Math. Control Signals Systems (1991) 4: 41–67

Mathematics of Control, Signals, and Systems

Necessary Conditions for Infinite-Dimensional Control Problems*

H. O. Fattorini† and H. Frankowska‡

Abstract. We consider infinite-dimensional nonlinear programming problems which consist of minimizing a function $f_0(u)$ under a target set constraint. We obtain necessary conditions for minima that reduce to the Kuhn–Tucker conditions in the finite-dimensional case. Among other applications of these necessary conditions and related results, we derive Pontryagin's maximum principle for a class of control systems described by semilinear equations in Hilbert space and study convergence properties of sequences of near-optimal controls for these systems.

Key words. Lagrange multiplier rule, Kuhn–Tucker conditions, Maximum principle, Optimal controls, Approximately optimal controls.

1. Introduction

We consider in this paper the following *abstract nonlinear programming problem:* Let V be a complete metric space, let E be a Hilbert space, let f, f_0 be two continuous functions,

$$f: V \to E \qquad f_0: V \to \mathbb{R}, \tag{1.1}$$

and let Y (the *target set*) be a subset of E. Characterize the elements $\bar{u}$ of V that satisfy

$$f_0(\bar{u}) = m = \min f_0(u) \qquad (u \in U) \tag{1.2}$$

$$\text{subject to} \quad f(u) \in Y. \tag{1.3}$$

The main application in mind is the control problem where

$$f(u) = y(\bar{t}, u), \qquad f_0(u) = y_0(\bar{t}, u), \tag{1.4}$$

with $y(t, u)$ given by a semilinear initial value problem in E,

$$y'(t, u) = Ay(t, u) + f(t, y(t, u), u(t)), \qquad y(0) = y^0, \tag{1.5}$$

---

* Date received: March 15, 1988. Date revised: June 9, 1989.

† Department of Mathematics, University of California, Los Angeles, California 90024, U.S.A. The work of this author was supported in part by the National Science Foundation under Grant DMS-8701877

‡ CEREMADE, Université de Paris-Dauphine, 75775 Paris CX (16), France.

and $y_0(t, u)$ being a suitable cost fuctional; here, the space V is a suitable control space of F-valued functions defined in $0 \leq t \leq \bar{t}$, where $\bar{t}$ is the time (or one of the times) taken to hit the target Y optimally. More generally, the maps $y(t, u)$, $y_0(t, u)$ may be *systems* as defined in [F1]; this formulation also includes, for instance control problems for delay differential equations.

The main result of the problem (1.2)–(1.3) is a result of Kuhn–Tucker type (Theorem 2.4) for solutions $\bar{u}$ (assumed to exist). This result, when applied to the maps (1.4) yields a version of Pontryagin's maximum principle for optimal controls $\bar{u}(\cdot)$. It also applies to the time optimal problem, but it may yield empty results, thus a separate treatment is necessary. This motivates consideration of the following *abstract time optimal problem:* Let $\{V_n\}$ be a sequence of complete metric spaces, let E be Hilbert space, let $\{f_n\}$ be a sequence of continuous functions,

$$f_n\colon V_n \to E, \tag{1.6}$$

and let Y a be subset of E. Assume that

$$f_n(V_n) \cap Y = \varnothing \tag{1.7}$$

for all n. Characterize the sequences $\{\bar{u}^n\}$, $\bar{u}^n \in V_n$, that satisfy

$$f_n(\bar{u}^n) \to \bar{y} \tag{1.8}$$

in the norm of E for some $\bar{y} \in Y$.

Motivation is provided by the usual time optimal control problem: if $\bar{u}$ is a time optimal control and $\{t_n\}$ is a sequence of times with $t_n < \bar{t}$, $t_n \to \bar{t}$, then $y(t_n, \bar{u}) \notin Y$, while, on the other hand, we have $\|y(t_n, \bar{u}) - \bar{y}\| = \|y(t_n, \bar{u}) - y(\bar{t}, \bar{u})\| \to 0$ as $n \to \infty$ since the optimal trajectory $y(t, \bar{u})$ (as any other trajectory) is continuous. The result for the abstract time optimal problem is Theorem 2.6.

The results in this paper and those in [F1] compare as follows:

(a) The treatment in [F1] applies to systems and a particular type of variation (spike variation) of controls. The present results apply to arbitrary functions and cost functionals (thus nonevolutionary problems can also be handled) and to arbitrary variations of all parameters.
(b) The results in [F1] apply to very restricted target sets, whereas arbitary closed target sets are handled here.
(c) The separate treatments of the point target and set target cases in [F1], are unified.

In the case of finite-dimensional problems, our results are related to work by Clarke (see [C2] and bibliography therein) and Clarke and Löwen [CL] where a "proximal normal analysis" approach to minimization problems is carried out, leading to different approximate multiplier rules.

Since the results in this paper are obtained using Ekeland's variational principle, which refers to *approximate* solutions of minimization problems, the same approach yields information about sequences $\{u^n\}$ of *suboptimal* elements u^n that satisfy (1.2)–(1.3) approximately, that is,

$$f_0(u^n) \leq \min f_0(u) + \varepsilon_n, \qquad \operatorname{dist}(f(u^n), Y) \leq \varepsilon_n$$

for a small ε_n (see Section 5 for a precise formulation). Results on suboptimal sequences are treated in Sections 5 and 6. In Section 5 we obtain *sequence maximum principles* (or sequence Kuhn–Tucker theorems) for the separate elements of a suboptimal sequence $\{u^n\}$. In Section 6 these sequence maximum principles are upgraded to *convergence principles*, that is, it is shown that (a suitable subsequence) of the Kuhn–Tucker multipliers is convergent to a nonzero multiplier; convergence is weak or strong depending on the type of hypotheses that are placed on the target sets and on the functions f_0, f. Finally, we point out in Section 6 how convergence principles can be used to show strong convergence of sequences of suboptimal controls in the case of the systems defined by (1.5). We note in passing that results on suboptimal controls are probably much more interesting that results on optimal controls. In fact, optimal controls may not exist, and when they do, they are almost always impossible to compute explicity in truly infinite-dimensional problems. Thus, information on suboptimal controls (that can be approximated numerically) is more useful. Finally, this information also bears on the problem of *robustness* of control systems, that is, insensitivity of optimal controls with respect to small changes in the parameters of the system.

2. The Maximum Principles

We recall some definitions from nonsmooth analysis. Let Y be an arbitrary subset of a Hilbert space E. Given $\bar{y} \in H$, the *contingent cone to Y at $\bar{y}$* is the set $K_Y(\bar{y})$ defined as follows: $w \in K_Y(\bar{y})$ if and only if there exist a sequence $\{h_k\}$ of positive numbers with $h_k \to 0$ and a sequence $\{y_k\} \subset Y$ such that $(y_k - \bar{y})/h_k \to w$ as $k \to \infty$ (equivalently, if and only if there exist a sequence $\{h_k\}$ as above and a sequence $\{w_k\} \subset E$ such that $w_k \to w$ and $\bar{y} + h_k w_k \in Y$). $K_Y(\bar{y})$ is a closed (in general, nonconvex) cone. The *Clarke tangent cone $C_Y(\bar{y})$ to Y at $\bar{y} \in Y$* consists of all $w \in E$ such that, for every sequence $\{h_k\}$ of positive numbers and every sequence $\{\bar{y}_k\} \subset Y$ such that $h_k \to 0$, $\bar{y}_k \to \bar{y}$, there exists a sequence $\{y_k\} \subset Y$ such that $(y_k - \bar{y}_k)/h_k \to w$ (equivalently $w \in C_Y(\bar{y})$ if there exists a sequence $\{w_k\} \subset H$ with $w_k \to w$ and $\bar{y}_k + h_k w_k \in Y$). It can be shown that $C_Y(\bar{y})$ is convex and closed (see p. 407 [AE]). It is obvious from the definitions that $C_Y(\bar{y}) \subset K_Y(\bar{y})$. These cones are in general different, although if Y is a convex set we have

$$C_Y(\bar{y}) = K_Y(\bar{y}) = \mathrm{Cl}\left[\bigcup_{\lambda \geq 0} \lambda(Y - \bar{y})\right] \tag{2.1}$$

[AE, p. 407]. The negative polar cone $N_Y(\bar{y}) = C_Y(\bar{y})^-$ of all $z \in E$ with $\langle z, w\rangle \leq 0$, $w \in C_Y(\bar{y})$, is called the *Clarke normal cone to Y at $\bar{y}$*.

Let V be a metric space, let E be a linear topological space, and let $g\colon V \to E$ be an arbitrary map, not necessarily defined for all $u \in V$. Given a point $u \in V$ where $g(u)$ exists, a vector $\xi \in E$ is called a *(first-order) variation of g at u* if and only if for some $\delta > 0$ and for all h, $0 < h \leq \delta$, there exists an element $v = v(h) \in V$ where g is defined as well and

$$d(v(h), u) \leq h, \tag{2.2}$$

$$\lim_{h\to 0+} h^{-1}(g(v(h)) - g(u)) = \xi. \tag{2.3}$$

The *variation set* of all variations of g at u is called $\partial g(u)$. For instance, if g is a function defined in $\mathbb{R}^m$, a directional derivative at u in any direction v, $\|v\| = 1$, belongs to $\partial g(u)$.

Finally, we recall Kuratowski's lim inf Z_n, where $\{Z_n\}$ is a sequence of subsets of a metric space V: $z \in \liminf_{n\to\infty} Z_n$ if and only if there exists a sequence $\{z_n\}$, $z_n \in Z_n$, such that $z = \lim_{n\to\infty} Z_n$. A similar definition applies to $\liminf_{w\to\bar{w}} Z(w)$, where $\{Z(w);\ w \in W\}$ is a family of subsets of a metric space V indexed by elements of another metric space W.

Theorem 2.1. *Let the target set Y be closed and let $\bar{u}$ be a solution of the abstract nonlinear programming problem (1.2)–(1.3). Then there exist a sequence $\{\delta_n\}$ of positive numbers with $\delta_n \to 0$, a sequence $\{u^n\} \subset V$, a sequence $\{y^n\} \subset Y$ such that*

$$d(u^n, \bar{u}) + \|y^n - f(\bar{u})\| \le \delta_n, \tag{2.4}$$

and a sequence $\{(\mu_n, z_n)\} \subset \mathbb{R} \times E$ satisfying

$$\mu_n \ge 0, \qquad \|(\mu_n, z_n)\| = 1, \tag{2.5}$$

such that, for every $(\eta^n, \xi^n) \in \overline{\mathrm{conv}}\,\partial(f_0, f)(u^n)$ and every $w^n \in K_Y(y^n)$,

$$\mu_n\eta^n + \langle z_n, \xi^n - w^n\rangle \ge \delta_n(1 + \|w^n\|). \tag{2.6}$$

Moreover, for every cluster point (μ, z) of the sequence $\{(\mu_n, z_n)\}$ in $\mathbb{R} \times E$ (E endowed with its weak topology) we have

$$\mu \ge 0, \qquad z \in N_Y(f(\bar{u})) \tag{2.7}$$

and for every $(\eta, \xi) \in \liminf_{n\to\infty} \overline{\mathrm{conv}}\,\partial(f_0, f)(u^n)$

$$\mu\eta + \langle z, \xi\rangle \ge 0. \tag{2.8}$$

Proof. Given an arbitrary sequence $\{\varepsilon_n\}$ of positive numbers tending to zero we define for each n the real-valued function F_n by

$$F_n(u, y) = \{(\max(0, f_0(u) - m + \varepsilon_n))^2 + \|f(u) - y\|^2\}^{1/2} \tag{2.9}$$

in the space $V \times Y$ endowed with the product distance $\tilde{d}((u, y), (u', y')) = d(u, u') + \|y - y'\|$; this space is a complete metric space and m is the minimum in (1.2). The function F_n is continuous and positive in $V \times Y$ and

$$F_n(\bar{u}, f(\bar{u})) = \varepsilon_n \le \inf F_n + \varepsilon_n, \tag{2.10}$$

thus, by Ekeland's principle [E2] there exist $(u^n, y^n) \in V \times Y$ ($n = 1, 2, \ldots$) such that

$$F_n(u^n, y^n) \le F_n(\bar{u}, f(\bar{u})) = \varepsilon_n, \tag{2.11}$$

$$d(u^n, \bar{u}) + \|y^n - f(\bar{u})\| \le \sqrt{\varepsilon_n}, \tag{2.12}$$

$$F_n(v, y) \ge F_n(u^n, y^n) - \sqrt{\varepsilon_n}(d((v, u^n) + \|y - y^n\|)) \qquad ((v, y) \in V \times Y). \tag{2.13}$$

Let $w^n \in K_Y(y^n)$, so that there exist a sequence $\{h_k\}$ of positive numbers with $h_k \to 0$

and a sequence $\{w_k\}$ in E with $w_k \to w^n$ and $y^n + h_k w_k \in Y$. Applying (2.13) to the function (2.9) we obtain

$$\{(\max(0, f_0(v) - m + \varepsilon_n))^2 + \|f(v) - y^n - h_k w_k\|^2\}^{1/2}$$
$$\geq \{(\max(0, f_0(u^n) - m + \varepsilon_n))^2 + \|f(u^n) - y^n\|^2\}^{1/2} - \sqrt{\varepsilon_n}(d((v, u^n) + h_k\|w_k\|)). \tag{2.14}$$

Let $(\eta^n, \xi^n) \in \partial(f_0, f)(u^n)$. We write (2.14) for $v = v(h_k)$, $v(h)$ being the function used in the definition (2.2)–(2.3) of variation. Substituting in (2.14), dividing by h_k, and letting $k \to \infty$ we obtain (2.6) with

$$(\mu_n, z_n) = (\lambda_n, x_n)/\|(\lambda_n, x_n)\|, \tag{2.15}$$

$$(\lambda_n, x_n) = (\max(0, f_0(u^n) - m + \varepsilon_n), f(u^n) - y^n). \tag{2.16}$$

Obviously, (2.6) extends to $\overline{\text{conv}}\,\partial(f_0, f)(u^n)$.

Inequality (2.8) is obtained from (2.6) setting $w^n = 0$ and taking limits as $n \to \infty$. The first condition (2.7) is obvious; thus it only remains to prove the second, i.e., that $z \in N_Y(f(\bar{u}))$. Let $w \in C_Y(f(\bar{u}))$ and let $\{h_n\}$ be a sequence of positive numbers such that

$$h_n/\|(\lambda_n, x_n)\| \to 0 \qquad (n \to \infty), \tag{2.17}$$

where (λ_n, x_n) is the vector (2.16); as $\|(\lambda_n, x_n)\| = F_n(u^n, y^n) \leq \varepsilon_n \to 0$ we have $h_n \to 0$ as well, thus by definition of $C_Y(f(\bar{u}))$ we can pick a sequence $\{w^n\}$ in E such that $w^n \to w$, $y^n + h_n w^n \in Y$. Then we obtain from (2.13) that

$$\begin{aligned}
&F_n(u^n, y^n + h_n w^n)\\
&\quad = \{\|(\lambda_n, x_n)\|^2 - 2h_n\langle f(u^n) - y^n, w^n\rangle + h_n^2\|w^n\|^2\}^{1/2}\\
&\quad = \|(\lambda_n, x_n)\|\{1 - 2(h_n/\|(\lambda_n, x_n)\|)\langle z_n, w^n\rangle + h_n^2\|w^n\|^2/\|(\lambda_n, x_n)\|^2\}^{1/2}\\
&\quad = \|(\lambda_n, x_n)\|\{1 - (h_n/\|(\lambda_n, x_n)\|)\langle z_n, w^n\rangle + o(h_n/\|(\lambda_n, x_n)\|)\}\\
&\quad \geq F_n(u^n, y^n) - \sqrt{\varepsilon_n} h_n\|w^n\|\\
&\quad = \|(\lambda_n, x_n)\| - \sqrt{\varepsilon_n}(h_n/\|(\lambda_n, x_n)\|)\|(\lambda_n, x_n)\|\,\|w^n\|.
\end{aligned} \tag{2.18}$$

Subtracting $\|(\lambda_n, x_n)\|$ from both sides, dividing by $h_n/\|(\lambda_n, x_n)\|$, and letting $n \to \infty$, we deduce that

$$\langle z, w\rangle \leq 0. \tag{2.19}$$

Since $w \in C_Y(f(\bar{u}))$ is arbitrary, $z \in N_Y(f(\bar{u}))$. The proof of Theorem 2.1 is complete. ■

This result is inadequate in two senses. The first is its dependence on the unknown sequence $\{u^n\}$ (although, as we shall see below, in some cases large subsets of $\liminf_{n\to\infty} \overline{\text{conv}}\,\partial(f_0, f)(u^n)$ that do not depend on $\{u^n\}$ can be identified). The second is that the vector (μ, z) may vanish, rendering (2.8) inoperative. In order to avoid this, we need additional assumptions on Y.

Let $\{\Delta_n; n = 1, 2, \ldots\}$ be a sequence of sets in E. Following [F1] we say that $\{\Delta_n\}$ has *finite codimension* in E if and only if there exists a closed subspace H with dim $H^{\perp} < \infty$ such that

$$\Delta_H = \bigcap_{n \geq 1} \Pi_H(\overline{\text{conv}}(\Delta_n)) \tag{2.20}$$

has nonempty interior in H, where Π_H is the projection of E into H.

The following result is a generalization of Lemma 5.6 in [F1].

Lemma 2.2. *Let the sequence $\{\Delta_n\}$ have finite codimension. Assume that the following condition holds:*

$$\bigcup_{n \geq 1} (I - \Pi_H)(\Delta_n) \text{ is bounded.} \tag{2.21}$$

Let $\{z_n\}$ be a sequence in E such that

$$0 < c \leq \|z_n\| \leq C, \tag{2.22}$$

$$\sup_{y \in \Delta_n} \langle z_n, y \rangle \leq \varepsilon_n \to 0. \tag{2.23}$$

Then every weakly convergent subsequence of $\{z_n\}$ has a nonzero limit.

Proof. We denote the subsequence by $\{z_n\}$ as well. Write

$$z_n = x_n + \tilde{x}_n \qquad (x_n \in H, \tilde{x}_n \in H^{\perp}).$$

Since $\{z_n\}$ is weakly convergent, so are $\{x_n\}$ and $\{\tilde{x}_n\}$. If $\tilde{x}_n \to \tilde{x} \neq 0$ we are through; thus we may assume that $\tilde{x}_n \to 0$, where convergence is strong on account of the finite dimension of $H^{\perp}$. It follows that

$$\|x_n\| \geq c' > 0 \tag{2.24}$$

for sufficiently large n. We note that (2.23) can be extended from Δ_n to $\overline{\text{conv}}(\Delta_n)$; also, if (2.21) is bounded, the set

$$\bigcup_{n \geq 1} (I - \Pi_H)(\overline{\text{conv}}(\Delta_n))$$

will be bounded as well. Using the fact that $\tilde{x}_n \to 0$ in (2.23) we deduce that for all $y^n \in \overline{\text{conv}}(\Delta_n)$

$$\begin{aligned}\langle x_n, \Pi_H(y^n) \rangle &= \langle z_n, y^n \rangle - \langle \tilde{x}_n, (I - \Pi_H)(y^n) \rangle \\ &\leq \sup_{y \in \overline{\text{conv}}(\Delta^n)} \langle z_n, y \rangle + \sup_{y \in \overline{\text{conv}}(\Delta^n)} \langle \tilde{x}_n, (I - \Pi_H) y \rangle \to 0. \end{aligned} \tag{2.25}$$

Accordingly, it follows that

$$\langle x_n, x \rangle \leq \delta_n \to 0 \qquad (x \in \Delta_H). \tag{2.26}$$

Let $B(\bar{x}, \rho) = \{x \in H; \|x - \bar{x}\| \leq \rho\}$ $(\rho > 0)$ be a ball in H contained in Δ_H. Then it follow from (2.26) that

$$\langle x_n, \bar{x} \rangle + \rho \|x_n\| = \langle x_n, \bar{x} + \rho x_n / \|x_n\| \rangle \leq \delta_n \to 0$$

which, in view of (2.24) shows that $\langle x_n, \bar{x} \rangle$ does not tend to zero. Accordingly, $\{x_n\}$ cannot be weakly convergent to zero. This ends the proof. ∎

Remark 2.3. Boundedness of the set (2.21) cannot in general be given up; see the example in [F1, p. 163]. However, we can bypass the boundedness assumption by requiring that for some $x \in H$ and $\rho > 0$ the sequence

$$\{\Delta_{n.x.\rho}\} = \{\Delta_n \cap B(x, \rho)\} \tag{2.27}$$

should be of finite codimension, where $B(x, \rho)$ is the ball in H with center x and radius ρ. Here, $B(x, \rho)$ can be replaced by $(H \oplus B^{\perp}(\tilde{x}, \rho))$, where $B^{\perp}(x, \rho)$ is the ball of center x and radius ρ in $H^{\perp}$. Of course, this is stronger than assuming finite codimension of the original sequence $\{\Delta_n\}$. See Lemma 5.5 of [F1] for a variant of Lemma 2.2 where $\{\Delta_n\}$ consists of only one set but where boundedness of (2.21) is unnecessary.

Theorem 2.4. *Let the assumptions of Theorem 2.1 hold. Assume that, for every sequence $\{u^n\} \subset V$ with $d(u^n, \bar{u}) \to 0$ fast enough and every sequence $\{y^n\} \subset Y$ with $y^n \to f(\bar{u})$, there exists $\rho > 0$ such that, either* (a) *the sequence*

$$\{\Delta_n\} = \{\{0\} \times \overline{\text{conv}}(K_Y(y^n) \cap B(0, \rho)) - \overline{\text{conv}}\, \partial(f_0, f)(u^n)\} \tag{2.28}$$

has finite codimension in $\mathbb{R} \times E$ and satisfies (2.21) *or* (b) *the sequence* (2.27) *has finite codimension for some $x \in H$, $\rho > 0$. Then the vector (μ, z) in* (2.8) *does not vanish.*

Proof. It suffices to note that (2.6) implies

$$\langle (\mu_n, z_n), (\kappa_n, \zeta_n) \rangle \leq \delta_n(1 + \rho) \to 0 \qquad ((\kappa_n, \zeta_n) \in \Delta_n)$$

and apply Lemma 2.2 and Remark 3.3. ∎

Theorem 2.4 may be called a *Kuhn–Tucker theorem* for (1.2)–(1.3); we may also call it a *maximum principle* in view of the principal application.

The analog of Theorem 2.1 for the time optimal problem is

Theorem 2.5. *Let the target set Y be closed and let $\{\bar{u}^n\}$ be a solution of the abstract time optimal control problem. Then there exist a sequence $\{\delta_n\}$ of positive numbers with $\delta_n \to 0$, a sequence $\{u^n\}$, $u^n \in V_n$, a sequence $\{y^n\} \subset Y$ such that*

$$d_n(u^n, \bar{u}) + \|y^n - \bar{y}\| \leq \delta_n, \tag{2.29}$$

and a sequence $\{z_n\} \subset E$ with

$$\|z_n\| = 1 \tag{2.30}$$

and such that, for every $\xi^n \in \overline{\text{conv}}\, \partial f_n(u^n)$ and every $w^n \in K_Y(y_n)$, we have

$$\langle z_n, \xi^n - w^n \rangle \geq \delta_n(1 + \|w^n\|). \tag{2.31}$$

Moreover, for every cluster point z of $\{z_n\}$ in the weak topology of E we have

$$z \in N_Y(\bar{y}) \tag{2.32}$$

and, for every $\xi \in \liminf_{n\to\infty} \overline{\text{conv}}\, \partial f_n(u^n)$, we have

$$\langle z, \xi \rangle \geq 0. \tag{2.33}$$

Proof. The proof is similar to that of Theorem 2.1. The functions F_n are defined by

$$F_n(u, y) = \|f_n(u) - y\| \tag{2.34}$$

in the spaces $V_n \times Y$, endowed with distance $\tilde{d}_n((u, y), (u', y')) = d_n(u, u') + \|y - y'\|$. We have

$$F_n(\bar{u}^n, \bar{y}) = \|f(\bar{u}^n) - \bar{y}\| = \varepsilon_n, \tag{2.35}$$

where $\varepsilon_n \to 0$ in view of (1.8). Applying Ekeland's variational principle we obtain an element $(u^n, y^n) \in V_n \times Y$ such that (2.11)–(2.13) hold (with d_n instead of d). Given $w^n \in K_Y(y^n)$, the sequences $\{h_k\}$, $\{w_k\}$ are the same in Theorem 2.1. The inequality corresponding to (2.14) is

$$\|f(v) - y^n - h_k w_k\| \geq \|f(u^n) - y^n\| - \sqrt{\varepsilon_n}(d_n(v, u^n) + h_k\|w_k\|). \tag{2.36}$$

If $\xi^n \in \partial f_n(u^n)$ we write (2.36) for $v = v(h_k)$ (see the comments after (2.14)), divide by h_k, and let $k \to \infty$. The result is (2.31), with $\delta_n = \sqrt{\varepsilon_n}$,

$$z_n = x_n/\|x_n\|, \tag{2.37}$$

$$x_n = f_n(u^n) - y^n. \tag{2.38}$$

To show that $z \in N_Y(\bar{y})$ we take $w \in C_Y(\bar{y})$, a sequence $\{h_n\}$ of positive numbers with

$$h_n/\|x_n\| \to 0 \qquad (n \to \infty) \tag{2.39}$$

and a sequence $\{w^n\}$ in E such that $w^n \to w$, $y^n + h_n w^n \in Y$. Then

$$\begin{aligned} F_n(u^n, y^n + h_n w^n) &= \{\|x_n\|^2 - 2h_n\langle f(u^n) - y^n, w^n\rangle + h_n^2\|w^n\|^2\}^{1/2} \\ &= \|x_n\|\{1 - 2(h_n/\|x_n\|)\langle z_n, w^n\rangle + h_n^2\|w^n\|^2/\|x_n\|^2\}^{1/2} \\ &= \|x_n\|\{1 - (h_n/\|x_n\|)\langle z_n, w^n\rangle + o(h_n/\|x_n\|)\} \\ &\geq F_n(u^n, y^n) - \sqrt{\varepsilon_n} h_n\|w^n\| \\ &= \|x_n\| - \sqrt{\varepsilon_n}(h_n/\|x_n\|)\|x_n\|\,\|w^n\| \end{aligned} \tag{2.40}$$

and we argue in the same way as after (2.18). This ends the proof. ∎

Theorem 2.6. *Let the assumptions of Theorem 2.5 hold. Assume that for every sequence $\{u^n\}$, $u^n \in V_n$, with $d_n(u^n, \bar{u}^n) \to 0$ fast enough and every sequence $\{y^n\} \subset Y$ with $y^n \to \bar{y}$ there exists $\rho > 0$ such that, either* (a) *the sequence*

$$\{\Delta_n\} = \{\overline{\text{conv}}(K_Y(y^n) \cap B(0, \rho)) - \overline{\text{conv}}\,\partial f_n(u^n)\} \tag{2.41}$$

has finite codimension in E and satisfies (2.21) *or* (b) *the sequence* (2.27) *has finite codimension for some $x \in H$, $\rho > 0$. Then the vector z in* (2.23) *does not vanish.*

Proof. Here, (2.31) implies that

$$\langle z_n, \zeta_n\rangle \leq \delta_n(1 + \rho) \to 0 \qquad (\zeta_n \in \Delta_n),$$

thus the result follows from Lemma 2.2 and Remark 2.3. ∎

When the target set Y is *convex* much additional information about the sequence $\{(\mu_n, z_n)\}$ in Theorem 2.1 and the sequence $\{z_n\}$ in Theorem 2.5 can be obtained. Moreover, the proofs of the corresponding results are simpler. We note first that, by virtue of (2.1), the Clarke normal cone $N_Y(\bar{y})$ consists of all $z \in E$ such that $\langle z, y - \bar{y}\rangle \leq 0$ for all $y \in Y$. Given $x \in E$ we denote by $y = \Pi_Y(x)$ the element of Y of minimum distance to Y, called the *projection of x into Y*; we have $x - y \in N_Y(\bar{y})$.

Theorem 2.7. *Let the target set Y be convex and closed, and let $\bar{u}$ be a solution of the abstract nonlinear programming problem* (1.2)–(1.3) *Then there exist a sequence $\{\delta_n\}$ of positive numbers, a sequence $\{u^n\} \subset V$ such that*

$$\delta_n \to 0, \qquad d(u^n, \bar{u}) \to 0, \tag{2.42}$$

and a sequence of vectors $\{(\mu_n, z_n)\} \subset \mathbb{R} \times E$ with

$$\mu_n \geq 0, \qquad z_n \in N_Y(\Pi_Y(f(u^n))), \qquad \|(\mu_n, z_n)\| = 1, \tag{2.43}$$

and such that, for every $(\eta^n, \xi^n) \in \overline{\text{conv}}\, \partial(f_0, f)(u^n)$, we have

$$\mu_n \eta^n + \langle z_n, \xi^n\rangle \geq \delta_n. \tag{2.44}$$

Proof. Let $\{\varepsilon_n\}$ be a sequence of positive numbers with $\varepsilon_n \to 0$. This time we use the function

$$F_n(u) = \{\max(0, f_0(u) - m + \varepsilon_n))^2 + d(f(u), Y)^2\}^{1/2}, \tag{2.45}$$

where m is the minimum is (1.2) and $d(x, Y)$ denotes the distance from x to Y in E. Since $\bar{u}$ is an optimal element, $f(\bar{u}) \in Y$ and $f_0(\bar{u}) = m$, hence

$$F_n(\bar{u}) = \varepsilon_n \leq \inf F_n + \varepsilon_n, \tag{2.46}$$

thus we can apply Ekeland's variational principle to deduce the existence of an element $u^n \in V$ such that

$$F_n(u^n) \leq F_n(\bar{u}) = \varepsilon_n, \tag{2.47}$$

$$d(u^n, \bar{u}) \leq \sqrt{\varepsilon_n}, \tag{2.48}$$

$$F_n(v) \geq F_n(u^n) - \sqrt{\varepsilon_n} d(v, u^n) \qquad (v \in V). \tag{2.49}$$

Define a function $G_n: V \to \mathbb{R}$ by

$$G_n(u) = \{\max(0, f_0(u) - m + \varepsilon_n))^2 + \|f(u) - \Pi_Y(f(u^n))\|^2\}^{1/2}. \tag{2.50}$$

Using (2.49) we obtain

$$\begin{aligned} G_n(v) &\geq \{(\max(0, f_0(v) - m + \varepsilon_n))^2 + d(f(v), Y)^2\}^{1/2} \\ &= F_n(v) \geq F_n(u^n) - \sqrt{\varepsilon_n} d(v, u^n) \\ &= \{(\max(0, f_0(u^n) - m + \varepsilon_n))^2 + \|f(u^n) - \Pi_Y(f(u^n))\|^2\}^{1/2} - \sqrt{\varepsilon_n} d(v, u^n) \\ &= G(u^n) - \sqrt{\varepsilon_n} d(v, u^n) \qquad (v \in V) \end{aligned} \tag{2.51}$$

Let (η^n, ξ^n) be an element of $\partial(f_0, f)(u^n)$ and let $v(h)$ be as in Theorem 2.1. We write

(2.51) for $v = v(h)$ and divide by h, obtaining

$$\mu_n \eta^n + \langle z_n, \xi^n \rangle \geq -\sqrt{\varepsilon_n} \tag{2.52}$$

with

$$(\mu_n, z_n) = (\lambda_n, x_n)/\|(\lambda_n, x_n)\|, \tag{2.53}$$

$$(\lambda_n, x_n) = (\max(0, f_0(u^n) - m + \varepsilon_n), f(u^n) - \Pi_Y(f(u^n))), \tag{2.54}$$

so that all the statements in Theorem 2.7 are satisfied with $\delta_n = \sqrt{\varepsilon_n}$. ■

We observe that (2.44) and the seemingly more general inequality (2.6) are equivalent here; in fact, $y^n = \Pi_Y(f(u^n))$ and $z_n \in N_Y(\Pi_Y(f(u^n)))$.

The counterpart of Theorem 2.7 for the abstract time optimal problem is

Theorem 2.8. *Let the target set Y be convex and closed, and let $\{\bar{u}^n\}$ be a solution of the abstract time optimal control problem. Then there exist a sequence $\{\delta_n\}$ of positive numbers, a sequence $\{u^n\} \subset V$ such that*

$$\delta_n \to 0, \qquad d(u^n, \bar{u}^n) \to 0, \tag{2.55}$$

and a sequence of vectors $\{z^n\} \subset E$ satisfying

$$z_n \in N_Y(\Pi_Y(f_n(u^n))), \qquad \|z_n\| = 1 \tag{2.56}$$

and such that, for every $\xi^n \in \overline{\text{conv}}\, \partial f_n(u^n)$, we have

$$\langle z_n, \xi^n \rangle \geq -\delta_n. \tag{2.57}$$

The proof is similar to that of Theorem 2.7; we use the functions $F_n(u) = d(f(u), Y)$, $G_n(u) = \|f(u) - \Pi_Y(f(u^n))\|$.

Remark 2.9. All of our results only involve elements which are close to the optimal element $\bar{u}$ in the distance of V. Accordingly, we can formulate a *local* version of the abstract nonlinear programming problem, where we replace the whole space V by the ball $B(\bar{u}, \delta)$ for δ sufficiently small, with no change in the results or proofs. This allows us, for instance, to dispense with global solvability conditions in differential systems (see Section 4, Lemma 4.1).

3. Nontriviality of the First Multiplier

In results such as Theorem 2.4 we are only guaranteed that $(\mu, z) \neq 0$. The case $\mu = 0$ is exceptional in that the necessary conditions do not involve the cost functional f_0. In the case $\mu > 0$, the multiplier rule obtained in Section 4 is called *normal*. We examine here a condition, introduced in [F8], that guarantees normality.

Let $\{u^n\}$ be a convergent sequence in V. Define

$$P(\{u^n\}) = \text{Cl}\left[\bigcup_{\lambda \geq 0} \lambda \Pi_E\left(\liminf_{n\to\infty} \overline{\text{conv}}\, \partial(f_0, f)(u^n)\right)\right], \tag{3.1}$$

where we denote by Π_E the canonical projection of $\mathbb{R} \times E$ into E.

Theorem 3.1. *Let the assumptions of Theorem 2.4 be satisfied and let $\bar{u}$ be a solution of the abstract optimal control problem. Assume that, for every sequence $\{u^n\} \subset V$ such that $d(u^n, \bar{u}) \to 0$ fast enough, we have*

$$N_Y(f(\bar{u})) \subseteq \mathrm{Cl}(C_Y(f(\bar{u})) - P(\{u^n\})). \tag{3.2}$$

Then, for every cluster point (μ, z) of the sequence $\{(\mu_n, z_n)\}$ in Theorem 2.4, we have

$$\mu > 0. \tag{3.3}$$

Proof. Let $\{u^n\}$ be the sequence in Theorem 2.4 corresponding to the multiplier (μ, z). If $\mu = 0$, then $z \neq 0$ and

$$\langle z, \xi \rangle \geq 0 \tag{3.4}$$

for all $\xi \in \Pi_E(\liminf_{n\to\infty} \overline{\mathrm{conv}}\, \partial(f_0, f)(u^n))$, thus for all $\xi \in P(\{u^n\})$. Since $z \in N_Y(f(\bar{u}))$ and (3.2) holds, there exist sequences $\{w_k\} \subset C_Y(f(\bar{u}))$ and $\{\xi_k\} \subset P(\{u^n\})$ such that $z = \lim_{k\to\infty}(w_k - \xi_k)$. Taking k sufficiently large, we may write $\xi_k = w_k - z + v_k$ ($\|v_k\| \leq \alpha\|z\|$) with $\alpha < 1$, so that

$$\langle z, \xi_k \rangle = \langle z, w_k \rangle - \|z\|^2 + \langle v_k, z \rangle \leq -(1-\alpha)\|z\|^2 \tag{3.5}$$

since $z \in N_Y(f(\bar{u}))$. This obviously produces a contradiction with (3.4). ■

Remark 3.2. Since $\mathrm{Cl}(C_Y(f(\bar{u})) - P(\{u^n\}))$ is a (closed, convex) cone, (3.2) will hold if

$$0 \in \mathrm{Int}(\mathrm{Cl}(C_Y(f(\bar{u})) - P(\{u^n\}))). \tag{3.6}$$

This last condition is equivalent to

$$\mathrm{Cl}(C_Y(f(\bar{u})) - P(\{u^n\})) = E. \tag{3.7}$$

We note finally that when $\mathrm{Int}(C_Y(f(\bar{u}))) \neq \varnothing$, (3.6) or (3.7) are equivalent to

$$\mathrm{Int}(C_Y(f(\bar{u}))) \cap P(\{u^n\}) \neq \varnothing. \tag{3.8}$$

In fact, it is obvious that (3.8) implies that $0 \in \mathrm{Int}(C_Y(f(\bar{u})) - P(\{u^n\}))$, which in turn implies (3.6). On the other hand, if (3.8) does not hold there exists $z \in E$, $z \neq 0$, such that

$$\langle z, w \rangle \leq \langle z, \xi \rangle \qquad (w \in C_Y(f(\bar{u})),\ \xi \in P(\{u^n\})).$$

It follows that $\langle z, v \rangle \leq 0$ for $v \in C_Y(f(\bar{u})) - P(\{u^n\})$ (*a fortiori* also for $v \in \mathrm{Cl}(C_Y(f(\bar{u})) - P(\{u^n\}))$, thus $0 \notin \mathrm{Int}(\mathrm{Cl}(C_Y(f(\bar{u})) - P(\{u^n\})))$.

Similar considerations apply to the case $\mathrm{Int}(P(\{u^n\})) \neq \varnothing$.

We note that the assumptions of Theorem 3.1 involve the point $f(\bar{u})$, which is in general not known in advance; thus in applications we must verify the assumptions for every possible $f(\bar{u})$, that is, for every point of Y.

4. Systems Described by Abstract Differential Equations

Consider the semilinear initial value problem

$$y'(t) = Ay(t) + f(t, y(t), u(t)) \qquad (0 \leq t \leq T), \tag{4.1}$$

$$y(0) = y^0, \tag{4.2}$$

where A is the infinitesimal generator of a strongly continuous semigroup $\{S(t); t \geq 0\}$ in a Hilbert space E and $f(t, y, u)$ is a function from $[0, T] \times E \times U$ into E, where U (the *control set*) is a bounded subset of another Hilbert space F. The *control space* $W(0, T; U)$ is the set of all (equivalence classes of) strongly measurable F-valued functions $u = u(t)$ defined in $0 \leq t \leq T$ and such that $u(t) \in U$ a.e. (almost everywhere). This space is equipped with the *Ekeland distance*

$$d(u, v) = \text{meas}\{t; u(t) \neq v(t)\} \tag{4.3}$$

(see [E1] and [E2]) which makes it complete. We assume that the function $f(t, y, u)$ has a Fréchet derivative $\partial_y f(t, y, u)$ with respect to y and that f (resp. $\partial_y f$) is continuous (resp. strongly continuous) and bounded on bounded subsets of $[0, T] \times E \times U$. Solutions of (4.1)–(4.2) are, by definition, solutions of the integral equation

$$y(t) = S(t)y_0 + \int_0^t S(t-\sigma)f(\sigma, y(\sigma), u(\sigma))\, d\sigma \qquad (0 \leq t \leq T). \tag{4.4}$$

The solution corresponding to u is called $y(t, u)$. Given u in $W(0, T; U)$ the hypotheses guarantee existence and uniqueness of a solution of (4.3) in an interval $0 \leq t \leq \bar{t}$, where possibly $\bar{t} < T$. For assumptions that guarantee existence in the whole interval see, for instance, [F2] and [F3]. However, we shall not require global existence.

Lemma 4.1. *Let $\bar{u}(\cdot) \in w(0, \bar{t}; U)$ be a control whose trajectory $y(t, \bar{u})$ exists in $0 \leq t \leq \bar{t}$. Then there exists $\varepsilon > 0$ such that if $u(\cdot) \in W(0, \bar{t}; U)$ and $d(u, \bar{u}) \leq \varepsilon$, then the trajectory $y(t) = y(t, u)$ also exists in $0 \leq t \leq \bar{t}$.*

The proof is an elementary application of Gronwall's inequality and is omitted. Lemma 4.1 in combination with Remark 2.9 allows us to work in a ball $B(\bar{u}, \varepsilon)$ instead of the whole space $W(0, \bar{t}; E)$ in case global solvability does not hold.

We apply the results in Sections 2 and 3 to the optimal control problem described in Section 1 with cost functional

$$y_0(t, u) = \int_0^t f_0(\sigma, y(\sigma, u), u(\sigma))\, d\sigma, \tag{4.5}$$

where $f_0(t, y, u)$ is a real-valued function defined in $[0, T] \times E \times U$ possessing a Fréchet derivate $\partial_y f_0(t, y, u)$ with f_0 (resp. $\partial_y f_0$) continuous (resp. strongly continuous) and bounded on bounded subsets of $[0, T] \times E \times U$. The space V is $W(0, T; U)$ and the maps f, f_0 are defined by (1.4). In the time optimal case, the sequence $\{f_n\}$ of maps is given by

$$f_n(u) = y(t_n, u), \tag{4.6}$$

where $\{t_n\}$ is a sequence in $0 \leq t \leq \bar{t}$ ($\bar{t}$ the optimal time) such that $t_n \to \bar{t}$. Continuity (in fact, Lipschitz continuity) of these maps is proved in [F1]. We define a set of variations $\xi(\bar{t}, s, u, v) \in \partial f(u)$ (with $0 \leq s \leq \bar{t}$ and $v \in U$) taking $v(h)(t) = u_{s,h,v}(t)$ in (2.2)–(2.3), where $u_{s,h,v}(t)$ is the *spike variation* of the control $u(\cdot)$ defined by $u_{s,h,v}(t) = v$ in $s - h < t \leq s$, $u_{s,h,v}(t) = u(t)$ elsewhere. It is proved in [F1] that if s is a left

Lebesgue point of the function $f(s, y(s, u), u(s))$ in $0 \le s \le \bar{t}$, then $\xi(\bar{t}, s, u, v)$ exists and equals

$$\xi(\bar{t}, s, u, v) = S(\bar{t}, s; u)\{f(s, y(s, u), v) - f(s, y(s, u), u(s))\}. \tag{4.7}$$

where $S(t, s, u)$ is the solution operator of the linear equation

$$z'(t) = \{A + \partial_y f(t, y(t, u), u(t))\} z(t), \tag{4.8}$$

that is, the only strongly continuous solutions of the operator equation

$$S(t, s; u)z = S(t - s)z + \int_0^t S(t - \sigma)\partial_y f(\sigma, y(\sigma, u), u(\sigma))S(\sigma, s; u)z \, d\sigma$$

in $0 \le s \le t \le T$.

Note that $\xi(t, s, u, v)$ can only depend on the equivalence class of u; thus the variation $\xi(t, s, u, v)$ will exist when s is a Lebesgue point of any function in the equivalence class.

Variations of $y_0(\bar{t}, u)$ are constructed in the same way and denoted by $\xi_0(t, s, u, v)$; it is also shown in [F1] that if s is a left Lebesgue point of both $f(s, y(s, u), u(s))$ and $f_0(s, y(s, u), u(s))$ in $0 \le s \le \bar{t}$, then $\xi_0(t, s, u, v)$ exists and equals

$$\begin{aligned}\xi_0(\bar{t}, s, u, v) &= f_0(s, y(s, u), v) - f_0(s, y(s, u), u(s)) \\ &\quad + \int_s^{\bar{t}} \langle \partial_y f_0(\sigma, y(\sigma, u), u(\sigma)), \xi(\sigma, s, u, v)\rangle \, d\sigma.\end{aligned} \tag{4.9}$$

In order to interpret the maximum principle (2.8) we must identify elements of $\limsup_{n\to\infty} \overline{\text{conv}}\, \partial(f_0, f)(u^n)$, or $\limsup_{n\to\infty} \overline{\text{conv}}\, \partial f_n(u^n)$ for (2.33). We do this below.

Lemma 4.2. *Let $\{t_n\}$ be a sequence in $0 \le t \le \bar{t}$, and let $\{u^n\}$ be a sequence in* $W(0, t_n; U)$ such that

$$\sum_{n=1}^{\infty} (\bar{t} - t_n) < \infty, \qquad \sum_{n=1}^{\infty} d_n(u^n, \bar{u}) < \infty. \tag{4.10}$$

Then there exists a set $e = e(\{t_n\}, \{u^n\})$ of full measure in $0 \le s \le \bar{t}$ such that

$$\xi(t_n, s, u^n, v) \to \zeta(\bar{t}, s, \bar{u}, v) \qquad (s \in e), \tag{4.11}$$

$$\xi_0(t_n, s, u^n, v) \to \zeta_0(\bar{t}, s, \bar{u}, v) \qquad (s \in e). \tag{4.12}$$

Proof. By (4.10), if $d_n = \{t; 0 \le t \le t_n, \bar{u}(t) \ne u^n(t)\} \cup (t_n, \bar{t}]$, then $\sum \text{meas}(d_n) < \infty$. Accordingly, the set

$$d = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} d_n \tag{4.13}$$

has measure zero. Let $e(u^n)$ (resp. $e_0(u^n)$) be the set of left Lebesgue points of $f(\sigma, y(\sigma, u^n), u^n(\sigma))$ (resp. $f_0(\sigma, y(\sigma, u^n), u^n(\sigma))$) in $[0, t_n]$, and let $e_n = (e(u^n) \cap e_0(u^n)) \cup [t_n, \bar{t}]$. Define

$$e = \left(\bigcap_{n=1}^{\infty} e_n\right) \backslash d. \tag{4.14}$$

Obviously, e is total in $[0, \bar{t}]$. Assume that $s \in e$. Then there exists n_o such that $s \in e(u^n) \cap e_0(u^n)$ and $s \notin d_n$ $(n \geq n_0)$. Accordingly, (4.7) and (4.9) hold for each u^n and $u^n(s) = \bar{u}(s)$ for $n \geq n_0$. To take limits in (4.7) and (4.9) we use continuity of the solution operator. $S(t, s; u)$ is continuous in $0 \leq s \leq t \leq T$, $u \in W(0, T; U)$, u measured in the Ekeland distance (see [F1]). ■

Lemma 4.2 shows that if $t_n = \bar{t}$ and $\{u^n\}$ satisfies (4.10), then

$$(\xi_0(\bar{t}, s, u^n, v), \xi(\bar{t}, s, \bar{u}, v)) \in \liminf_{n \to \infty} \partial(f_0, f)(u^n) \tag{4.15}$$

for every $s \in e$. Likewise, if $\{t_n\}$ and $\{u^n\}$ satisfy (4.10),

$$\xi(\bar{t}, s, \bar{u}, v) \in \liminf_{n \to \infty} \partial f_n(u^n). \tag{4.16}$$

Using (4.15) we deduce from (2.8) that

$$\mu\xi(\bar{t}, s, u, v) + \langle z, \xi(\bar{t}, s, u, v)\rangle \geq 0 \qquad (s \in e, v \in V). \tag{4.17}$$

It is shown in [F1] that (4.17) can be written in the more familiar form

$$\begin{aligned} &\mu f_0(s, y(s, \bar{u}), \bar{u}(s)) + \langle z(s), f(s, y(s, \bar{u})), \bar{u}(s)\rangle \\ &\qquad = \min_{v \in U} \{\mu f_0(s, y(s, \bar{u}), v) + \langle z(s), f(s, y(s, \bar{u}), v)\rangle\}, \end{aligned} \tag{4.18}$$

where $z(s)$ is the solution of the adjoint initial value problem

$$z'(s) = -(A^* + \partial_y f(s, y(s, \bar{u}), \bar{u}(s))^*)z(s) - \mu\partial_y f_0(s, y(t, \bar{u}), u(s)) \qquad (0 \leq s \leq \bar{t}), \tag{4.19}$$

$$z(\bar{t}) = z. \tag{4.20}$$

In the time optimal problem, we deduce from (2.33) that

$$\langle z, \xi(\bar{t}, s, \bar{u}, v)\rangle \geq 0 \qquad (s \in e, v \in V) \tag{4.21}$$

which can be written in the form

$$\langle z(s), f(s, y(s, \bar{u}), \bar{u}(s))\rangle = \min_{v \in U} \langle z(s), f(s, y(s, \bar{u}), v)\rangle \tag{4.22}$$

for $s \in e$, where $z(s)$ is given by (4.19)–(4.20) with $f_0 = 0$.

We look below at conditions (2.28) (resp. (2.41)) that guarantee that the multiplier (μ, z) in (2.8) (resp. the multiplier z in (2.33)) is nontrivial. By virtue of (4.7), the set $\Pi_E(\overline{\text{conv}}\, \partial(f_0, f)(u^n))$ will contain all elements of the form

$$\bar{t}^{-1} \int_0^{\bar{t}} S(\bar{t}, s; u^n)\{f(s, y(s, u^n), v(s)) - f(s, y(s, u^n), u^n(s))\}\, ds \tag{4.23}$$

with $v(\cdot) \in W(0, \bar{t}; U)$.

For $\tilde{u} \in W(0, \bar{t}; U)$ consider the control system

$$z'(s) = \{A + \partial_y f(s, y(s, \tilde{u}), \tilde{u}(s))\}z(s) + v(s) \qquad (0 \leq s \leq \bar{t}), \tag{4.24}$$

$$z(0) = w, \tag{4.25}$$

for a fixed control $\tilde{u}$. Given a control space V and a set $Z \in E$ denote by $R(\tilde{t}, \tilde{u}; V, Z)$ the reachable subspace of (4.24)–(4.25) corresponding to all $v(\cdot) \in V$ and $w \in Z$. Then the set of all elements of the form (4.23) is $\bar{t}^{-1}R(\bar{t}, u^n; V^n, \{0\})$, where the control space V^n consists of all functions of the form

$$v(s) = f(s, y(s, u^n), u(s)) - f(s, y(s, u^n), u^n(s)) \tag{4.26}$$

with $u(s) \in W(0, \bar{t}; U)$. In certain cases involving abstract hyperbolic equations it is possible to show that all the $R(\bar{t}, u^n; V^n; \{0\})$ contain a fixed open set; thus the conditions in Theorem 2.4 are satisfied for any target set Y and the multiplier in (2.8) is nontrivial. On the other hand, if Y is "sufficiently large," Theorem 2.4 will hold without conditions on $R(\bar{t}, u^n; V^n; \{0\})$. For details, see Section 4 of [F1].

The preceding observations apply as well to the time optimal problem.

In relation with the results in Section 3, we have

Theorem 4.3. *Under the conditions of Theorem* 2.6, *assume that*

$$N_Y(y(\bar{t}, \bar{u})) \subseteq \mathrm{Cl}(C_Y(y(\bar{t}, \bar{u})) - R(\bar{t}, \bar{u}; \bar{V}, \{0\})), \tag{4.27}$$

where $\bar{V}$ consists of all integrable functions $v(s)$ satisfying

$$v(s) \in \mathrm{Cl}\left[\bigcup_{\lambda \geq 0} \lambda\{\overline{\mathrm{conv}} f(s, y(s, \bar{u}), U) - f(s, y(s, \bar{u}), \bar{u}(s))\}\right]. \tag{4.28}$$

Then (2.8) *holds with* $\mu = 1$.

Proof. It is enough to show that $P(\{u^n\})$, defined as in (3.1), contains $R(\bar{t}, \bar{u}; \bar{V}, \{0\})$; this follows from the definition of $P(\{u^n\})$ and from an approximation argument. ∎

Remark 4.4. Let

$$U(s) = \overline{\mathrm{conv}} f(t, y(s, \bar{u}), U). \tag{4.29}$$

Using the definition of a contingent cone, we show that the control space $\bar{V}$ in Theorem 4.3 consists of all integrable functions $v(s)$ such that

$$v(s) \in K_{U(s)}(f(s, y(s, \bar{u}), \bar{u}(s))) \qquad \text{a.e. in} \quad 0 \leq s \leq \bar{t}. \tag{4.30}$$

Remark 4.5. The treatment of optimal control problems as nonlinear programming problems (1.2)–(1.3) does not take advantage of the fact that the arrival time $\bar{t}$ may not be fixed (except of course in the time optimal problem). For control systems described by ordinary differential equations, the difference between the free-time and the fixed-time problem is given by the fact that the Hamiltonian is zero. In infinite-dimensional spaces, the Hamiltonian cannot in general be defined for systems described by (4.1)–(4.2) and even when it can, it may not be directly related to the maximum principle under use; thus the vanishing of the Hamiltonian can only be proved in very special situations (see [F7] for details).

We show finally how the results in Section 2 apply to a control system where not only the control but the initial condition is a variable. The system is defined by (4.1) with the variable initial condition

$$y(0) = x \in X, \tag{4.31}$$

where X is a closed set in E. Trajectories of this control system are denoted by $y(t, x, u)$. We take a cost functional of the form

$$y_0(t, x, u) = g(t, y(t, x, u)), \tag{4.32}$$

where $g(t, y)$ has a Fréchet derivative $\partial_y g(t, y)$ in $[0, T] \times E$ and $g(t, y)$, $\partial_y g(t, y)$ are continuous. We minimize $y_0(t, x, u)$ among all controls $u \in W(0, t; U)$ and initial conditions $x \in X$ with target condition

$$y(t, x, u) \in Y, \tag{4.33}$$

where Y is a closed set. We assume that an optimal initial condition–control pair $(\bar{x}, \bar{u})$ exists and denote by $\bar{t}$ the (not necessarily unique) optimal arrival time to the target, and by $\bar{y}$ the hitting point $y(\bar{t}, \bar{x}, \bar{u})$. To apply Theorem 2.1, the basic space is $V = X \times W(0, \bar{t}; U)$ with the product distance, and f, f_0 are

$$f(x, u) = y(\bar{t}, x, u), \qquad f_0(x, u) = y_0(\bar{t}, x, u). \tag{4.34}$$

Let x belong to X and let w be an element of the contingent cone $K_X(x)$; moreover, let $v \in U$ and $u \in W(0, \bar{t}; U)$. Let s be a left Lebesgue point of $f(s, y(s, x, u), u(s))$. Then, if we set

$$\eta = \langle \partial_y g(\bar{t}, y(\bar{t}, x, u)), \xi \rangle \tag{4.35}$$

$$\xi = S(\bar{t}, 0; u)w + S(\bar{t}, s, u)\{f(s, y(s, x, u), v) - f(s, y(s, x, u), u(s))\}, \tag{4.36}$$

$(\eta, \xi) \in \partial(f_0, f)(x, u)$. Theorem 2.1 produces a multiplier (μ, z) such that (2.8) holds for every $(\eta, \xi) \in \overline{\text{conv}}\, \partial(f_0, f)(x^n, u^n)$ and every $w^n \in K_Y(y^n)$, where $x^n \to \bar{x}$, $y^n \to \bar{y}$, $u^n \to \bar{u}$ as fast as desired. Denote by $S_X(\bar{x})$ the set

$$S_X(\bar{x}) = \liminf_{x \to \bar{x}, x \in X} \overline{\text{conv}}\, K_X(x). \tag{4.37}$$

Let $w \in S_X(\bar{x})$. Consider a sequence $\{x_n\}$ in X with $x_n \to \bar{x}$ and a sequence $\{w_n\}$, $w_n \in K_Y(x^n)$, with $w_n \to w$. Taking limits in (4.35)–(4.36) we deduce that the element (η, ξ), where

$$\eta = \langle \partial_y g(\bar{t}, y(\bar{t}, \bar{x}, \bar{u})), \xi \rangle, \tag{4.38}$$

$$\xi = S(\bar{t}, 0; \bar{u})w + S(\bar{t}, s, \bar{u})\{f(s, y(s, \bar{x}, \bar{u}), v) - f(s, y(s, \bar{x}, \bar{u}), \bar{u}(s))\}, \tag{4.39}$$

belongs to $\liminf_{n\to\infty} \partial(f_0, f)(\bar{t}, x^n, u^n)$ for almost all s in $0 \le s \le \bar{t}$. Inequality (2.8) can be reworked in the form

$$\langle \mu\partial_y g(\bar{t}, y(\bar{t}, \bar{x}, \bar{u})) + z, S(\bar{t}, 0, \bar{u})w + S(\bar{t}, s, \bar{u})f(s, y(s, y(s, \bar{x}, \bar{u}), \bar{u}(s)) \rangle$$
$$= \min_{\substack{v \in U \\ w \in S_X(\bar{x})}} \langle \mu\partial_y g(\bar{t}, y(\bar{t}, \bar{x}, \bar{u})) + z, S(\bar{t}, 0, \bar{u})w + S(\bar{t}, s, \bar{u})f(s, y(s, \bar{x}, \bar{u}), v) \rangle.$$

Finally we note that assumption (2.20), guaranteeing $(\mu, z) \neq 0$, can be checked as in the observations following (4.25): we show that the set $\Pi_E(\overline{\text{conv}}\, \partial(f_0, f)(t, x^n, u^n))$ contains $R(\bar{t}, u^n; V^n, K_X(x^n))$. Finally, the following analog of Theorem 4.2 holds:

if $\bar{V}$ is the control space defined there and

$$N_Y(y(\bar{t}, \bar{x}, \bar{u})) \subseteq \mathrm{Cl}(C_Y(y(\bar{t}, \bar{x}, \bar{u})) - R(\bar{t}, \bar{u}; \bar{V}; S_X(\bar{x}))),$$

then the vector (μ, z) satisfies $\mu > 0$.

Remark 4.6. The results in this section extend to the case where U is unbounded and, controls are taken from the space $W^1(0, \bar{t}; U) = W(0, \bar{t}; U) \cap L^1(0, \bar{t}; U)$. Assume an optimal control $\bar{u}(\cdot)$ exists. It is easy to see that Lemma 4.1 can be extended to this situation. Given now the ε provided by Lemma 4.1 and an arbitrary constant N, consider the subspace $V_{\varepsilon,N}$ of $W(0, \bar{t}; U)$ consisting of all $u(\cdot)$ with

$$d(u, \bar{u}) \leq \varepsilon, \qquad \|u(t)\| \leq N \qquad \text{where} \quad u(t) \neq \bar{u}(t).$$

Then V_N is a complete matric space. Applying (2.8) into the system defined by (4.1)–(4.2) we obtain (4.18) for $\|v\| \leq N$, thus for all N, since N is arbitrary.

5. Suboptimal Controls: the Sequence Minimum Principle

For the control system (1.4), a sequence $\{u^n(\cdot)\}$ of controls, $u^n \in W(0, t_n, U)$, is called *(t_n, ε_n)-suboptimal* if

$$\mathrm{dist}(y(t_n, u), Y) \leq \varepsilon_n, \qquad y_0(t_n, u) \leq m + \varepsilon_n, \tag{5.1}$$

with $\varepsilon_n \to 0$, where m is the minimum in (1.2). The same definition applies to the time optimal problem.

In the last two sections we study the problem of convergence of sequences of suboptimal controls. We formulate the problem in the following abstract way: Let $\{V_n\}$ be a sequence of complete metric spaces, let E be a Hilbert space, and let $\{f_n\}$, $\{f_{0n}\}$ be two sequences of continuous functions,

$$f_n: V_n \to E, \qquad f_{0n}: V_n \to \mathbb{R}. \tag{5.2}$$

Let Y be a subset of E, and

$$m = \inf\{f_{0n}(u); n \geq 1, u \in V_n, f_n(u) \in Y\}. \tag{5.3}$$

Characterize the sequences $\{u^n\}$, $u^n \in V_n$, such that

$$f_{0n}(u^n) \leq m + \varepsilon_n, \tag{5.4}$$

$$\mathrm{dist}(f_n(u^n), Y) \leq \varepsilon_n \to 0. \tag{5.5}$$

A sequence $\{u^n\}$ as above is called *$\{\varepsilon_n\}$-suboptimal.* For the system (1.4) $V_n = W(0, t_n, U)$, $f_n(u) = y(t_n, u)$, and $f_{0n}(u) = y_0(t_n, u)$.

The corresponding version for the time optimal problem is: Let $\{V_n\}$ be a sequence of complete metric spaces, let E be a Hilbert space, let $\{f_n\}$ be a sequence of continuous functions,

$$f_n: V_n \to E, \tag{5.6}$$

and let Y be a subset of E. Assume that

$$f_n(V_n) \cap Y = \varnothing. \tag{5.7}$$

Characterize the sequences $\{u^n\}$, $u^n \in V_n$, that satisfy

$$\operatorname{dist}(f_n(u^n), Y) \le \varepsilon_n \to 0. \tag{5.8}$$

For (1.4), $V_n = W(0, t_n, U), f_n(u) = y(t_n, u)$, where $t_n < \bar{t} =$ optimal time. The problem is the same as the abstract time optimal problem except for condition (5.8); the corresponding condition in Section 1 is

$$f_n(u^n) \to \bar{y} \in Y. \tag{5.9}$$

However, the difference is not essential, since we will assume (5.9) in Section 6 in order to obtain convergence results.

The strategy to prove convergence is essentially that of [F2]–[F4], although in these works the conditions on the target set Y are very restrictive, whereas here Y is just a closed set. The proof consists of three steps. In the first we establish a *sequence maximum principle* (Theorem 5.1 below), which is a separate approximate maximum principle for each of the members of the suboptimal sequence $\{u^n\}$. Each approximate maximum principle will involve a multiplier (μ_n, z_n) (a multiplier z_n in the time optimal case). The second step (which we treat in next section) consists in upgrading the sequence maximum principle to a *convergence principle*, where we prove that (a subsequence of) the sequence of multipliers $\{(\mu_n, z_n)\}$ or $\{z_n\}$ is convergent; the convergence principle is *weak* or *strong* depending on whether the convergence of the sequence of multipliers is weak or strong. In the third step, we translate convergence of the sequence of multipliers into convergence of the sequence $\{u^n\}$ for the system (1.4).

In our first result, the assumptions on the target set are those in Theorem 2.1.

Theorem 5.1. *Let the target set Y be closed. Let $\{u^n\}$ be a $\{\varepsilon_n\}$-suboptimal sequence, and let $\delta_n = 8^{1/4}\varepsilon_n{}^{1/2}$. Then there exist a sequence $\{\tilde{u}^n\}$, $\tilde{u}^n \in V_n$, a sequence $\{\tilde{y}^n\} \subset Y$ such that*

$$d_n(\tilde{u}^n, u^n) + \|\tilde{y}^n - f_n(\tilde{u}^n)\| \le \delta_n, \tag{5.10}$$

and a sequence $\{(\mu_n, z_n)\} \subset \mathbb{R} \times E$ with

$$\mu_n \ge 0, \qquad \|(\mu_n, z_n)\| = 1, \tag{5.11}$$

and such that if $(\eta^n, \zeta^n) \in \partial(f_{0n}, f_n)(\tilde{u}^n)$ and $w^n \in K_Y(\tilde{y}^n)$, then

$$\mu_n \eta^n + \langle z_n, \zeta^n - w^n \rangle \ge -\delta_n(1 + \|w^n\|). \tag{5.12}$$

Proof. The proof of Theorem 5.1 is essentially the same as that of Theorem 2.1. The role of the functions (2.9) is played by

$$F_n(u, y) = \{\max(0, f_{0n}(u) - m + \varepsilon_n))^2 + \|f_n(u) - y\|^2\}^{1/2} \tag{5.13}$$

in the spaces $V_n \times Y$. Again, each F_n is continuous and positive. In view of (5.5) there exists a sequence $\{y^n\} \subset Y$ such that

$$\|f_n(u^n) - y^n\| \le 2\varepsilon_n, \tag{5.14}$$

hence

$$F_n(u^n, y^n) \le ((2\varepsilon_n)^2 + (2\varepsilon_n)^2)^{1/2} = \sqrt{8}\varepsilon_n = \delta_n^2. \tag{5.15}$$

Using Ekeland's variational principle we deduce the existence of an element $(\tilde{u}^n, \tilde{y}^n) \in V_n \times Y$ such that

$$F_n(\tilde{u}^n, \tilde{y}^n) \leq F_n(u^n, y^n) \leq \delta_n^2, \tag{5.16}$$

$$d_n(\tilde{u}^n, u^n) + \|\tilde{y}^n - f_n(\tilde{u}^n)\| \leq \delta_n, \tag{5.17}$$

$$F_n(v, y) \geq F_n(\tilde{u}^n, \tilde{y}^n) - \delta_n(d_n(v, \tilde{u}^n) + \|y - \tilde{y}^n\|) \qquad ((v, y) \in V_n \times Y). \tag{5.18}$$

The sequence $\{\tilde{u}^n\}$ is that claimed in the statement of Theorem 5.1. The rest of the proof is essentially the same as the that of the first part of Theorem 2.1, thus we omit the details: the vector (μ_n, z_n) is

$$(\mu_n, z_n) = (\lambda_n, x_n)/\|(\lambda_n, x_n)\|, \tag{5.19}$$

with

$$(\lambda_n, x_n) = (f_{0n}(\tilde{u}^n) - m + \varepsilon_n, f_n(\tilde{u}^n) - \tilde{y}^n). \quad \blacksquare \tag{5.20}$$

The following result corresponds to the time optimal problem.

Theorem 5.2. *Let the target set Y be closed. Let $\{u^n\}$ be a $\{\varepsilon_n\}$-suboptimal sequence and let $\delta_n = 8^{1/4}\varepsilon_n^{1/2}$. Then there exists a sequence $\{\tilde{u}^n\}$, $\tilde{u}^n \in V_n$, a sequence $\{\tilde{y}^n\} \subset Y$ such that*

$$d_n(\tilde{u}^n, u^n) + \|\tilde{y}^n - f_n(u^n)\| \leq \delta_n, \tag{5.21}$$

and a sequence $\{z_n\} \subset E$ such that

$$\|z_n\| = 1, \tag{5.22}$$

and such that, for $\xi^n \in \overline{\text{conv}}\, \partial f_n(\tilde{u}^n)$ and $w^n \in K_Y(\tilde{y}^n)$, we have

$$\langle z_n, \xi^n - w^n \rangle \geq -\delta_n(1 + \|w^n\|). \tag{5.23}$$

The proof follows closely that of Theorem 4.3 and we omit it; the basic functions are defined by (2.34). The vector z_n is

$$z_n = x_n/\|x_n\|, \tag{5.24}$$

with

$$x_n = f_n(\tilde{u}^n) - \tilde{y}^n. \tag{5.25}$$

Remark 5.3. As in Section 2, simplified proofs can be used when the target set Y is convex, and we obtain additional information on the multiplier z_n. In Theorem 5.1 we use the functions

$$F_n(u) = \{\max(0, f_{0n}(u) - m + \varepsilon_n))^2 + d(f(u), Y)^2\}^{1/2} \tag{5.26}$$

instead of (5.13). The sequence $\{y^n\}$ is unnecessary; we obtain simply a sequence $\{\tilde{u}^n\}$, $\tilde{u}^n \in V_n$, such that

$$d_n(\tilde{u}^n, u^n) \leq 2\sqrt{\varepsilon_n} \tag{5.27}$$

and a sequence $\{(\mu_n, z_n)\} \subset \mathbb{R} \times E$ such that (5.11) and (5.12) hold with $w^n = 0$; moreover,

$$z_n \in N_Y(\Pi_Y(f_n(\tilde{u}^n))). \tag{5.28}$$

The vector (μ_n, z_n) is given by (5.19)–(5.20) with $\tilde{y}^n = \Pi_Y(f(\tilde{u}^n))$. The same observations apply to Theorem 5.2: in particular, (5.28) holds.

6. Convergence Principles

In this section we study convergence of (suitable subsequences of) the multiplier sequence $\{(\mu_n, z_n)\}$ in Theorem 5.1 and of the sequence $\{\mu_n\}$ in Theorem 5.2. In case of weak convergence, we show that the limit multipliers do not vanish.

In all the results below, we need the following compactness assumption.

Let $\{\tilde{u}^n\}$ be an arbitrary sequence, $\tilde{u}^n \in V_n$. Then there exists a subsequence of $\{\tilde{u}^n\}$ (denoted by the same symbol) and an element $\bar{y} \in Y$ such that

$$f_n(\tilde{u}^n) \to \bar{y} \qquad \text{strongly in } E. \tag{6.1}$$

Our first result upgrades Theorem 5.1 to a weak convergence principle.

Theorem 6.1. *Let the target set Y be closed, and let the compactness assumption hold. Assume in addition that for every sequence $\{\tilde{u}^n\}$, $\tilde{u}^n \in V_n$, and for every convergent sequence $\{\tilde{y}^n\} \subset Y$ there exist subsequences (denoted by the same symbols) and a number $\rho > 0$ such that, either* (a) the sequence

$$\{\Delta_n\} = \{\{0\} \times \overline{\text{conv}}(K_Y(\tilde{y}^n) \cap B(0, \rho)) - \overline{\text{conv}}\, \partial(f_{0n}, f_n)(\tilde{u}^n)\} \tag{6.2}$$

has finite codimension in $\mathbb{R} \times E$ and satisfies (2.21) *or* (b) *the sequence* (2.27) *has finite codimension for some $x \in H$, $\rho > 0$.*

Let $\{(\mu_n, z_n)\}$ be the sequence constructed in Theorem 5.1. Then there exists a subsequence (which we denote with the same symbol) such that

$$(\mu_n, z_n) \to (\mu, z) \neq 0 \qquad \text{weakly in} \quad \mathbb{R} \times E, \tag{6.3}$$

where

$$\mu \geq 0, \qquad z \in N_Y(\bar{y}). \tag{6.4}$$

Proof. It is essentially the same as that of Theorem 2.4. We start with the sequences $\{\tilde{u}^n\}$ and $\{\tilde{y}^n\}$ in Theorem 5.1; note that in view of (6.1) and of (5.10) we may assume that $\{\tilde{y}^n\}$ is convergent so that, selecting further subsequences if necessary, we ensure that the sequence $\{\Delta_n\}$ in (6.2) satisfies the assumptions in Theorem 6.1. We obtain from (5.12)

$$\langle (\mu_n, z_n), (\eta^n, \xi^n) \rangle \leq \delta_n(1 + \rho) \to 0 \qquad ((\eta^n, \xi^n) \in \Delta_n),$$

thus the result is a consequence of Lemma 2.2 and Remark 2.3. ■

The companion result for the abstract time suboptimal problem is

Theorem 6.2. *Let the target set Y be closed, and let the compactness assumption hold. Assume in addition that for every sequence $\{\tilde{u}^n\}$, $\tilde{u}^n \in V_n$, and for every convergent sequence $\{\tilde{y}^n\} \subset Y$ there exist subsequences (denoted by the same symbols) and a number $\rho > 0$ such that, either* (a) *the sequence*

$$\{\Delta_n\} = \{\overline{\text{conv}}(K_Y(\tilde{y}^n) \cap B(0, \rho)) - \overline{\text{conv}}\, \partial f_n(\tilde{u}^n)\} \tag{6.5}$$

has finite codimension in E and satisfies (3.21), *or* (b) *the sequence* (3.27) *has finite codimension for some $x \in H$, $\rho > 0$. Let $\{z_n\}$ be the sequence constructed in Theorem 5.2. Then there exists a subsequence (which we denote with the same symbol) such that*

$$z_n \to z \neq 0 \qquad \text{weakly in } E \tag{6.6}$$

with

$$z \in N_Y(\bar{y}). \tag{6.7}$$

The proof is similar to that of Theorem 6.1.

Let Z be a convex set in a Hilbert space H and let $\bar{y}$ be a point on the boundary of Z. We say that Z is *flat at* $\bar{y}$ if and only if the normal cone $N_Y(\bar{y})$ at $\bar{y}$ is a half-line, that is, for some $\zeta \in H$ we have

$$N_Y(\bar{y}) = (Y - \bar{y})^- = \{\lambda\zeta; \lambda \geq 0\}. \tag{6.8}$$

The vector ζ (normalized to $\|\zeta\| = 1$) is called the *outer unit normal vector to Y at $\bar{y}$*. We say that Z is *strongly flat* at $\bar{y}$ if and only if it is flat and, for every θ, $0 \leq \theta < \pi/2$, there exists $b = b(\theta) > 0$ such that

$$\bar{y} + \Gamma(-\xi, \theta, b) \subset Y, \tag{6.9}$$

where $\Gamma(z, \theta, b)$ is defined by $\Gamma(z, \theta, b) = \{y \in H; \langle z, y\rangle \geq \|y\| \cos\theta, \|y\| \leq b\}$. Obviously, if Y is strongly flat at $\bar{y}$, it is flat at $\bar{y}$. The converse may not be true: if $H = l^2$ and Y is the convex set of all sequences $\{x_n; n \geq 0\}$ with $x_0 \geq \max(0, x_n - 1/n)$, then Y is flat (but not strongly flat) at 0.

Lemma 6.3. *Let $\Delta \subseteq H$ be strongly flat at 0. Then, if $\{z_n\}$ satisfies*

$$\|z_n\| \to r \tag{6.10}$$

for some $r \geq 0$ and

$$\langle z_n, y\rangle \leq \varepsilon_n \to 0 \qquad (y \in \Delta), \tag{6.11}$$

we have

$$z_n \to r\xi \qquad \text{strongly}, \tag{6.12}$$

where ξ is the normal vector to Δ at 0.

Proof. It is enough to consider the case $r = 1$. Also, by uniqueness of the limit, it suffices to show (6.12) for a suitable subsequence of $\{z_n\}$. Noting that (6.9) must be satisified at $\bar{y} = 0$ for $\theta < \pi/2$ arbitrary, we deduce that for every such θ there exists $\rho > 0$ with $B(-\rho\zeta, \rho \sin\theta) = \{y \in H; \|y + \rho\zeta\| \leq \rho \sin\theta\} \subset \Gamma(-\zeta, \theta, b) \subset \Delta$.

Accordingly, we have

$$\langle z_n, -\rho\zeta\rangle + \rho \sin\theta \|z_n\| = \langle z_n, -\rho\zeta + \rho \sin\theta z_n/\|z_n\|\rangle \le \delta_n. \tag{6.13}$$

Using (6.13) for a sequence $\{\theta_m\}$, $\theta_m \to \pi/2$, we deduce that there exists a subsequence of $\{z_n\}$ (named in the same way) such that

$$\langle z_n, -\zeta\rangle \to -1$$

which ends the proof of Lemma 6.3. ∎

Theorem 6.4. *Let $\{\Delta_n\}$ be the sequence of sets in Theorem 6.1. Assume that the set*

$$\Delta = \bigcap_{n \ge 1} \overline{\mathrm{conv}}(\Delta_n) \tag{6.14}$$

is strongly flat at 0. Let $\{(\mu_n, z_n)\}$ be the sequence constructed in Theorem 5.1. Then

$$(\mu_n, z_n) \to (\mu, z) \ne 0 \qquad \text{strongly in} \quad \mathbb{R} \times E. \tag{6.15}$$

Proof. The result follows from Lemma 6.4; in fact, since $\|(\mu_n, z_n)\| = 1$ we may assume that (for a subsequence) $\mu_n \to \mu$ so that $\|z_n\| \to r = 1 - \mu$.

Similar reasoning leads to

Theorem 6.5. *Let $\{\Delta_n\}$ be the sequence of sets in Theorem 6.2. Assume the set (6.14) is strongly flat at 0. Let $\{z_n\}$ be the sequence constructed in Theorem 5.2. Then*

$$z_n \to z \ne 0 \qquad \text{strongly in } E. \tag{6.16}$$

The results in this section generalize those in [F2]–[F4], where convergence of multipliers is obtained under much more restrictive conditions. We note that Theorem 6.1 can be considered a generalization of a well-known result on convergence of Kuhn–Tucker multipliers for perturbed nonlinear programming problems [L, p. 317].

We consider a control system described by the quasilinear initial value problem (4.1)–(4.2). We assume that there exists an *a priori* bound on the solutions,

$$\|y(t, u)\| \le C \qquad (0 \le t \le \bar{t})$$

which guarantees global existence. See [F2] and [F3] for indications on how to establish this bound. The compactness assumption is proved in [F3] for

$$f(t, y, u) = f(t, y) + Bu, \tag{6.17}$$

thus we only sketch the extension to the general case.

Lemma 6.6. *Let $\bar{t} > 0$, $I(\bar{t}) = \{(s, t); 0 \le s \le t \le \bar{t}\}$. Assume that the semigroup $S(t)$ is compact for all $t > 0$. Then the operator*

$$(\Pi v)(s, t) = \int_0^t S(t - \sigma)v(\sigma)\, d\sigma \tag{6.18}$$

from $L^2(0, \bar{t}; F)$ *(with its weak topology) into* $C(I(\bar{t}); E)$ *(the space of all E-valued continuous functions defined in* $I(\bar{t})$ *with it usual supremum norm) is compact.*

Proof. The result is a minor generalization of Lemma 6.1 in [F2]. ■

The compactness assumption is a consequence of the following result.

Lemma 6.7. *Assume that* $S(t)$ *is compact for* $t > 0$. *Let* $\{t_n\}$ *be a sequence with* $t_n \to \bar{t}$, $\{u^n\}$ *a sequence of controls,* $u^n \in W(0, t_n; U)$. *Then there exist an E-valued continuous function* $y(t)$ *in* $0 \leq t \leq \bar{t}$ *and a subsequence of* $\{u^n\}$ *(that we denote by the same symbol) with*

$$y(t, u^n) \to y(t) \tag{6.19}$$

uniformly in $0 \leq t \leq \bar{t}$.

Proof. We note that

$$y(t, u^n) = S(t)y^0 + \int_0^t S(t-\sigma)f(\sigma, y(\sigma, u^n), u^n(\sigma))\, d\sigma. \tag{6.20}$$

Select a subsequence of $\{f(\sigma, y(\sigma, u^n), u^n(\sigma))\}$ converging weakly in $L^2(0, \bar{t}; F)$ to $v(\cdot)$ and apply Lemma 6.6. The limit is $S(t)y^0 + (\Pi v)(0, t)$, which is not necessarily a trajectory of the system. ■

The compactness assumption can be proved without compactness of the semigroup, but only for certain cost functionals and control sets. See, for instance, the result in [F3, Theorem 3.2] for the time optimal problem, where Y is a closed ball and $f(t, y, u)$ is linear in u, but where $S(t)$ is an arbitrary semigroup.

We sketch below how the convergence principles are used in [F2]–[F4] and [F6] to deduce strong convergence of sequences of suboptimal controls in the time optimal case. Expressions like "$u^n \to \bar{u}$" (where, for instance, $\bar{u}$ belongs to $L^2(0, \bar{t}; F)$ and $u^n \in L^2(0, t_n; F)$) are understood by thinking of the u^n as elements of $L^2(0, \bar{t}; F)$ extending the function ($u = 0$ in $t > t_n$ if $\bar{t} > t_n$) or chopping it off at $\bar{t}$ if $\bar{t} \leq t_n$.

We deduce from the sequence maximum principle (5.23) and from (4.7) that

$$\langle S(t_n, s; \tilde{u}^n)^* z_n, f(s, y(s, \tilde{u}^n), v) - f(s, y(s, \tilde{u}^n), \tilde{u}^n(s))\rangle \geq -\delta_n, \tag{6.21}$$

in this case,

$$\langle S(t_n, s; \tilde{u}^n)^* z_n, v - u^n(s)\rangle \geq -\delta_n \tag{6.22}$$

with $\delta_n \to 0$, where $S(t, s; u)$ is the solution operator of (4.8). With a compact semigroup and a nonlinearity of the form (6.17) it is shown in [F2] that $S(t_n, s; \tilde{u}^n)^*$ is compact and that $S(t_n, s; \tilde{u}^n)^*$ is convergent to $S(\bar{t}, s; \bar{u})^*$ in the uniform topology of operators, where $\bar{u}$ is the weak limit of a subsequence of $\{\tilde{u}^n\}$.

Let U be an arbitrary set in F. Given a vector $y \neq 0$, define

$$U(y, \delta) = \{u \in U; \langle y, v - u\rangle \geq -\delta \quad (v \in U)\}$$

(for the geometry of the situation see [F2]). Assume that U is such that

$$\operatorname{diam} U(y, \delta) \to 0 \qquad (\delta \to 0) \tag{6.23}$$

for all $y \in F$, $y \neq 0$, and the solution operator $S(t, s; u)$ satisfies

$$S(t, s; u)^* z \neq 0 \qquad \text{if} \quad u \in W(0, \bar{t}; U), \qquad s \leq t, \quad z \neq 0, \tag{6.24}$$

then, if the weak convergence principle holds, $S(t, s; \tilde{u}^n)^* z_n$ in (6.21) converges strongly to $S(\bar{t}, s; \bar{u})^* z$ ($z_n =$ weak $\lim z_n$) and it follows from (6.23) that a subsequence of $\{\tilde{u}^n\}$ is convergent in $L^p(0, \bar{t}; F)$, $1 \leq p < \infty$; it suffices to apply (6.23) with $y = S(\bar{t}, s; \bar{u})z$.

The following result differs from those in [F2]–[F4], and [F6] in that it deals with a general nonlinearity $f(t, y, u)$ in a situation where optimal controls may not exist.

Theorem 6.8. *Assume the Hilbert space E is separable. Let $\{t_n\}$ be a sequence with $t_n \to \bar{t}$, and let $\{u^n\}$ be a sequence with $u^n \in W(0, t_n; U)$. Then there exist a subsequence of $\{u^n\}$ (which we denote by the same symbol) and a linear bounded operator*

$$B^*: E \to L^\infty(0, \bar{t}; E) \tag{6.25}$$

such that, for each $z \in E$,

$$S^*(t, s; u^n)z \to S^*(t, s)z \tag{6.26}$$

uniformly in $I(\bar{t}) = \{(s, t); 0 \leq s \leq t \leq \bar{t}\}$, where $S^*(t, s)$ is the only continuous solution of the equation

$$S^*(t, s)z = S^*(t - s)z + \int_s^t S(t - \sigma)^*(B^* S^*(\sigma, s)z)(\sigma)\, d\sigma. \tag{6.27}$$

We note that $(B^*z)(t)$ stands for the image of z in $L^\infty(0, \bar{t}; U)$ by the operator B^* in (9.7). Note also that, for fixed t, $z \to (B^*z)(t)$ may not be a bounded operator; in fact the operator $z \to (B^*z)(t)$ may not even be defined for any t, since the function $t \to (B^*z)(t)$ may be defined in a different set (of full measure) of $0 \leq t \leq \bar{t}$ for each z. The symbol $(B^*g(\sigma))(\sigma)$ (as on the right-hand side of (6.27)) denotes the value at σ of the function $\sigma \to (B^*z(\sigma))(\sigma)$. That this makes sense even for a single σ is not obvious (the functions $B^*z(\sigma)$ are defined modulo null sets that depend on $z(\sigma)$), thus this point needs clarification. We prove below that, in fact, for each $g \in L^\infty(0, \bar{t}; E)$ the function $(B^*g(\sigma))(\sigma)$ is well defined a.e., essentially bounded, and the operator

$$g(\sigma) \to (B^*g(\sigma))(\sigma) \tag{6.28}$$

is bounded from $L^\infty(0, \bar{t}; E)$ into itself. Let $g(\sigma) = \sum \chi_k(\sigma) y_k$ be a step function (a finite linear combination of characteristic functions of disjoint measurable sets). Then $(B^*g(\sigma))(\sigma) = \sum \chi_k(\sigma)(B^*y_k)(\sigma)$ is a well-defined element of $L^\infty(0, \bar{t}; E)$ for almost all $\sigma \in [0, \bar{t}]$ and

$$\begin{aligned} \|(B^*g(\sigma))(\sigma)\| &\leq \|\textstyle\sum \chi_k(\sigma)(B^*y_k)(\sigma)\| \leq \sum \chi_k(\sigma)\|B^*y_k\|_\infty \\ &\leq \textstyle\sum \chi_k(\sigma)\|B^*\|\|y_k\| = \|B^*\|(\sum \chi_k(\sigma)\|y_k\|) = \|B^*\|\|g(\sigma)\|, \end{aligned}$$

where $\|B^*\|$ is the norm of the operator (6.25). Let now $g(\sigma)$ be an arbitrary integrable function. Select a sequence $\{g_m\}$ of step functions such that $g_m(\sigma) \to g(\sigma)$ a.e. Then $\|(B^*g_m(\sigma))(\sigma) - (B^*g_n(\sigma))(\sigma)\| \leq \|B^*\|\|g_m(\sigma) - g_n(\sigma)\|$, so that $\{B^*g_m(\sigma)(\sigma)\}$ is convergent a.e. to a measurable function, declared to be $(B^*g(\sigma))(\sigma)$. Obviously this definition does not depend on the approximating sequence and we have

$$\|(B^*g(\sigma))(\sigma)\| \leq \|B^*\|\|g(\sigma)\|. \tag{6.29}$$

This gives meaning to the integral equation (6.27) and allows us to solve it in the usual way by successive approximations.

Proof of Theorem 6.8. Let $\{\varphi_m; m \geq 1\}$ be a complete orthonormal system in E. Using the diagonal process, select a subsequence of $\{u^n\}$ (denoted the same) such that for each m the sequence

$$B(\sigma, u^n)^*\varphi_m = \partial_y f(\sigma, y(\sigma, u^n), u^n(\sigma))^*\varphi_m \tag{6.30}$$

converges weakly in $L^2(0, \bar{t}; E)$ to a limit $b_m(\sigma) \in L^2(0, \bar{t}; F)$ as $n \to \infty$. In view of the boundedness of $\{u^n\}$ and $\{y(t, u^n)\}$ and of the hypotheses on $\partial_y f$ we know that each $b_m(\sigma) \in L^\infty(0, \bar{t}; E)$. More generally, since for any finite sum $\sum \alpha_m\varphi_m$ we have

$$B(\sigma, u^n)^*(\textstyle\sum \alpha_m\varphi_m) = \partial_y f(\sigma, y(\sigma, u^n), u^n(\sigma))^*(\sum \alpha_m\varphi_m) \to \sum \alpha_m b_m(t)$$

weakly in $L^2(0, \bar{t}; F)$ as $n \to \infty$, it follows again from uniform boundedness of $B(t, u^n)$ that there exists a constant C such that $\|\sum \alpha_m b_m(t)\| \leq C\|\sum \alpha_m\varphi_m\|$ a.e. in $0 \leq t \leq \bar{t}$. Using this for tails of the series $\sum \alpha_m\varphi_m$ we conclude that the partial sums of $\sum \alpha_m b_m$ are a Cauchy sequence in $L^\infty(0, \bar{t}; E)$, thus we can define an operator B^* of the required form (6.25) for arbitrary $y \in E$ by developing y in a Fourier series, $y = \sum \alpha_m\varphi_m$, and setting

$$(B^*y)(t) = B^*\left(\sum_{m=1}^{\infty} \alpha_m\varphi_m\right)(t) = \sum_{m=1}^{\infty} \alpha_m b_m(t).$$

We claim that for each $y \in E$ we have

$$B(t, u^n)^*y \to (B^*y)(t) \qquad \text{weakly in} \quad L^2(0, \bar{t}; E). \tag{6.31}$$

This is obvious if $y = y_m$ is a finite linear combination of the φ_m. Now let y be an arbitrary element of E and $\varepsilon > 0$. Developing y in Fourier series we can write $y = y_m + \tilde{y}_m$, where y_m is a finite partial sum of the series and $\|\tilde{y}_m\| \leq \varepsilon$. Let $v(\sigma) \in L^2(0, \bar{t}; E)$. We have

$$\begin{aligned}\int_0^{\bar{t}} \langle B(\sigma, u^n)^*y - (B^*y)(\sigma), v(\sigma)\rangle\, d\sigma &= \int_0^{\bar{t}} \langle B(\sigma, u^n)^*y_m - (B^*y_m)(\sigma), v(\sigma)\rangle\, d\sigma \\ &\quad + \int_0^{\bar{t}} \langle B(\sigma, u^n)^*\tilde{y}_m - (B^*\tilde{y}_m)(\sigma), v(\sigma)\rangle\, d\sigma.\end{aligned} \tag{6.32}$$

The second term in (6.32) can be bounded by $C\|v\|\varepsilon$; the first can be estimated by an expression of the same form taking $m \geq m_0(n)$.

The proof ends as follows. Consider the integral equation defining the solution operator $S^*(t, s; u^n)$:

$$S^*(t, s; u^n)z = S^*(t - s)z + \int_0^t S(t - \sigma)^*B(\sigma, u^n)^*S^*(\sigma, s; u^n)z\, d\sigma. \tag{6.33}$$

This equation is solved by successive approximation starting with $(S^*)_0(\sigma, s; u^n)z = z$. We define the sequence $\{(S^*)_m(t, s; u^n); m \geq 1\}$ by

$$(S^*)_{m+1}(t, s; u^n)z = S^*(t - s)z + \int_0^t S(t - \sigma)^* B(\sigma, u^n)^* (S^*)_m(\sigma, s; u^n)z \, d\sigma. \quad (6.34)$$

Selecting a subsequence of the sequence $\{u^n\}$, we may assume that $B(\sigma, u^n)^*(S^*)_0(\sigma, s; u^n)z$ is weakly convergent in $L^2(0, \bar{t}; E)$; hence by Lemma 6.6, $(S^*)_1(t, s; u^n)z$ is uniformly convergent as $n \to \infty$ in the interval $0 \leq t \leq \bar{t}$. Refining the subsequence and applying Lemma 6.6 at each step, and then selecting a diagonal subsequence we conclude that all the sequences $\{(S^*)_{m+1}(t, s; u^n)z, n = 1, 2, \ldots\}$ are uniformly convergence in $0 \leq t \leq \bar{t}$. We then take advantage of the uniform convergence of the approximations $\{(S^*)_m(t, s; u^n)\}$. Since the details are similar to those in the proof of Theorem 6.2 in [F2], we omit them. ■

We are now in a position to establish a convergence result. Let $\{u^n\}$ be the (t_n, ε_n)-suboptimal sequence $\{u^n\}$ with $\varepsilon_n \to 0$. We start with (6.21). Theorem 6.8 shows that we have $S(t_n, s; \tilde{u}^n)^* \to S^*(t, s)$ strongly, uniformly in $0 \leq s \leq t_n$. Assuming now the target set Y satisfies the necessary assumptions for the strong convergence principle (Theorem 6.5) we may assure, if necessary passing to a subsequence, that $\{z_n\}$ is strongly convergent to $z \neq 0$. On the other hand, by Lemma 6.7, if necessary passing again to a subsequence, we may assume that $y(t, \tilde{u}^n)$ is uniformly convergent to a continuous function $y(s)$. Thus, we may rewrite (6.21) as follows:

$$\langle S^*(\bar{t}, s)z, \langle f(s, y(s), v) - f(s, y(s), \tilde{u}^n(s)) \rangle \rangle \geq -\tilde{\delta}_n \to 0. \quad (6.35)$$

Assuming the set $U_s = f(s, y(s), U)$ satisfies (6.23) and that (6.24) holds for $S(t, s)$, $L^p(0, \bar{t}; U)$-convergence of $f(s, y(s), \tilde{u}^n(s))$ is obtained.

References

[AE] J.-P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley-Interscience, New York, 1984.

[C1] F. Clarke, The maximum principle under minimum hypotheses, *SIAM J. Control Optim.*, **14** (1976), 1078–1091.

[C2] F. Clarke, *Optimization and Nonsmooth Analysis*, Wiley—Interscience, New York, 1983.

[CL] F. Clarke and P. Löwen, State constraints in optimal control: a proximal normal analysis approach, *SIAM J. Control Optim.* (to appear).

[E1] I. Ekeland, Sur les problémes variationnels, *C. R. Acad. Sci. Paris*, **275** (1972), 1057–1059.

[E2] I. Ekeland, Nonconvex minimization problems, *Bull. Amer. Math. Soc. (N.S.)*, **1** (1979), 443–474.

[F1] H. O. Fattorini, A unified theory of necessary conditions for nonlinear nonconvex control systems, *Appl Math. Optim.*, **15** (1987), 141–185.

[F2] H. O. Fattorini, *Optimal Control of Nonlinear Systems: Convergence of Suboptimal Controls, I*, pp. 159–199, Lecture Notes in Pure and Applied Mathematics, Vol. 108, Marcel Dekker, New York, 1987.

[F3] H. O. Fattorini, *Optimal Control of Nonlinear Systems: Convergence of Suboptimal Controls, II*, pp. 230–266, Lecture Notes on Control and Information Sciences, Vol. 97, Springer-Verlag, Berlin, 1987.

[F4] H. O. Fattorini, *Convergence of Suboptimal Controls for Point Targets*, pp. 91–107. International Series of Numerical Mathematics, Vol. 78, Birkhäuser, Basel, 1987.

[F5] H. O. Fattorini, Some remarks on convergence of suboptimal controls, in *Software Optimization*, pp. 359–363, A. V. Balakrishnan Anniversary Volume, New York, 1988.

[F6] H. O. Fattorini, Convergence of suboptimal controls: the point target case, *SIAM J. Control Optim.* (to appear).

[F7] H. O. Fattorini, Constancy of the Hamiltonian in infinite-dimensional systems, in *Proceedings of the 4th International Conference on Control of Distributed Parameter Systems*, Vorau, 1988, pp. 123–133, International Series of Numerical Mathematics, Vol. 91, Birkhäuser, Basel, 1989.

[FF1] H. O. Fattorini and H. Frankowska, Necessary conditions for infinite-dimensional control problems, *Proceedings of the 8th International Conference on Analysis and Optimization of Systems*, Antibes-Juan Les Pins, June 1988, Lecture Notes on Control and Information Sciences, Springer-Verlag, Berlin (to appear).

[FF2] H. O. Fattorini and H. Frankowska, Explict convergence estimates for suboptimal controls, I, II, *Problems Control Inform. Theory* (to appear).

[F8] H. Frankowska, The maximum principle for differential inclusions with end point constraints, *SIAM J. Control Optim.*, **25** (1987), 145–157.

[F9] H. Frankowska, *On the Linearization of Nonlinear Control Systems and Exact Reachability*, pp. 132–143, Lecture Notes on Control and Information Sciences, Vol. 114, Springer-Verlag, Berlin, 1988.

[L] D. G. Luenberger, *Linear and Nonlinear Programming* (2nd edn.), Addison-Wesley, Reading, MA, 1984.