \]

Using this identification, we rewrite inequality (3.21) in the form

\[ \int_{[0, \bar{\Omega}] \times \Omega} (y(t, x) - y(t, x, \zeta, \bar{u}))\mu(\,dt\,dx) \leq 0 \] (6.15)

for all \( y(\cdot, \cdot) \in \mathbf{M}(\bar{\Omega}) \). Assume that \( \mathbf{M} \) is the unit ball of \( C(\bar{\Omega}) \). Then in view of the isomorphism (a) in Lemma 6.4 \( \mathbf{M}(\bar{\Omega}) \) consists of all functions \( y(t, x) \) continuous in \([0, \bar{\Omega}] \times \Omega \) and satisfying

\[ |y(t, x)| \leq 1 \]

so that (6.15) implies, via a simple 2-dimensional generalization of Lemma 5.1 that \( \mu = 0 \) in \( e(y(\cdot, \cdot, \zeta, \bar{u})) \), where

\[ e(y(\cdot, \cdot)) = \{ (t, x) \in [0, \bar{\Omega}] \times \bar{\Omega}; |y(t, x)| < 1 \} . \]

7. Pointwise evanescence of the costate. Example 6.3 leads to the following

Problem 7.1. Let \( z \in \Sigma(\Omega) \) \( \in \Sigma([0, \bar{\Omega}] \times \Omega) \), \( \mu \in \Sigma(0, \bar{\Omega}; \Sigma(\Omega)) \) \( \in \Sigma([0, \bar{\Omega}] \times \bar{\Omega}) \), \( (\mu, z) \neq 0 \), and let \( \bar{z}(t, x) \) be the solution of

\[ d\bar{z}(t, x) = -A'_1(\bar{z'})\bar{z}(t, x)\,dt - \mu(\,dt) , \quad \bar{z}(\bar{t}, \cdot) = z . \] (7.1)

What can be said about the set

\[ e = \{ (t, x) \in [0, \bar{\Omega}] \times \bar{\Omega}; z(t, x) = 0 \} ? \] (7.2)
The motivation for posing this problem is formula (6.14), which determines the control only outside of \( e \). Since not much information on the measure \( \mu \) is available, nothing obvious can be said of \( e \cap \text{supp}(\mu) \) (\( \text{supp}(\mu) = \text{support of } \mu \)), thus the question should be restricted to \( e \setminus \text{supp}(\mu) \).

We contribute a simple result for the case \( \mu = 0 \), where the objective is to show that the set \( e \subseteq [0, \bar{t}] \times \Omega \) has measure zero. The equation is the homogeneous version of (7.2),

\[
\bar{z}(t, x) = -A_1^\prime(\beta')\bar{z}(t, x), \quad \bar{z}(\bar{t}, \cdot) = \bar{z}
\]  

(7.3)

whose solution operator is \( S_e(\bar{t} - t; A, \beta)^* \), extended to \( E^*(A) = \Sigma(\Omega)(A, (\beta)) \) as in Lemma 2.3. Now, \( S_e(t; A, \beta)^* : \Sigma(\Omega) \to C(\Omega)^\circ = L^1(\Omega) \), thus if we consider the solution of (7.3) in an interval \( 0 \leq t \leq \bar{t} < \tilde{t} \), its final condition \( S_e(\bar{t} - t; A, \beta)^*z \) belongs to \( L^1(\Omega) \). It follows that when we need consider \( z \in L^1(\Omega) \) in (7.3) (this observation extends to the general equation (7.1) if \( \mu(t) = 0 \) for \( t \) near the endtime \( \bar{t} \); this condition is in turn guaranteed if \( Y \subseteq \text{Int}(M) \)).

**Theorem 7.2.** Assume the domain \( \Omega \) is of class \( C(\infty) \), that the operator \( A(\beta) \) is self-adjoint \( (A' = A, \beta' = \beta) \) with coefficients infinitely differentiable in \( \Omega \) and that \( \gamma(\cdot) \) is infinitely differentiable on \( \Gamma \) if \( \beta \) is a variational boundary condition. Then the set \( e \) has measure zero.

**Proof.** By hypoellipticity of the operator \( \partial/\partial t - A_1^\prime(\beta') = \partial/\partial t - A_1(\beta) \) the solution \( \bar{z}(t, x) \) of (7.2) is infinitely differentiable in \([0, \bar{t}] \times \Omega \); in particular, \( z(t, \cdot) \in C(\infty)(\Omega) \) for \( t < \bar{t} \), so that we may exchange \( A_1(\beta) \) by \( A_2(\beta) \) in (7.3) for \( t < \bar{t} \) and (via another replacement of \( \bar{t} \) by \( \tilde{t} < \bar{t} \)) we may assume from the beginning that \( z \in L^2(\Omega) \). In (7.3),

Let \( \{\phi_n(x)\} \) be the eigenfunctions of \( A_2(\beta) \) in \( L^2(\Omega) \) (corresponding to eigenvalues \( \{-\lambda_n\} \)); they are infinitely differentiable in \( \Omega \) and

\[
\|\phi_n\|_{C(\Omega)} = 0(n^k) \quad \text{as} \quad n \to \infty
\]  

(7.4)

(see [31]). Write \( z(t, x) = \Sigma c_n(t)\phi_n(x) \) (convergence in \( L^2(\Omega) \)). We have

\[
c'_n(t) = \int_\Omega z(t, x)\phi_n(x)dx
\]  

(7.5)

\[
\begin{align*}
&= -\int_\Omega A_2(\beta)z(t, x)\phi_n(x)ds \\
&= -\int_\Omega z(t, x)A_2(\beta)\phi_n(x)dx = \lambda_n c_n(t)
\end{align*}
\]

so that

\[
z(\bar{t} - t, x) = \sum c_n e^{-\lambda_n(\bar{t} - t)}\phi_n(x),
\]

\( \{c_n\} \) the Fourier coefficients of \( z \). In view of (7.4) the series is uniformly convergent in \( \Omega \) for \( t < \bar{t} \). Assume now that the set \( e \) in (7.2) has positive measure. Then, if \( \chi \) is its characteristic function,

\[
\int_\Omega \int_0^\bar{t} \chi(t, x)dxdt = \text{meas}(e) > 0
\]
so that there exists a set $d \subseteq \bar{\Omega}$ of positive measure such that, for every $x \in d$ we have
\[ \sum c_n e^{-\lambda_n t} \phi_n(x) = 0 \]
in a set of positive measure, thus, by analyticity, in $0 < t \leq \bar{t}$. However, a Dirichlet series can vanish identically only if all its coefficients are zero \cite{5} so that $\phi_n(x) = 0$ for each $x \in d$ for all $n$. If $x$ is a density point of $d$ then all partial derivatives of $\phi_n$ vanish at $x$; iterating the argument, all partial derivatives of all orders of $\phi_n$ vanish in a set of positive measure, in particular, $\phi_n$ has a zero of infinite order at some point $x_0 \in \Omega$. Now, $\phi_n(x)$ is a solution of the elliptic equation
\[ A_2(\beta) \phi_n(x) + \lambda_n \phi_n(x) = 0 \]
so we can bring to bear results on unique continuation on solutions of elliptic equations having zeros of infinite order \cite{25}, \cite{27}, \cite{29}, \cite{30}, \cite{32}. In particular, using \cite[Theorem 2.1]{30} we deduce that $\phi_n(x) \equiv 0$, which is a contradiction and completes the proof of Theorem 7.2.

Obviously, there is ample room for improvement in Theorem 7.2. Theorem 2.1 in \cite{30} requires infinitely differentiability from the leading coefficients of the operator but no smoothness from lower order coefficients, thus all that is needed is to justify (7.5). For a unique continuation result that does not require infinitely differentiable leading coefficients, see \cite{25}.

8. A distributed parameter system with scalar control. The system is
\[
\begin{align*}
    y(t, x) &= y_{xx}(t, x) + b(x)u(t) \quad (0 < x < \pi), \\
    y(t, 0) &= y(t, \pi) = 0
\end{align*}
\]
in $E = C_0[0, \pi]$, with $A = d^2/dx^2$ with domain $D(A)$ consisting of all twice continuously differentiable $y(\cdot)$ with $Ay \in C_0[0, \pi]$. The eigenvalues of $A$ are \{-n^2; n = 1, 2, \ldots \} corresponding to eigenfunctions $\{\sin nx\}$ and $A$ generates the analytic compact semigroup
\[ S(t)(\Sigma c_n \sin nx) = \sum_{n=1}^{\infty} e^{-n^2 t} c_n \sin nx \quad (8.2) \]
in $C_0[0, \pi]$. We take $F = \mathbb{R}$ and assume $b(x) = \Sigma b_n \sin nx$ with all $b_n \neq 0$. If $\zeta = \Sigma a_n \sin nx$ we have
\[ y(t, \zeta; u) = \sum_{n=1}^{\infty} \left( a_n + b_n \int_0^t e^{-n^2(t-\tau)} u(\tau)d\tau \right) \sin nx \quad (8.3) \]
(the series in (8.2) and (8.3) convergent in the norm of $L^2(0, \pi)$), so that the equation $y(t, 0; u) = y = \Sigma c_n \sin nx$ reduces to the moment problem
\[ \int_0^t e^{-n^2 \tau} u(t-\tau)d\tau = \frac{c_n}{b_n} \quad (n = 0, 1, \ldots) . \quad (8.4) \]
If $\phi_n(t)$ is a sequence biorthogonal to $\{e^{-n^2 t}\}$ (that is, if $\int \phi_j(t)e^{-n^2 t}dt = \delta_{jn}$) then we obtain a solution of (8.4) in the form

$$u(t - \tau) = \sum_{n=1}^{\infty} \frac{c_n}{b_n} \phi_n(\tau)$$  \hspace{1cm} (8.5)

provided that the series is convergent and term-by-term integration can be justified. A biorthogonal sequence satisfying

$$\|\phi_n\|_{C(0,t)} \leq Cn^3 e^{\pi n}$$  \hspace{1cm} (8.6)

(with $C$ depending on $t$) was constructed in [17] using previous results in [24] (this sequence is reasonably optimal in that there exists no other biorthogonal sequence with $\|\phi_n\| = 0(e^{an})$ for $a < \pi$). We may then set

$$Ky = K(\Sigma c_n \sin nx) = \sum_{n=1}^{\infty} \frac{n^{4+\rho} e^{n\pi}}{b_n} c_n \sin nx$$  \hspace{1cm} (8.7)

($\rho > 0$ fixed), the domain $D(K)$ consisting of all $y \in C_0[0, \pi]$ such that the right side belongs to $C_0[0, \pi]$. There is no simple characterization of the domain in terms of summability conditions on the Fourier coefficients, but $K$ has a bounded inverse given by

$$K^{-1}y = K^{-1}(\Sigma c_n \sin nx) = \sum_{n=1}^{\infty} n^{4-\rho} e^{-n\pi} b_n c_n \sin nx$$  \hspace{1cm} (8.8)

and $S(t)E \subseteq D(K)$, $S(t)K^{-1} = K^{-1}S(t)$. Moreover, if $y \in D(K)$ then the $\{n^{4+\rho} e^{n\pi} c_n/b_n\}$ are the sine Fourier coefficients of the continuous function $Ky$, thus they are bounded by a constant times $\|Ky\|_E$. Combining with (8.6), there exists a solution of (8.4) with

$$\|u\|_{C(0,\bar{t})} \leq M\|Ky\|_E = M\|y\|_{D(K)}$$  \hspace{1cm} (8.9)

and it follows that $R(\bar{t}, 0; U)$ contains a ball in $D(K)$, thus the same is true of $R(\bar{t}, \zeta; U)$ for any $\zeta$. We may then apply Theorem 3.1 in the case of point targets $y(t, \zeta, u) = \bar{y} \in D(K)$. We do this with control set

$$U = [-1, 1]$$

and state constraint $y(t, \cdot) \in M = \text{closed convex set in } C[0, \pi]$ with nonempty interior; a possible choice is $M = \text{unit ball of } C[0, \pi]$ in which case the state constraint is

$$|y(t, x)| \leq 1 \hspace{1cm} (0 \leq t \leq \bar{t}, 0 \leq x \leq \pi).$$

We deduce existence of a pair $(\mu, z)$ with $\mu \in \Sigma([0, \bar{t}] \times [0, \pi])$, $z \in E_1^*(K, B) = \Sigma_0[0, \pi]_1(K, B)$, $(\mu, z) \neq 0$ such that

$$\bar{u}(t) = \frac{\sigma(t)}{\sigma'(t)} \text{ where } \sigma(t) \neq 0,$$  \hspace{1cm} (8.10)
\[ \sigma(t) = \int_0^\pi b(x) \tilde{z}(t, x) \, dx, \quad (8.11) \]
\[ \tilde{z}(t, x) \text{ the solution of the backwards heat equation} \]
\[ d\tilde{z}(t, x) = -\tilde{z}_{xx}(t, x) \, dt - \mu(dt) \quad (0 < x < \pi), \quad (8.12) \]
\[ \tilde{z}(t, 0) = \tilde{z}(t, \pi) \]

with final condition \( z(\bar{t}, x) = z \). If the constraint is never saturated then Theorem 5.2 shows that \( \sigma(t) \) is nonzero except perhaps in a sequence \( \{\bar{t}_n\} \) whose only point of accumulation, if any, is \( \bar{t} \). We note that the condition that \( \bar{y} \in D(K) \) can be checked without computing sine Fourier coefficients; in fact, it will hold if \( \bar{y}(\cdot) \) can be extended to a function \( \bar{y}(x + i\xi) \) (a) 2\(\pi\)-periodic and odd in \( x \), (b) analytic in the strip \( |\xi| < \pi \), (c) infinitely differentiable in the strip \( |\xi| \leq \pi \). (See [17].)

**Remark 8.1.** Theorem 4.4 can be used with other operators \( K \). For instance, we may take \( (\Sigma a_n \sin nx) = \Sigma a_n^k \sin nx \) and \( Y \) a closed ball of center \( \bar{y} \in D(K) \) and positive radius in \( D(K) \). This situation roughly corresponds to approximation of the target in Sobolev norm of order \( k \).

9. **Other cost functionals.** We consider the optimal problem for
\[ y'(t) = Ay(t) + Bu(t), \quad y(0) = \zeta \quad (9.1) \]
in a fixed or variable time interval \( 0 \leq t \leq \bar{t} \) with cost functional
\[ y_0(\bar{t}, u) = \int_0^{\bar{t}} f_0(y(\tau), u(\tau)) \, d\tau. \quad (9.2) \]

The problem is a bit of a hybrid: linear system, nonlinear functional. We assume that the Fréchet derivative \( \partial_y f_0(y, u) \) exists in \( E \times U \) and is continuous in \( y \) for \( u \) fixed, that \( f_0(y, u(\cdot)) \) is measurable and that \( \partial_y f_0(y, u(\cdot)) \) is strongly measurable (as an \( E^* \)-valued function) for \( u(\cdot) \in C_{ad}(0, \bar{t}; U) \). Further, for every \( c > 0 \) there exists \( K(c), L(c) \) such that
\[ \| f_0(y, u) \|_E \leq K(c), \quad \| \partial_y f_0(y, u) \|_E \leq L(c) \quad (u \in U, \|y\| \leq c). \quad (9.3) \]

Let \( u(\cdot) \in C_{ad}(0, \bar{t}; U) \), \( 0 < s \leq \bar{t}, \, v \in U \), and let \( u_{s,v,h}(\cdot) \) be the spike perturbation of \( u(\cdot) \). Then
\[ y(t, u_{s,v,h}) - y(t, u) = \int_{s-h}^s S^0(t - \tau)^* B(v - u(\tau)) \, d\tau \]
\[ = \int_{s-h}^s S(s - \tau) S^0(t - s)^* B(v - u(\tau)) \, d\tau \]
where the function \( S^0(t - s)^* u(\tau) \) is strongly measurable (see §2). We can then use the theory of Lebesgue points to show.

**Lemma 9.1.** There exists a set \( e \) of full measure in \([0, \bar{t}]\) such that
\[ \lim_{h \to 0+} \frac{1}{h} \left( y(t, u_{s,v,h}) - y(t, u) \right) = S^0(t - s)^* B(v - u(s)) \quad (t > s) \]
for \( s \in e \).

Convergence is uniform in \( t \geq s + \varepsilon \), and approximations remain bounded, which plays a role in the proof of the result below, which can be found in [19, 20].
Lemma 9.2. There exists a set $e$ of full measure in $[0, \bar{t}]$ such that for every $s \in e$ we have
\[
\lim_{h \to 0^+} \frac{1}{h} (y_0(\bar{t}, u_{s,v,h}) - y_0(\bar{t}, u)) = \int_0^\bar{t} \langle \partial_y f_0(y(\tau, u), u(\tau)), S^\circ (\tau - s)^* B(v - u(s)) \rangle d\tau + f_0(y(s, u), v) - f_0(y(s, u), u(s)).
\]

Existence results require weak lower semicontinuity of the functional; if \{u_n(\cdot)\} is $L^1(0, \bar{t}; X)$-weakly convergent in $L^\infty(0, \bar{t}; X^*)$ to $\bar{u}(\cdot)$ then
\[
y_0(\bar{t}, \bar{u}) \leq \liminf_{n \to \infty} y_0(\bar{t}, u_n)
\]
and the proof of the maximum principle requires convexity: if $u, v \in U$ and $0 \leq \alpha \leq 1$ then
\[
y_0(\bar{t}, \alpha u + (1 - \alpha)v) \leq \alpha y_0(\bar{t}, u) + (1 - \alpha)y_0(\bar{t}, v).
\]
The results below are on the target condition
\[
y(\bar{t}, \zeta, u) \in Y
\]

Theorem 9.3. Assume that the state constraint set $M$ and the target set $Y$ are closed. Then, if there exists a minimizing sequence \{u_\bar{t}(\cdot)\} in $L^\infty(0, \bar{t}; X^*)$ driving $\zeta$ to $Y$ in time $t_\bar{t}$ satisfying the state constraint and \{t_\bar{t}\} is bounded the optimal problem has a solution.

The proof is the same as that of Theorem 2.7.

The operator $K$ below satisfies the requirements in Section 2. We assume from now on that $Y \subseteq D(K)$.

Theorem 9.4. Let $\bar{u}(\cdot)$ be an optimal control in the interval $0 \leq t \leq \bar{t}$. Assume that $U$, $M$ and $Y$ are convex and closed, $\text{Int}(U) \neq \emptyset$ in $X^*$ and $\text{Int}(M) \neq \emptyset$ in $E$, and that
\[
R(\bar{t}, \zeta; U) - Y
\]
contains a ball in $D(K)$. Then there exists $(z_0, \mu, z) \in \mathbb{R} \times \Sigma(0, \bar{t}; E^*) \times E^*_1(K, B),\ (z_0, \mu, z) \neq 0$ with $z_0 \geq 0, \mu \in \mathcal{N}_M(\bar{t}) \cdot (\bar{y}(\cdot, \bar{u}), \bar{u}(\cdot)), z \in \mathcal{N}_Y(\bar{y}(\cdot, \bar{u}))$ (the latter normal cone defined according to the duality (3.1)) and such that
\[
z_0 f_0(y(s, \zeta, u), \bar{u}(s)) + \langle B^* \bar{z}(s), \bar{u}(s) \rangle = \min_{v \in U} \{z_0 f_0(y(s, \zeta, u), v) + \langle B^* \bar{z}(s), v \rangle \}
\]
a.e. in $0 \leq s \leq \bar{t}$, $\bar{z}(\cdot)$ the solution of
\[
d\bar{z}(t) = -A^*\bar{z}(t)dt - z_0 \partial_y f_0(y(t, u), \bar{u}(t))dt - \mu(dt), \quad \bar{z}(\bar{t}) = z.
\]

Proof. Note first that the assumptions on $\partial_y f_0(y(t, \bar{u}), \bar{u}(t))$ qualify $\partial_y f_0(y(t, \bar{u}), \bar{u}(t))dt$ as an element of $\Sigma(0, \bar{t}; E^*)$, thus the costate $\bar{z}(\cdot)$ is well defined. Given $a \in \mathbb{R}$, $M_K(a, Y) \subseteq \mathbb{R} \times C(0, \bar{t}; E) \times E$ consists of all elements of the form
\[
(r, y(\cdot), Ky) \quad (r \leq a, y(\cdot) \in M(\bar{t}), y \in Y)
\]
and (modifying the definition in Section 2), \( R_K(\bar{t}, \zeta; U) \subseteq \mathbb{R} \times C(0, \bar{t}; E) \times E \) consists of all elements of the form

\[
(r, y(\cdot, \zeta, u), K y(\bar{t}, \zeta, u)) \quad (r \geq y_0(\bar{t}, u), u(\cdot) \in C_{ad,K}(0, \bar{t}, u)).
\]

(9.11)

Obviously, \( M_K(\bar{t}, a, Y) \) is convex; convexity of \( R_K(t, \zeta; U) \) follows from convexity of \( y_0(\bar{t}, u) \). If \( m \) is the minimum of the functional \( y_0(\bar{t}, u) \) (under state and control constraints and target condition) then for every \( \varepsilon > 0 \) the sets \( M_K(m - \varepsilon, t, \bar{y}) \) and \( R_K(t, \zeta, U) \) are obviously disjoint, and in fact lie at a positive distance. If this were not true we would have three sequences \( \{r_n\} \subseteq \mathbb{R} \), \( \{u_m(\cdot)\} \subseteq C_{ad,K}(0, \bar{t}; U) \) and \( y_m(\cdot) \in M(\bar{t}) \) such that

\[
y_0(\bar{t}, u_m) - r_m \to 0, \quad r_m \leq m - \varepsilon,
\]

(9.12)

\[
y(t, \zeta, u_m) - y_m(t) \to 0
\]

(9.13)

uniformly in \( 0 \leq t \leq \bar{t} \) and

\[
K y(\bar{t}, \zeta, u_m) - K y_m \to 0
\]

(9.14)

as \( m \to \infty \). Operating as in Lemma 2.7 we may select a subsequence of \( \{u_m\} \) such that \( u_m(\cdot) \to \bar{y}(\cdot) L^1(0, \bar{t}; E) \)-weakly in \( L^\infty(0, \bar{t}; X^*) \) and \( y(t, \zeta, u_m) \to y(t, \zeta, \bar{u}) \).

Applying \( K^{-1} \) to both sides we deduce that \( y(t, \zeta; u_m) - y_m \to 0 \) so that \( y(t, \zeta, \bar{u}) \in Y \) with \( y_0(\bar{t}, \bar{u}) \leq m - \varepsilon \) by (9.12) and weak lower semicontinuity of \( y_0(\bar{t}, u) \), a contradiction to optimality of \( \bar{u} \).

Since \( \text{dist}(M_K(m - \varepsilon, \bar{t}, Y), R_K(\bar{t}, \zeta; U)) > 0 \) we apply the separation theorem and construct \( (z_0, \mu, w_0) \in \mathbb{R} \times \Sigma(0, \bar{t}; E^*) \times E^*, \|z_0, \mu, w_0\| = 1 \) such that

\[
z_0 r + \int_0^t \langle \mu_n(dt), y(t) \rangle + \langle w_n, Ky \rangle
\]

\[
\leq z_0 y_0(\bar{t}, u) + \int_0^t \langle \mu_n(dt), y(t, \zeta, u) \rangle + \langle w_n, Ky(t, \zeta, u) \rangle
\]

(9.15)

for \( r \leq m - \varepsilon \), \( y(\cdot) \in M(\bar{t}) \), \( y \in Y \) and \( u(\cdot) \in C_{ad,K}(0, \bar{t}; U) \). We select a weakly convergent subsequence of \( (z_0, \mu_n, w_n) \), and prove that \( (z_0, \mu, w_0) \neq 0 \). If \( z_0 \to z_0 \neq 0 \) there is nothing to prove; if \( z_0 \to 0 \), then the proof that \( (\mu, w) \neq 0 \) is exactly the same as that in Theorem 3.1, as is the proof that

\[
\int_0^t \|B^* K^* S(t)^* w\| dt \leq C,
\]

(9.16)

the construction of the element \( z \in E^*_1(K, B) \) and the taking of the limit in (9.15) as \( n \to \infty \). We end up with

\[
z_0 r + \int_0^t \langle \mu(dt), y(t) \rangle + \langle w, Ky \rangle
\]

\[
\leq z_0 y_0(\bar{t}, u) + \int_0^t \langle \mu(dt), y(t, \zeta, u) \rangle + \langle w, Ky(t, \zeta, u) \rangle
\]

(9.17)
for $r \leq m$, $y(\cdot) \in M(\tilde{l})$, $y \in Y$ and $u(\cdot) \in C_{ad.K}(0,\tilde{l};U)$. We take first $y(t) = y(\tilde{l}, \zeta, \tilde{u})$, $r = y_0(\tilde{l}, \tilde{u})$ and $y = y(\tilde{l}, \zeta, \tilde{u})$, obtaining

$$z_0 y_0(\tilde{l}, \tilde{u}) + \int_0^\ell \langle B^* \tilde{z}(t), \tilde{u}(t) \rangle \, dt \quad (9.18)$$

$$\leq z_0 y_0(\tilde{l}, u) + \int_0^\ell \langle B^* \tilde{z}(t), u(t) \rangle \, dt$$

for all $u(\cdot) \in C_{ad.K}(0,\tilde{l};U)$, with $\tilde{z}(t)$ the solution of $d\tilde{z}(t) = -A^* \tilde{z}(t) \, dt - \mu(\tilde{z}(t))$, $\tilde{z}(\tilde{l}) = z$. We use inequality (9.18) with spike perturbations as after (3.19). The effect of a spike on the cost functional has been elucidated in Lemma 9.2. We note that, by Lemma 2.6 applied to the $E^*$-valued measure $\partial y f_0(y(\tau, u), u(\tau)) \, d\tau$ we have

$$\int_s^t \langle \partial_y f_0(y(\tau, u), u(\tau)), S^o (\tau - s)^* B(v - u(s)) \rangle \, d\tau$$

$$= \int_s^t \langle S(\tau - s)^* \partial_y f_0(y(\tau, u), u(\tau)), B(v - u(s)) \rangle \, d\tau$$

thus we obtain (9.8) applying the definition (2.13) of solutions of (2.11). It follows immediately from (9.15) and the definition of $M_{\tilde{K}}(m - \varepsilon, \tilde{l}, Y)$ that $z_{0n} \geq 0$, thus $z_0 \geq 0$. That $\mu \in N_{M_{\tilde{K}}}(y(\cdot, \zeta, \tilde{u}))$ and $z \in N_Y(y(\tilde{l}, \zeta, \tilde{u}))$ is proved as in Theorem 3.1.

10. Final observations

**Remark 10.1.** The requirement that $\bar{y} \in D(A)$ in Theorem 4.1 raises the question of whether there actually exist trajectories with $y(\tilde{l}, \zeta, u) \neq D(A)$ or, on the contrary,

$$R^\infty(\tilde{l}) = D(A). \quad (10.1)$$

This question was addressed in [16] for a general semigroup $S(t)$ (not necessarily analytic or compact). It was concluded in [16] (correctly) that (10.1) implies

$$\int_0^1 \|S(t)^* z\| \, dt < \infty \quad (10.2)$$

for all $z \in E^*$ and then (incorrectly) that (10.2) implies that $A$ has to be bounded. In fact, there are unbounded infinitesimal generators that satisfy (10.2) for all $z \in E^*$; an example (and the only one we know) is the following: $E$ is the space $c_0$ of all sequences $\{\xi_n; n = 1, 2, \ldots\}$ satisfying $\lim \xi_n = 0$, endowed with the norm $\|\{\xi_n\}\| = \sup |\xi_n|$, and $A\{\xi_n\} = \{-u\xi_n\}$, $D(A)$ consisting of all $\{\xi_n\} \in E$ such that the right hand side belongs to $E$. $A$ is the infinitesimal generator of the (analytic, compact) semigroup

$$S(t)\{\xi_n\} = \{e^{-nt}\xi_n\}. \quad (10.4)$$

The dual space $E^* = \ell^1$ consists of all sequences $\{\eta_n\}$ with $\|\{\eta_n\}\| = \Sigma |\eta_n| < \infty$ equipped with $\|\cdot\|$, and the adjoint semigroup is again given by (10.4); we check easily that (10.2) holds for every $\{\eta_n\} \in E^*$. 
On the other hand, it is easy to see that (10.2) can never hold for, say, an unbounded self adjoint infinitesimal generator in Hilbert space.

**Remark 10.2.** Existing proofs of the maximum principle for nonlinear systems usually rely on Kuhn - Tucker type theorems for infinite dimensional nonlinear programming problems [19], [20], [21]. Applied to a system like (8.1), application of these theorems would require computation of differentials or “directional derivatives” of the solution map \( u \rightarrow y(t, \zeta, u) \) in the norm of \( D(K) \) rather than in the norm of the original space \( E \). We don’t know if these computations are possible; in particular we don’t know if any version of the maximum principle is available for nonlinear perturbations of (8.1) (by terms of the form \( f(y(t, x)) \)).

**REFERENCES**

[1] R.A. Adams, “Sobolev Spaces,” Academic Press, New York, 1975.
[2] H. Amann, M. Hieber and G. Simonett, *Bounded \( H_\infty \)-calculus for elliptic operators*, Differential and Integral Equations, 7 (1994) 13–47.
[3] R.G. Bartle, *A general bilinear vector integral*, Studia Math., 15 (1956) 337–352.
[4] R. Bellman, I. Glicksberg and O. Gross, *On the “bang-bang” control problem*, Quart. Applied Math., 14 (1956) 11–18.
[5] V. Bernstein, “Lecons sur les Progrés Récents de la Théorie des series de Dirichlet,” Gauthier–Villars, Paris, 1933.
[6] J. Bourgain, *Some remarks on Banach spaces in which martingale difference sequences are unconditional*, Ark. Mat., 21 (1983) 163–168.
[7] D.L. Burkholder, *A geometric condition that implies the existence of certain singular integral of Banach space valued functions*, Wadsworth Math. Series, Wadsworth, Belmont (1983) 270–286.
[8] P. Butzer and H. Berens, “Semi-Groups of Operators and Approximation,” Springer, Berlin–Heidelberg 1967.
[9] I. De Simon, *Un’applicazione della teoria degli integrali singolari allo studio delle equazioni differenziali astratte del primo ordine*, Rend Sem. Mat. Univ. Padova, 34 (1964) 205–223.
[10] J. Diestel and J.J. Uhl, “Vector Measures,” Amer. Math. Society, Providence, R.I., 1977.
[11] J. Dieudonne, *Sur le théorème de Lebesgue-Nikodym (III)*, Ann. Université Grenoble, 23 (1947-48) 25–53; (IV), J. Indian Math. Soc., 22 (1951) 77–86.
[12] G. Dore and A. Venni, *On the closedness of the sum of two closed operators*, Math. Z., 196 (1987) 189–201.
[13] N. Dunford, and J.T. Schwartz, “Linear Operators,” Interscience–Wiley, New York, 1957.
[14] H.O. Fattorini, *Time-optimal control of solutions of operational differential equation*, J. SIAM Control, 2 (1964) 54–59.
[15] , *Some remarks on complete controllability*, SIAM J. Control, 4 (1966) 686–694.
[16] , *The time-optimal control problem in Banach spaces*, Appl. Math. Optimization, 1 (1974) 163–188.
[17] , *Boundary control of temperature distributions in a parallelepipedon*, SIAM J. Control, 13 (1975), 1–13.
[18] , “The Cauchy Problem, Encyclopedia of Mathematics and its Applications vol. 18,” Cambridge University Press, Cambridge 1983.
[19] , *A unified theory of necessary conditions for nonlinear nonconvex control systems*, Applied Math. Optim., 15 (1987) 141–185.
[20] , *Optimal control problems for distributed parameter systems in Banach spaces*, Applied Math. & Optimization, 28 (1993) 225–257.
[21] , *Relaxation in infinite dimensional control systems*, in “L. Markus Festschrift Volume” (ed. K.D. Elworthy, W.N. Everitt, and E.B. Lee), Marcel Dekker, 1993, 505–522.
[22] , *Optimal control problems with state constraints for semilinear distributed parameter systems*, Jour. Optimization Theory & Appl., to appear.
[23] , *Optimal control problems with state constraints for semilinear distributed parameter systems: the abstract parabolic case*, to appear.
[24] H.O. Fattorini and D.L. Russell, *Exact controllability theorems for linear parabolic equations in one space dimension*, Arch. Rat. Mech. Anal., 43 (1971) 272–292.
[25] N. Garofalo and Fanghua Lin, Unique continuation for elliptic operators: a geometric-variational approach, Comm. Pure Appl. Math., 40 (1987) 347–366.
[26] E. Hille and R.S. Phillips, “Functional Analysis and Semi-Groups,” AMS Colloquium Pubs., vol. 31, Providence, 1957.
[27] D. Jerison and C.E. Kenig, Unique continuation and absence of positive eigenvalues for Schrödinger operators, Annals of Math., 121 (1985), 463–488.
[28] K. Miller, Nonunique continuation for uniformly parabolic and elliptic equations in self-adjoint divergence form with Hölder continuous coefficients, Arch. Rat. Mech. Anal., 54, 105–117.
[29] L. Robbiano, Dimension des zéros d’une solution faible d’un opérateur elliptique, J. Math. Pures et Appl., 67 (1988) 339–357.
[30] C. Sogge, Strong uniqueness theorems for second order elliptic differential equations, Amer. J. Math., 112 (1990), 943–984.
[31] C. Sogge, “Fourier Integrals in Classical Analysis,” Cambridge University Press, New York, 1993.
[32] E.M. Stein, Appendix to “Unique continuation”, Annals of Math., 121 (1985), 489–494. Received October, 1994.