MEMORANDUM

No 22/2000

Nonsmooth maximum principle for control problems in Banach state space

By
Atle Seierstad

ISSN: 0801-1117

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.oekonomi.uio.no/
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no/
e-mail: frisch@frisch.uio.no

List of the last 10 Memoranda:

No 21	By Diderik Lund: Imperfect loss offset and the after-tax expected rate of return to equity, with an application to rent taxation. 20 p.
No 20	By Christian Brinch: Identification of Structural Duration Dependence and Unobserved Heterogeneity with Time-varying Covariates. 19 p.
No 19	By Knut Røed and Morten Nordberg: Have the Relative Employment Prospects for the Low-Skilled Deteriorated After All? 21 p.
No 18	By Jon Vislie: Environmental Regulation under Asymmetric Information with Type-dependent outside Option. 20 p.
No 17	By Tore Nilssen and Lars Sørgard: Strategic Informative Advertising in a TV-Advertising Duopoly. 21 p.
No 16	By Michael Hoel and Perry Shapiro: Transboundary Environmental Problems with a Mobile Population: Is There a Need for Central Policy? 19 p.
No 15	By Knut Røed and Tao Zhang: Labour Market Transitions and Economic Incentives. 21 p.
No 14	By Dagfinn Rime: Private or Public Information in Foreign Exchange Markets? An Empirical Analysis. 50 p.
No 13	By Erik Hernæs and Steinar Strøm: Family Labour Supply when the Husband is Eligible for Early Retirement. 42 p.
No 12	By Erik Biørn: The rate of capital retirement: How is it related to the form of the survival function and the investment growth path? 36 p.

A complete list of this memo-series is available in a PDF® format at:
http://www.oekonomi.uio.no/memo/index.shtml

Nonsmooth maximum principle for control problems in Banach state space.

by
Atle Seierstad
University of Oslo, Norway

Abstract Necessary conditions in the form of a maximum principle is proved for the optimal control of solutions to nonsmooth nonlinear differential equations in Banach space. The conditions constitute a generalization to infinite dimensions of the maximum principle of Clarke. The approach is closely related to that of Yong (1990) and utilizes an approximating smoother system that exhibits Gâteaux differentiability. The results are applicable to Volterra integral equations and mild solutions of certain types of weakly nonlinear evolution equations.

1. Introduction The purpose of this paper is to give necessary conditions in the form of a maximum principle for the optimal control of solutions to nonsmooth nonlinear differential equations in Banach space. The conditions constitute a generalization to infinite dimensions of the maximum principle of Clarke, involving a sort of pseudoHamiltonian. The approach is closely related to that of Yong (1990), but here a general Banach space is considered, and the differential equation has no special structure. On the other hand, Yong considers a semilinear (or weakly nonlinear) evolution equation. In a remark below it is noted how the results in this paper can be translated to be applicable to Volterra integral equations, and then to mild solutions of certain types of semilinear evolution equations. As in the paper of Yong, an approximating smoother system is utilized, which in the present paper exhibits Gâteaux differentiability. The smoothing differs slightly from that of Yong. A selection of references to control problems in Banach space is included, mostly they involve applications to partial differential equations. Works discussing nonsmooth problems include those of Barbu, Fattorini and Frankowska (1991) and Fattorini (1993), (1999), (at least in the abstract parts of these works).

2. Terminology and notation

Two real Banach spaces X and Y are given. For any set A, cl$A$ means the norm-closure (or metric) closure of A. The topological dual of X is X^*, and if A is a bounded linear map from X into Y, then A^* is the topological dual map from Y^* into X^*. For any locally Lipschitz continuous real-valued function $\Psi(x)$ on X, $d^0\Psi(x)(w)$ is the Clarke generalized directional derivative, at x, in direction w, see Clarke (1983). The Clarke generalized gradient of a real-valued locally Lipschitz continuous function $f(x)$ on X is written $\partial f(x)$. If needed, we write ∂_x when it is taken with respect to x. Measurability and integrability of vector valued functions are used in the sense of Dunford and Schwartz (1967), often called strong or Bochner measurability (integrability). Let $J := [0, T], T$ a fixed number > 0. The set of (Lebesgue) measurable functions for which $\int_J |x(t)|^p < \infty$, $p \in [1, \infty)$, is denoted $L_p(J, X)$. If $p = \infty$, $L_\infty(J, X)$ consists of measurable essentially bounded functions. A function $x(.) : J \to X$, is defined to be antidifferentiable if it is absolutely continuous and has a derivative a.e. which is integrable. (Then $x(t) = x(0) + \int_0^t \dot{x}(s)ds$.) For any set $M, 1_M$ is the corresponding indicator function.

3. The control system

The control system is defined by the differential equation

$$dx(t)/dt = g(t, x(t), u(t)), t \in J, x(0) = x_0 \in X. \tag{1}$$

Here, x_0 is a fixed point, $g : J \times X \times U \to X$ is a fixed function, and U is a given metric space. The controls $u(t) : J \to U$ are (Lebesgue) measurable, i.e. the $u(.)$s' are a.e. limits of step functions. The set of all such control functions is denoted $\tilde{\mathcal{U}}$. For each x and each $u(.) \in \tilde{\mathcal{U}}, t \to g(t, x, u(t))$ is assumed to be (Lebesgue) measurable. The solutions $x(t)$ of (1) are antidifferentiable functions taking values in X. The constraints in the problem are:

$$\text{(i) } G(x(T)) \in C, C \text{ a closed convex set}, \qquad \text{(ii) } u(.) \in \mathcal{U}. \tag{2}$$

Here, G is a given locally Lipschitz continuous function from X into Y, C is a fixed subset in Y, and $\mathcal{U}$ is given family of measurable functions $u(.) : J \to U$, ($\mathcal{U}$ is a subset of $\tilde{\mathcal{U}}$).
The criterion to be maximized is

$\phi(x(T))$, where $\phi : X \to \mathbb{R}$ is a given locally Lipschitz continuous function. (3)

The maximization problem is thus

$$\max_{x(.),u(.)} \phi(x(T)) \text{ subject to } (1), G(x(T)) \in C, \text{ and } u(.) \in \mathcal{U}. \tag{4}$$

A "system pair" is a pair $(x(.), u(.))$ such that (1) and (2)(ii) are satisfied, with $x(.)$ antidifferentiable. If the system pair also satisfies $G(x(T)) \in C$,

it is called "admissible". If $x(.)$ is unique for a given $u(.)$, we often write $x(.) = x^u(.)$, (below, conditions will be imposed securing uniqueness for $u(.)$'s close to the optimal $u^*(.)$). The control problem (4) amounts to maximizing $\phi(x(T))$ in the class of admissible pairs $(x(.), u(.))$.
A free end problem is a problem where the condition $G(x(T)) \in C$ is omitted, or, equivalently, where $Y = \{0\} \subset \mathbb{R}$, (and $C = \{0\}, G(.) \equiv 0$). The free end case is referred to as $Y = \{0\}$. Below, $(x^*(.), u^*(.))$ denotes an optimal admissible pair, assumed to exist in the problem. Two assumptions are made:

For some number $\varsigma > 0$,

for all $u(.) \in \mathcal{U}$, there exist measurable functions $M_{u(.)}(t) \in L_1(J, \mathbb{R})$, $M^{u(.)}(t) \in L_2(J, \mathbb{R})$, such that, for all $x \in B(x^*(t), \varsigma)$ and for all t, $|g(t, x, u(t))| \le M_{u(.)}(t)$ and such that, for all t, $x \to g(t, x, u(t))$ is Lipschitz continuous in $x \in B(x^*(t), \varsigma)$ of rank $\le M^{u(.)}(t)$. Finally, ϕ and G are Lipschitz continuous in $B(x^*(T), \varsigma)$ of ranks M_ϕ and M_G, respectively. (5)

$\mathcal{U}$ is closed under switching, i.e. $u_1(.), u_2(.) \in \mathcal{U}$, M measurable $\Rightarrow$ $u_3(.) \in \mathcal{U}$, where $u_3(t) = u_1(t)$ for $t \in M, u_3(t) = u_2(t)$ for $t \notin M$. Moreover, $\mathcal{U}$ is essentially closed in the pseudometric $\sigma(u_1(.), u_2(.)) := \text{meas}\{t : u_1(t) \neq u_2(t)\}$ (6)

The assumptions made about $X, Y, \phi, G, g, C, U, x^*(.), u^*(.)$ and $\mathcal{U}$ in this section are used throughout this paper.

4. Necessary conditions

The following theorem holds in the free end, Gâteaux differential case, (for the definition of the latter term, see Fattorini (1999), p. 310).

Theorem 1 (Free end, Gâteaux derivative in x.) Let $Y = \{0\}$. Assume that for each $t \in J$, $\hat{x} \to g(t, x^*(t) + \hat{x}, u^*(t))$ has a Gâteaux derivative $g_x(t, x^*(t), u^*(t))$ at $\hat{x} = 0$. Assume, furthermore, that ϕ has a Gâteaux derivative at $x^*(T)$. Then the following maximum principle holds: For all $u(.) \in \mathcal{U}$, for t not in a null set $N_{u(.)}$,

$$\langle g(t, x^*(t), u(t)), p(t) \rangle \le \langle g(t, x^*(t), u^*(t)), p(t) \rangle. \tag{7}$$

Here $p(t) : J \to X^*$ is absolutely continuous and satisfies

$$p(T) = [\phi_x(x^*(T))] \tag{8}$$

and, for all $\hat{x} \in X$,

$$d\langle \hat{x}, p(t)\rangle/dt = \langle -g_x(t, x^*(t), u^*(t))\hat{x}, p(t)\rangle \tag{9}$$

for all t not in a null set $N_{\hat{x}}$. □

The next theorem to be stated concerns the nondifferentiable case. To formulate it, a few definitions are needed. Let $\to^*$ denote weak* convergence and $\to$ denote norm convergence.
For any locally Lipschitz continuous function $f : X \to Z$, (Z a Banach space), for any given $\beta > 0, (\hat{x}, \hat{z}^*) \in X \times Z^*, |\hat{z}^*| \leq \beta$, define

$$d^\beta(f, \hat{x}, \hat{z}^*)(w) := \limsup_{z^* \to^* \hat{z}^*, z^* \in clB(0,\beta), y \to \hat{x}, \lambda \searrow 0} \langle \lambda^{-1}[f(y + \lambda w) - f(y)], z^*\rangle. \tag{10}$$

Below, write $d_x^\beta(g(t, \hat{x}, u), \hat{x}^*) := d^\beta(f, \hat{x}, \hat{x}^*)$ for $f(x) = g(t, x, u)$, $\beta(t) := (M_\phi + M_G)\exp(\int_t^T M^*(s)ds)$ and $\gamma(t) := M^*(t)\beta(t)$, ($M_G = 0$ in the free end case.)

Theorem 2 (Free end, local Lipschitz continuity in x.) Let $Y = \{0\}$. Assume that $M^{u(.)}(.) \leq M^*(.)$ and $M_{u(.)}(.) \leq M_*(.)$ for all $u(.) \in \mathcal{U}$, where $M^*(.) \in L_2(J, \mathbb{R}), M_*(.) \in L_1(J, \mathbb{R})$. There exists an absolutely continuous function $p(.) : J \to X^*$, such that, for all $u(.) \in \mathcal{U}$, for t not in a null set $N_{u(.)}$, the inequality (7) holds. Moreover, $|p(t)| \leq \beta(t)$ for all t, and for all $w \in X$, for t not in a null set N_w,

$$-d\langle w, p(t)\rangle/dt \leq d_x^{\beta(t)}(g(t, x^*(t), u^*(t)), p(t))(w), \tag{11}$$

and, finally,

$$p(T) \in \partial\phi(x^*(T)). \tag{12}$$

□

In the end constrained case, the following four conditions are needed in order to formulate a constraints qualification (see (17) below), needed for our necessary conditions to hold.

$$\sigma(\hat{u}, u^*) \leq \varepsilon'. \tag{13}$$

For all $u(.) \in \mathcal{U}$, $\int_J \{\langle g(t, x(t), u(t)), p(t)\rangle - \langle g(t, x(t), \hat{u}(t)), p(t)\rangle\} dt \leq \varepsilon'$ (14)

For all $w \in X$, for $t \notin N_{w,\hat{u}(.)}$, ($N_{w,\hat{u}(.)}$ a null set),
$$-d\langle w, p(t)\rangle/dt \leq d_x^{\beta(t)}(g(t, x(t), \hat{u}(t)), p(t))(w). \tag{15}$$

For all $w \in X$, $\langle w, p(T) - \hat{p}\rangle \leq d^{M_G}(G, x(T), y^*)(w)$, $|\hat{p}| \leq \varepsilon'$,$|\ y^*| > 1/2$,
$$\langle C - G(x(T)), y^*\rangle \geq 0. \tag{16}$$

Theorem 3 (End constraint, local Lipschitz continuity in x.) Assume that $M^{u(.)}(.) \leq M^*(.)$ and $M_{u(.)}(.) \leq M_*(.)$ for all $u(.) \in \mathcal{U}$, where $M^*(.) \in L_2(J, \mathbb{R})$, $M_*(.) \in L_1(J, \mathbb{R})$. Assume also that there exists a vector $y \in Y$ and a number $\varepsilon' > 0$, such that for any quintuple $(x(.), \hat{u}(.), p(.), \hat{p}, y^*)$, where $\hat{p} \in X^*, y^* \in Y^*, (x(.), \hat{u}(.))$ is a system pair and $p(.) : J \to X^*$ absolutely continuous,

if $(x(.), \hat{u}(.), p(t), \hat{p}, y^*)$ satisfies the four conditions (13)-(16), then $\langle y,y^*\rangle \geq \varepsilon'$. (17)

Then there exist a number $\lambda_0 \geq 0$, and elements $\hat{p}$ and $\check{p}$ in X^* and y^* in Y^*, $(\lambda_0, y^*) \neq 0$, and an absolutely continuous function $p(.) : J \to X^*$, $|p(t)| \leq \beta(t)$, such that (7) and (11) hold, together with the following condition:

$$\langle C - G(x^*(T)), y^*\rangle \geq 0, p(T) = \hat{p} + \check{p}, \hat{p} \in \lambda_0 \partial\phi(x^*(T)),$$
$$\text{for all } w \in X, \langle w, \check{p}\rangle \leq d^{M_G}(G, x^*(T), y^*)(w) \tag{18}$$

Remark 1 In the three theorems above, $p(t)$ satisfies $|p(t)-p(s)| \leq \int_s^t \gamma(s)ds, t > s$. There exists a scalarwise integrable function $\check{p}(.) : J \to X^*$, with $|\check{p}(t)| \leq \gamma(t)$ a.e. and $|\check{p}(.)|$ measurable, such that, for any $w \in X$, for a.e. t, $d\langle w, p(t)\rangle/dt = \langle w, \check{p}(t)\rangle$. Given absolute continuity, this equality, together with $p(T) = p_T$, is equivalent to:

$$\langle w, p(t)\rangle = \langle w, p_T\rangle + \int_T^t \langle w, \check{p}(s)\rangle ds \text{ for all } t. \tag{19}$$

The function $\check{p}(t)$ is also written $\dot{p}(t)$, and is called a "scalarwise derivative" of $p(t)$. In the case X is separable or reflexive, $\dot{p}(.)$ belongs to $L_2(J, X^*)$.

Remark 2 If X is reflexive, then for some null set N, for $t \notin N$,

$$-\dot{p}(t) \in \tilde{\partial}_x^{\beta(t)}(g(t,x^*(t),u^*(t)),p(t)) := \{x^* \in X^* : \text{for all } w \in X, \langle w, x^* \rangle \leq \liminf_{\delta \searrow 0} \int_{t-\delta}^{t} d_x^{\beta(t)}(g(s,x^*(s),u^*(s)),p(s))(w)ds\}. \quad (20)$$

where the integral is a Lebesgue lower integral in case the integrand is non-measurable, i.e. the supremum of the integrals of all measurable functions smaller than or equal to the integrand for all t. The integrand is surely measurable in t, if either X is separable, or if g is simultaneously continuous in (t,u). If X is separable, then, for some null set $\tilde{N}$, for $t \notin \tilde{N}$,

$$-\dot{p}(t) \in \partial_x^{\beta(t)}(g(t,x^*(t),u^*(t)),p(t)) := \{x^* \in X^* : \text{for all } w \in X, \langle w, x^* \rangle \leq d_x^{\beta(t)}(g(t,x^*(t),u^*(t)),p(t))(w). \quad (20')$$

Remark 3 (A weaker constraints qualification.) Theorem 3 holds even if (17) is replaced by the weaker condition: There exist a finite dimensional subspace Y' of Y, and, if $Y \backslash Y' \neq \emptyset$, a vector $y \in Y$, and a number $\varepsilon' > 0$ such that for any quintuple $(x(.), \hat{u}(.), p(.), \hat{p}, y^*)$,($p(.)$ absolutely continuous, $x(.), \hat{u}(.)$ a system pair), if $(x(.), \hat{u}(.), p(t), \hat{p}, y^*)$ satisfies the four conditions (13)-(16) and $|\langle \hat{y}', y^* \rangle| \leq \varepsilon' |\hat{y}'|$ for all $\hat{y}' \in Y'$, then $\langle y, y^* \rangle \geq \varepsilon'$. (No constraints qualification is needed when $Y \backslash Y' = \emptyset$.) This weakened constraints qualification is automatically satisfied if $Y = Y' + Y''$, Y'' is a closed subspace, $Y' \cap Y'' = \{0\}$, and if for some $z \in Y$, some $\varepsilon > 0, \Pi'' B(z, \varepsilon) \subset \Pi''[(C - G(x^*(T))) \cap B(0,1)]$, where Π'' is the projection onto Y''.

Remark 4 (A different constraints qualification) A locally Lipschitz continuous function $h : X \to Y$ is said to have a "directional multiderivative" at x, if the set $\triangle h(x)(w,r) := \{\lambda^{-1}\{h(x + \lambda w) - h(x)\} : \lambda \in (0,r]\}$ is norm-compact for some $r > 0$, and its directional multiderivative is then defined as $Dh(x)(w) := \cap_{r>0} \triangle h(x)(w,r)$. Assume in the situation of Theorem 3, that $C = \{0\}$, that (17) does not hold, but that, for all $u(.) \in \mathcal{U}$ and all t, $\hat{x} \to g(t, \hat{x}, u(t))$ has a directional multiderivative at all $x \in B(x^*(t), \varsigma)$. Then, provided $M^{u(.)}(.) = M^{u^*(.)}(.) =$ constant and $M_{u(.)}(.) = M_{u^*(.)}(.) =$ constant, and condition (3), p. 306 in Seierstad (1997) holds, then the conclusion of Theorem 3 still holds, even with $\lambda_0 = 1$. For brevity, the condition (3) just mentioned is not stated here, it contains a certain type of approximate attainability condition on the directional multiderivatives of the end points $x^{u(.)}(T)$, $u(.)$ close to $u^*(.)$, obtained by perturbing $u(.)$ at small intervals. From this result, a result for general closed convex sets C can be obtained.

The proof consists in applying Theorem 2 to a free end problem where the restriction $G(x) = 0$ is replaced by a penalization term, using the exact penalization result, Theorem 1, in Seierstad (1997). □

Remark 5 (A case where $d_x^{\beta(t)}(g(t, x^*(t), u^*(t)), p(t))$ can be replaced by $d_x^0\langle g(t, x, u^*(t)), p(t)\rangle$.) Assume that, for any $(t,w) \in J \times X$, for any sequence of pairs (x_k,λ_k) converging to $(x^*(t), 0)$, $\lambda_k > 0$, the sequence $\lambda_k^{-1}\{g(t, x_k + \lambda_k w, u^*(t)) - g(t, x_k, u^*(t))\}$ contains a norm-convergent subsequence. Then, in (11) and (20),
$d_x^{\beta(t)}(g(t, x^*(t), u^*(t)), p(t))(w)$ can be replaced by $d_x^0\langle g(t, x^*(t), u^*(t)), p(t)\rangle(w)$, the Clarke generalized directional derivative of $x \to \langle g(t, x, u^*(t)), p(t)\rangle$ at $x = x^*(t)$. Moreover, in this case, $-\dot{p}(t) \in \partial_x\langle g(t, x^*(t), u^*(t)), p(t)\rangle$ a.e., if X is separable.

Remark 6 (Relation of (17) to another constraints qualification.) In the case $M_{u(.)}(.) = M_{u^*(.)}(.) =$ constant, $M^{u(.)}(.) = M^{u^*(.)}(.) =$ constant, the constraints qualification (17) is implied by a standard one used in the continuous Frechet derivative case (the case where Frechet derivatives of $x \to g(t, x, u)$ (for all t, u), of $\phi(x)$, and of $G(x)$ exist and are continuous in $B(x^*(t), \varsigma)$, respectively, $B(x^*(T), \varsigma)$). In the continuous derivative case, $(\lambda_0, y^*) \neq 0$ if, for some $z \in Y$, $\varepsilon > 0$, $B(z, \varepsilon) \subset cl\{(dG(x^*(T))/dx)q_u(T) - c + G(x^*(T)) : u = u(.) \in \mathcal{U}, c \in C\}$, where $q_u(.)$ is the solution of $dq_u/dt = g(t, x^*(t), u(t)) - g(t, x^*(t), u^*(t)) + (\partial g(t, x^*(t), u^*(t))/\partial x)q$ a.e., $q(0) = 0$. (A weakening similar to that of Remark 3 is possible.)

Remark 7 (Weakened boundedness and Lipschitz condition.) In (5) it can be assumed that $M^{u(.)} \in L_1(J, \mathbb{R})$, and in Theorems 2 and 3, it can be assumed that $M^*(.) \in L_1(J, X)$. Also, in Theorems 2 and 3, the definition of $\beta(t)$ can be changed to $\beta(t) := (M_\phi + M_G)\exp(\int_t^T M^{u^*(.)}(s)ds)$, ($M_G = 0$ in the free end case). A further generalization is that, in Theorems 2 and 3, (and Remarks 1-3), the conditions $M^{u(.)}(.) \leq M^*(.)$ and $M_{u(.)}(.) \leq M_*(.)$ for all $u(.) \in \mathcal{U}$ can be dropped, however it is then needed to replace $\mathcal{U}$ in (14) by some subset $\hat{\mathcal{U}}$ for which the two inequalities hold for some $M^*(.), M_*(.)$.

Remark 8 (Applications to Volterra integral equations.) The above results can be applied to Volterra integral equations, and hence to mild solutions of certain abstract weakly nonlinear evolution equations: Consider the Volterra integral equation

$$y(t) := x_0(t) + \int_0^t g(t,s,y(s),u(s))ds,\ t \in J, \tag{21}$$

where $x_0(.) \in C(J,X)$, $g : J \times J \times X \times U \to X$. Let $\pi_s : C(J,X) \to X$ be defined by $\pi_s \hat{x}(.) = \hat{x}(s)$, $\hat{x}(.) \in C(J,X)$. Next, consider for a moment the integral equation

$$z(\tau,t) := x_0(t) + \int_0^\tau g(t,s,z(s,s),u(s))ds,\ \tau,t \in J \tag{22}$$

This equation we may rewrite as

$$z(\tau,.) := x_0(.) + \int_0^\tau g(.,s,\pi_s z(s,.),u(s))ds,\ \tau \in J, \tag{23}$$

where $z(\tau,.) : J \to C(J,X)$. Taking the last space as our state space and writing $z(\tau,.) = z(\tau)$, this integral equation can equivalently be expressed as an ordinary differential equation

$$dz/ds = \check{g}(s,z(s),u(s)), z(0) = x_0, \check{g}(s,z(s),u(s)) := g(.,s,\pi_s z(s),u(s)), \tag{24}$$

$(x_0 = x_0(.))$. For this to work, we have to assume that g is separately continuous in t, that $g(.,s,x,u(s)) : J \times X \to C(J,X)$ is measurable in s, for each x, $u(.) \in \tilde{U}$, and that $g(.,s,x,u(.))$ is bounded by an integrable $M_{u(.)}(s)$ in $B(x^*(t),\varsigma)$ and is Lipschitz continuous in x here, with integrable Lipschitz rank $M^{u(.)}(s)$. (Again $(x^*(t),u^*(t))$ is a given optimal pair.) The criterion to be maximized is now $\phi(\pi_T z(T)) = \phi(y(T))$, and the end constraint takes the form $G(\pi_T z(T)) = G(y(T)) \in C$. For each solution $y(t) \in B(x^*(t),\varsigma)$ of (21), there is a solution $z(.)$ with $\pi_t z(t) \in B(x^*(t),\varsigma)$ of (24) with $\pi_t z(t) = y(t)$, and vice versa. □

5. Proofs

The proof of Theorem 1 is closely parallel to that of Theorem 1 in Pallu de la Barriere, p. 383, (1980). A proof is given in Appendix.

Proof of Theorem 2 (free end case) The proof is based on the use of Theorem 1, and is structured as follows. First, by using suitable mollifiers, the dependence on x in g is smoothened to such an extent that Gâteaux derivatives exist. The control $u^*(.)$ is approximately optimal in the smoothened problem. Ekeland's principle is used to obtain an optimal control $u_*(.)$ in the smoothened problem. To this problem, Theorem 1 is applied, yielding an adjoint function $p_*(t)$ satisfying the maximum condition (7) and the standard adjoint equation (9), as well as the transversality condition (8). The mollifiers are then shrunk "to nothing", and a cluster point of the $p_*(.)$-functions are shown to satisfy the conditions in Theorem 2. A lemma on mollifiers is needed.

Lemma 1 Let X be a separable Banach space. Let f be a Lipschitz continuous function on $B(\hat{x},\delta') \subset X$, with values in a Banach space Z. Let M^f be the Lipschitz rank of f. There exists a sequence of mollifications of f, written $f^{(k)}(x), k = 1, 2, ...,$ defined on $B(\hat{x},\delta'/2)$, such that $|f^{(k)}(x) - f(x)| \leq M^f/k$ for all x in $B(\hat{x},\delta'/2)$. Moreover, on $B(\hat{x},\delta'/2)$, $f^{(k)}(x)$ is Lipschitz continuous of rank $\leq M^f$ and is bounded by $M_f := |f(\hat{x})| + \delta' M^f$. In fact, $f^{(k)}(x)$ is bounded on $B(\hat{x},\delta'/2)$ by any bound $\hat{M}_f$ that f has on $B(\hat{x},\delta')$. Furthermore, $f^{(k)}(x)$ has a Gâteaux derivative $\nabla f^{(k)}(x)$ at each $x \in B(\hat{x},\delta'/2)$ and $|\nabla f^{(k)}(x)| \leq M^f$. Finally, let $\hat{z}^*$ be any element in $clB(0,\beta) \subset Z^*$ and let $D^*(f,\hat{x},\hat{z}^*)$ be the set of points $x^* \in X^*$, with the property that there exists a sequence $(x_k, z_k^*) \in X \times Z^*$, where $x_k \to \hat{x}$, $z_k^* \to^* \hat{z}^*$, $z_k^* \in clB(0,\beta)$, such that x^* is a weak* cluster point of a sequence of convex combinations $\hat{x}_j^* := \sum_{n=1}^{n_j} \theta_i \check{x}^*_{k_n^j}$, all $k_n^j \geq j$, of the elements $\check{x}_k^* := [\nabla f^{(k)}(x_k)]^* z_k^*$. Then $D^*(f,\hat{x},\hat{z}^*) \subset \partial^\beta(f,\hat{x},\hat{z}^*)$, where $\partial^\beta(f,\hat{x},\hat{z}^*) := \{x^* \in X^* : \langle w, x^*\rangle \leq d^\beta(f,\hat{x},\hat{z}^*)(w)$ for all $w \in X\}$; (for d^β, see (10)).

Proof Let $a_0 = 0 \in X$, and let $a_i, i = 1, 2, ...,$ be a sequence of unit vectors in X such that $\cup_m E_m$ is dense in X, where $E_m :=$linspan$\{a_1, ..., a_m\}, m = 1, 2, ...$. Let $\delta \in (0,\delta'/2)$ be given, and define $\alpha_{n,\delta}(\lambda) := \alpha_n(\lambda)$, $n = 1, 2, ...,$ to be a Lipschitz continuous nonnegative function on $(-\infty,\infty)$ vanishing outside $(-\delta/2^n, +\delta/2^n)$, bounded by $1/(\delta/2^n)$ and with $\int_{\mathbb{R}} \alpha_n(\lambda)d\lambda = 1$. The function is taken to be piecewise linear. (By necessity, the Lipschitz rank goes to infinity with n.)

Write $\lambda^n := (\lambda_1, ..., \lambda_n), \alpha^n(\lambda^n) = \alpha_1(\lambda_1) \cdot .... \cdot \alpha_n(\lambda_n)$ and let $\int^m$ denote an m-tuple integral over $\mathbb{R}^m$. For any $x \in B(\hat{x},\delta'/2)$, define the n- multiple integral $I^{\delta,f}_{n,k}(x) := I_{n,k}(x) := \int^n f(x - \sum_{i=0}^k \lambda_i a_i)\alpha^n(\lambda^n)d\lambda^n$, $k \leq n$. Note that $\alpha^n(\lambda^n)$ is non-vanishing only if $|\lambda_i| < \delta/2^i$, $i = 1, ..., n$, and $|\sum_{i=0}^k \lambda_i a_i| \leq \delta$ for such λ_i, so the calculation of $I_{n,k}(x)$ involves only values of f on $B(\hat{x},\delta'/2+\delta)$, when $x \in B(\hat{x},\delta'/2)$. Now, $|I_{n+1,n+1}(x) - I_{n,k}(x)| \leq$

$$\int^n | \int^1 f(x - \textstyle\sum_{i=0}^{n+1} \lambda_i a_i)\alpha_{n+1}(\lambda_{n+1})d\lambda_{n+1} - f(x - \sum_{i=0}^k \lambda_i a_i)|\alpha^n(\lambda^n)d\lambda^n \leq$$
$$\int^n \{\int^1 |f(x - \textstyle\sum_{i=0}^{n+1} \lambda_i a_i) - f(x - \sum_{i=0}^k \lambda_i a_i)|\alpha_{n+1}(\lambda_{n+1})d\lambda_{n+1}\}\alpha^n(\lambda^n)d\lambda^n \leq$$
$$\int^n (\int^1 \{M^f | \textstyle\sum_{i=k+1}^{n+1} \lambda_i a_i|\}\alpha_{n+1}(\lambda_{n+1})d\lambda_{n+1})\alpha^n(\lambda^n)d\lambda^n \leq$$

$$\int^n (\int^1 \{M^f \sum_{i=k+1}^{n+1} \delta/2^i\} \alpha_{n+1}(\lambda_{n+1}) d\lambda_{n+1}) \alpha^n(\lambda^n) d\lambda^n \le M^f \delta/2^k. \qquad (25)$$

To obtain the next to last inequality, note that when integrating with respect to λ_i , the integration can be confined to $(-\delta/2^i), \delta/2^i)$. Write $I_n^{\delta,f}(x) := I_{n,n}^{\delta,f}(x)$. By (25), $|I_{n+1}^{\delta,f}(x) - I_n^{\delta,f}(x)| \le M^f \delta/2^n$, so, for each $x \in B(\hat{x}, \delta'/2)$, $\{I_n^{\delta,f}(x)\}_n$ is a Cauchy sequence, with limit $f^{(\delta)}(x)$, $x \in B(\hat{x}, \delta'/2)$. When $\delta = 1/k$, we write $f^{(k)}(x)$ instead of $f^{(1/k)}$. We shall show that this sequence has the properties claimed in the lemma. Letting $k = 0$ in (25) yields $|I_{n+1}^{\delta,f}(x) - f(x)| \le M^f \delta, (f(x) = I_{n,0}^{\delta,f}(x))$, so, in the limit, for all $x \in B(\hat{x}, \delta'/2)$, $|f^{(\delta)}(x) - f(x)| \le M^f \delta$.

It is trivial that $I_n^{\delta,f}(x)$ is bounded by $\hat{M}_f$, for $x \in B(\hat{x}, \delta'/2)$. It is then easily seen that in $B(\hat{x}, \delta'/2)$, $I_n^{\delta,f}(x)$ has a Lipschitz rank $\le M^f$, taking limits, also $f^{(\delta)}$ is seen to have a Lipschitz rank $\le M^f$. The first claim in fact follows from linearity of mollification: $I_n^{\delta,(\alpha f + \beta h)}(x) = \alpha I_n^{\delta,f}(x) + \beta I_n^{\delta,h}(x)$, so $I_n^{\delta,f(.+z)-f(.)}(x) = I_n^{\delta,f(.+z)}(x) - I_n^{\delta,f(.)}(x)$. Thus, for $\check{x}, \check{x}+z \in B(\hat{x}, \delta'/2)$, since $|f(x+z) - f(x)| \le M^f|z|$ for $x, x+z \in B(\hat{x}, \delta')$, then $|I_n^{\delta,f}(\check{x}+z) - I_n^{\delta,f}(\check{x})| = |I_n^{\delta,f(.+z)}(\check{x}) - I_n^{\delta,f(.)}(\check{x})| = |I_n^{\delta,f(.+z)-f(.)}(\check{x})| \le M^f|z|$.

Next, we turn to the proof of Gâteaux differentiability. Let $x \in B(\hat{x}, \delta'/2)$. Choose some natural number m, and let $w := \sum_{i=0}^m w_i a_i \in E_m$. For $n > m$, write

$$I_n^m(x) := \int^{n-m} f(x - \sum_{i=m+1}^n \lambda_i a_i) \cdot \alpha_{m+1}(\lambda_{m+1}) \cdot .... \cdot \alpha_n(\lambda_n) d\lambda_{m+1} \cdot ... \cdot d\lambda_n,$$

and let $I^m(x) = \lim_{n\to\infty} I_n^m(x)$, (this limit exist for the same reasons as $I(x)$ exists, moreover $I_n^m(x)$ and $I^m(x)$ are bounded by $\hat{M}_f$ and have Lipschitz rank M^f). Note that $I_n^{\delta,f}(x) = \int^m I_n^m(x - \sum_{i=0}^m \lambda_i a_i) \alpha^m(\lambda^m) d\lambda^m$, and by dominated convergence, that $f^{(\delta)} = \int^m I^m(x - \sum_{i=0}^m \lambda_i a_i) \alpha^m(\lambda^m) d\lambda^m$.

Evidently, $(f^{(\delta)}(x+tw) - f^{(\delta)}(x))/t = \lim_n t^{-1}\{I_n^{\delta,f}(x+tw) - I_n^{\delta,f}(x)\} :=$

$$\begin{aligned}
&\lim_n \int^m t^{-1}[I_n^m(x + tw - \textstyle\sum_{i=0}^m \lambda_i a_i) - I_n^m(x - \sum_{i=0}^m \lambda_i a_i)] \alpha^m(\lambda^m) d\lambda^m = \\
&\int^m t^{-1}[I^m(x + tw - \textstyle\sum_{i=0}^m \lambda_i a_i) - I^m(x - \sum_{i=0}^m \lambda_i a_i)] \alpha^m(\lambda^m) d\lambda^m = \\
&\int ... \int t^{-1} I^m(x + \textstyle\sum_{i=0}^m (tw_i - \lambda_i) a_i) \alpha_1(\lambda_1) \cdot .... \cdot \alpha_m(\lambda_m) d\lambda_1 \cdot ... \cdot d\lambda_m - \\
&\int ... \int t^{-1} I^m(x - \textstyle\sum_{i=0}^m \lambda_i a_i) \alpha_1(\lambda_1) \cdot .... \cdot \alpha_m(\lambda_m) d\lambda_1 \cdot ... \cdot d\lambda_m = \\
&\int ... \int t^{-1} I^m(x - \textstyle\sum_{i=0}^m \lambda_i a_i) \alpha_1(\lambda_1 + tw_1) \cdot .... \cdot \alpha_m(\lambda_m + tw_m) d\lambda_1 \cdot ... \cdot d\lambda_m - \\
&\int^m t^{-1} I^m(x - \textstyle\sum_{i=0}^m \lambda_i a_i) \alpha^m(\lambda^m) d\lambda^m = \\
&\int^m t^{-1} I^m(x - \textstyle\sum_{i=0}^m \lambda_i a_i)[\alpha^m(\lambda^m + tw^m) - \alpha^m(\lambda^m)] d\lambda^m =: \zeta(t), \qquad (26)
\end{aligned}$$

where $w^m = (w_1, ..., w_m)$.

Now, $\nabla\alpha^m(\lambda^m)$ exists for a.e. λ^m, hence $\lim_{t\searrow 0}[\alpha^m(\lambda^m+tw^m)-\alpha^m(\lambda^m)]/t = \nabla\alpha^m(\lambda^m)w^m$ exists for a.e. λ^m. By Lipschitz continuity, $[\alpha^m(\lambda^m+tw^m) - \alpha^m(\lambda^m)]/t$ has a bound independent of λ^m and t. Moreover, $I^m(x-\sum_{i=0}^m \lambda_i a_i)$ is similarly bounded, hence, $\lim_{t\searrow 0}\zeta(t)$ exists and equals $\int^m I^m(x-\sum_{i=0}^m \lambda_i a_i)\nabla\alpha^m(\lambda^m)w^m d\lambda^m$. Because the last expression is linear in w^m, $w \to f^{(\delta)}(x+w)$ has a Gâteaux derivative at $w = 0$ on E_m, in fact on $\cup_m E_m$. Moreover, as $f^{(\delta)}$ has the Lipschitz rank M^f, then $|\zeta(t)| \le M^f|w|$, so $|\nabla f^{(\delta)}(x)| \le M^f$. By density of $\cup_m E_m$ in X, at $w = 0, w \to f^{(\delta)}(x+w)$ has a Gâteaux derivative on X, (see Appendix, Lemma A).

Finally, let us prove that $D^*(f,\hat{x},\hat{z}^*) \subset \partial^\beta(f,\hat{x},\hat{z}^*)$. Let w be any given unit vector in X. For any given $\varepsilon > 0$, there exist a $\kappa \in (0,\delta'/4)$ and a weak* neighbourhood W of $\hat{z}^*$ in $clB(0,\beta)$ such that $\langle[f(y+\lambda w)-f(y)]/\lambda, \tilde{z}^*\rangle \le d^\beta(f,\hat{x},\hat{z}^*)(w)+\varepsilon$ when $y \in B(\hat{x},2\kappa), \lambda \in (0,\kappa), \tilde{z}^* \in W$. Let $\delta < \kappa, \lambda < \kappa$. Then, for $x \in B(\hat{x},\kappa), \tilde{z}^* \in W$,

$$\langle[f^{(\delta)}(x+\lambda w)-f^{(\delta)}(x)]/\lambda, \tilde{z}^*\rangle =$$
$$\lim_n \int^n \langle \lambda^{-1}[f(x+\lambda w - \textstyle\sum_{i=0}^n \lambda_i a_i) - f(x-\sum_{i=0}^n \lambda_i a_i)], \tilde{z}^*\rangle \alpha^n(\lambda^n)d\lambda^n \le$$
$$\lim_n \int^n (d^\beta(f,\hat{x},\hat{z}^*)(w)+\varepsilon)\alpha^n(\lambda^n)d\lambda^n = d^\beta(f,\hat{x},\hat{z}^*)(w)+\varepsilon.$$

Hence, letting $\lambda \searrow 0$, we get that $\langle \nabla f^{(\delta)}(x)w, \tilde{z}^*\rangle \le d^\beta(f,\hat{x},\hat{z}^*)(w)+\varepsilon$ for all $x \in B(\hat{x},\kappa), \delta \in (0,\kappa), \tilde{z}^* \in W$.

Let $x^* \in D^*(f,\hat{x},\hat{z}^*)$. Then, by definition, there exists a sequence $(x_k, z_k^*), x_k \to \hat{x}$, $z_k^* \to^* \hat{z}^*, z_k^* \in clB(0,\beta)$, such that x^* is a weak* cluster point of a sequence of convex combination $\hat{x}_j^* := \sum_{n=1}^{n_j} \theta_n \check{x}^*_{k_n^j}$, all $k_n^j \ge j$, of the elements $\check{x}_k^* := [\nabla f^{(k)}(x_k)]^* z_k^*$. For k large, $(x_k, z_k^*) \in B(\hat{x},\kappa) \times W$. Thus, as all $\check{x}_k^*$ satisfies $\langle w, \check{x}_k^*\rangle \le d^\beta(f,\hat{x},\hat{z}^*)(w)+\varepsilon$ for k large, then $\langle w, \hat{x}_j^*\rangle \le d^\beta(f,\hat{x},\hat{z}^*)(w)+\varepsilon$ for j large, hence $\langle w, x^*\rangle \le d^\beta(f,\hat{x},\hat{z}^*)(w)+\varepsilon$. By the arbitraryness of ε and w, $x^* \in \partial^\beta(f,\hat{x},\hat{z}^*)$. □

Lemma 2 Let, in Lemma 1, f be Lipschitz continuous in $B(\hat{x}',\delta'')$ with rank M^f. Then, $d^\beta(f,\hat{x},\hat{z}^*)(w)$ is upper semicontinuous in $(\hat{x},\ \hat{z}^*) \in B(\hat{x}',\delta'') \times clB(0,\beta)$, in the norm $\times$ weak* topology.

Proof: Let $\hat{z}^* \in clB(0,\beta), \hat{x} \in B(\hat{x}',\delta'')$, and let δ' be so small that $B(\hat{x},\delta') \subset B(\hat{x}',\delta'')$. For any given $w \in B(0,1)$, and any given $\varepsilon > 0$, there exist a $\kappa \in (0,\delta')$ and a weak* open neighbourhood W of $\hat{z}^*$ in $clB(0,\beta)$ such that $\langle[f(y+\lambda w)-f(y)]/\lambda, z^*\rangle \le d^\beta(f,\hat{x},\hat{z}^*)(w)+\varepsilon$ when $y \in B(\hat{x},\kappa/2), \lambda \in (0,\kappa/2), z^* \in W$. Let $(\check{x},\check{z}^*) \in B(\hat{x},\kappa/2) \times W$. Evidently, $d^\beta(f,\check{x},\check{z}^*)(w) \le$

$\sup_{y\in B(\hat{x},\kappa/2),\lambda\in(0,\kappa/2),z^*\in W}\langle[f(y+\lambda w)-f(y)]/\lambda,z^*\rangle \le d^\beta(f,\hat{x},\hat{z}^*)(w)+\varepsilon.\square$

Continued proof of Theorem 2 First, we assume that X is separable. For $\delta=1/k<\varsigma/2$, let $g^k(s,x,u):=f^{(k)}(x)$, for $f(x)=g(s,x,u)$. Similarly, write ϕ^k for $\phi^{(\delta)}$, when $\delta=1/k$. Then, by Lemma 1, $|g^k(s,x,u(s))-g(s,x,u(s))|\le M^{u(.)}(s)/k\le M^*(s)/k$, uniformly in $(s,x), x\in B(x^*(s),\varsigma/2)$ and $|\phi^k(x)-\phi(x)|\le M_\phi/k, x\in B(x^*(T),\varsigma/2)$. Let $x^{u,k}(t)$ be the solution of $dx^{u,k}/ds=g^k(s,x,u(s)), x^{u,k}(0)=x_0$. Consider first $u=u^*$. By a local existence theorem, $x^{u^*,k}(.)$ exists in $B(x^*(t),\varsigma/2)$, at least on a small interval $[0,t']$. Write $\alpha(t):=1+\int_0^t M^*(s)ds\exp(\int_0^t M^*(s)ds)$ and note that when $t\le t'$, then $|x^{u^*,k}(t)-x^*(t)|=$

$$\begin{aligned}&|\textstyle\int_0^t g^k(s,x^{u^*,k}(s),u^*(s))-g(s,x^*(s),u^*(s))ds|\le\\ &|\textstyle\int_0^t g^k(s,x^{u^*,k}(s),u^*(s))-g(s,x^{u^*,k}(s),u^*(s))ds|+\\ &|\textstyle\int_0^t g(s,x^{u^*,k}(s),u^*(s))-g(s,x^*(s),u^*(s))ds|\le\\ &(1/k)\textstyle\int_0^t M^*(s)ds+\int_0^t M^*(s)|x^{u^*,k}(s)-x^*(s)|ds.\end{aligned}$$

Let $1/k<\varsigma/4\alpha(T)$. By Gronwall's inequality, $|x^{u^*,k}(t)-x^*(t)|\le\alpha(t)/k\le\alpha(T)/k\le\varsigma/4, t\in[0,t']$. An existence and continuation argument gives that $x^{u^*,k}(t)$ exists on all J in $clB(x^*(t),\varsigma/4)$ and the preceding inequalities hold for $t'=T$. Using this and $|\phi^k(x)-\phi(x)|\le M_\phi/k$ yield

$$\begin{aligned}&|\phi^k(x^{u^*,k}(T))-\phi(x^*(T))|\le|\phi^k(x^{u^*,k}(T))-\phi^k(x^*(T))|+\\ &|\phi^k(x^*(T))-\phi(x^*(T))|\le\xi/k,\end{aligned}\qquad(27)$$

where $\xi:=M_\phi\alpha(T)+M_\phi$.

Define $\sigma_t(u,\hat{u}):=\int_0^t(1+M_*(s))1_{\{\tau:u(\tau)\neq\hat{u}(\tau)\}}(s)ds$. Note that $|\int_0^t(\dot{x}^{\hat{u},k}(s)-\dot{x}^{u,k}(s))ds|=$

$$\begin{aligned}&|\textstyle\int_0^t g^k(s,x^{\hat{u},k}(s),\hat{u}(s))-g^k(s,x^{u,k}(s),u(s))ds|\le\\ &|\textstyle\int_0^t g^k(s,x^{\hat{u},k}(s),\hat{u}(s))-g^k(s,x^{\hat{u},k}(s),u(s))ds|+\\ &|\textstyle\int_0^t g^k(s,x^{\hat{u},k}(s),u(s))-g^k(s,x^{u,k}(s),u(s))ds|\le\\ &2\sigma_t(u,\hat{u})+\textstyle\int_0^t M^*(s)|x^{\hat{u},k}(s)-x^{u,k}(s)|ds,\end{aligned}$$

when $k>4\alpha(T)/\varsigma$. For $\hat{u}=u^*$, by these inequalities, Gronwall's inequality, and an existence and continuation argument, for $u\in D_T:=\{u\in\mathcal{U}, \sigma_T(u,u^*)\le\varsigma/8\exp(\int_0^T M^*(s)ds)\}$, $k>4\alpha(T)/\varsigma$, the solution $x^{u,k}(t)$ exists on all J in $clB(x^{u^*,k}(t),\varsigma/4)\subset clB(x^*(t),\varsigma/2)$. In fact, the inequalities and Gronwall's inequality give, for $u,\hat{u}\in D_T$, that

$$|x^{u,k}(t) - x^{\hat{u},k}(t)| \leq 2\sigma_t(u, \hat{u}) \exp(\int_0^t M^*(s)ds). \tag{28}$$

For the metric σ_T on $\mathcal{U}$, the space $\mathcal{U}$ is complete, (we identify a.e. equal functions). Even the closed ball D_T is complete. Using (27), by Ekeland's variational principle, for any $k > 4\alpha(T)/\varsigma$, there exists a control $u_k \in D_T$, with $\sigma_T(u_k, u^*) \leq (\xi/k)^{\frac{1}{2}}$, which is optimal in the problem

$$\max_{u(.)}\{\phi^k(x^u(T)) - (\xi/k)^{\frac{1}{2}}\sigma_T(u, u_k)\}, \tag{29}$$

subject to

$$dx^u/dt = g^k(s, x, u(s)),\ x(0) = x_0,\ u(.) \in D_T. \tag{30}$$

Below, k is $> 4\alpha(T)/\varsigma$. Let $y^u(t)$ be the solution of $dy/dt = (1+M_*(t))1_{\{\tau:u(\tau)\neq u_k(\tau)\}}(t)$ a.e., $y(0) = 0$, and write $\phi^*(x, y) := \phi^k(x) - (\xi/k)^{\frac{1}{2}}y$. Then u_k maximizes $\phi^*(x^u(T), y^u(T))$ for $u \in D_T$. Applying Theorem 1 in the present situation yields an adjoint function $\tilde{p}_k(t) := (p_k(t), P_k(t))$ such that, for any given $u(.) \in D_T$, for t not in a null set $N_{u(.),u_k(.)}$,

$$\begin{aligned}&\langle (g^k(t, x^{u_k}(t), u(t)), (1 + M_*(t))1_{\{\tau:u(\tau)\neq u_k(\tau)\}}(t)), (p_k(t), P_k(t))\rangle \leq \\ &\langle (g^k(t, x^{u_k}(t), u_k(t)), 0), (p_k(t), P_k(t))\rangle.\end{aligned} \tag{31}$$

Here $p_k(t)$ is an absolutely continuous function from J into X^*, satisfying, for all $w \in X$, for t not in a null set N_w,

$$d\langle w, p_k(t)\rangle/dt = \langle -\nabla g^k(t, x^{u_k}(t), u_k(t))w, p_k(t)\rangle \tag{32}$$

where here, (and below), ∇ denotes a Gâteaux derivative with respect to x. Moreover, $P_k(t) \in \mathbb{R}^* = \mathbb{R}$, $dP_k/dt = 0$ a.e., and

$$(p_k(T), P_k(T)) = (\nabla\phi^k(x^{u_k}(T)), -(\xi/k)^{\frac{1}{2}}) \tag{33}$$

Note that, by Gronwall's inequality, (32) and (33), $|p_k(t)| \leq M_\phi \exp(-\int_T^t M^*(s)ds) =: \beta(t), (M_G = 0)$, so, recalling $\gamma(t) := M^*(t)\beta(t)$, we get $d\langle w, p_k(t)\rangle/dt \in \mathrm{cl}B(0, \gamma(t)|w|)$.

Let $\{b_i\}_{i=1}^\infty$ be dense in the unit ball in X. Note that the derivative $d\langle b_i, p_k(t)\rangle/dt$ exists in a set $J \setminus N^*$, where N^* is a null set independent of i and k, and $\gamma_{i,k}(t) := d\langle b_i, p_k(t)\rangle/dt \in L_2(J, \mathbb{R})$, (more precisely, let $\gamma_{i,k}(t) := 0 \cdot 1_{N^*} + d\langle b_i, p_k(t)\rangle/dt \cdot 1_{J\setminus N^*}$). In fact, for $t \notin N^*, |\gamma_{i,k}(t)| \leq \gamma(t)|b_i| \leq \gamma(t)$, (implying uniform countable addtivity of $E \to \int_E \gamma_{i,k}(s)ds), k = 1, 2, ...$), and

$|\gamma_{i,k}(t)-\gamma_{i',k}(t)|\le\gamma(t)|b_i-b_{i'}|$. Below, a certain subsequence k_m of the sequence $k=1,2,...$ is introduced. By diagonal selection, we can find a subsequence $\gamma_{i,k_{m_j}}(.)$ of $\gamma_{i,k_m}(.)$ such that $\gamma_{i,k_{m_j}}(.)$ is weakly convergent in $L_1(J,\mathbb{R})$ to some $\check{p}^i(.)\in L_1(J,\mathbb{R})$, for all i. We can also assume that $\langle b_i, p_{k_{m_j}}(T)\rangle$ is convergent for each i. For simplicity, assume that these convergence properties hold for the sequence k_m itself. Since $\{\gamma_{i,k_m}(.)\}_m$ is weakly convergent, for each fixed i, a sequence of convex combination of the functions $\gamma_{i,k_m}(.)$ converges in L_1-norm, to $\check{p}^i(.)$. By diagonal selection, a sequence of convex combinations $\{\hat{p}_{i,j}(.)\}_j$, $\hat{p}_{i,j}(.):=\sum_{n=1}^{n_j}\theta_n\gamma_{i,k^j_{m_n}}(.)$, all $k^j_{m_n}\ge j$, exists such that for all i, $\{\hat{p}_{i,j}(.)\}_j$ converges in L_1-norm to $\check{p}^i(.)\in L_1(J,\mathbb{R})$. We may even assume pointwise convergence on a set $J\backslash N'$, N' a null set. Evidently, $\gamma_{i,k}(t)\in clB(0,\gamma(t))$ a.e. $\Rightarrow \check{p}^i(t)\in clB(0,\gamma(t))$ a.e.

Let $N=N'\cup N^*$. For each $t\notin N$, it is easily seen that there exists an element $\check{p}(t)\in X^*$ such that $\check{p}^i(t)=\langle b_i,\check{p}(t)\rangle$ for all i. To give some details of the argument, let b_{i_j} be a subsequence of linearly independent vectors such that, for each k, $b_k\in \text{linspan}\{b_{i_j}: i_j\le k\}=E^k$. On $\cup_k E^k$ we define the linear functional $\check{p}(t)$ by $\check{p}(t)(x)=\sum\beta_j\check{p}^{i_j}(t)$, if $x=\sum\beta_j b_{i_j}$ (a finite sum). Fortunately, consistency holds: If $x=b_i$ and $b_i=\sum\beta_j b_{i_j}$, then, for a.e. t, $\check{p}^i(t)=\lim_k\gamma_{i,k}(t)=\lim_k d\langle\sum\beta_j b_{i_j},p_k(t)\rangle/dt=\sum\beta_j\lim_k d\langle b_{i_j},p_k(t)\rangle/dt=\sum\beta_j\lim_k\gamma_{i_j,k}(t)=\sum\beta_j\check{p}^{i_j}(t)=\check{p}(t)(b_i)$. From now on, we write $\langle x,\check{p}(t)\rangle$ instead of $\check{p}(t)(x)$. For all i, $|\langle b_i,\check{p}(t)\rangle|\le\gamma(t)$ for $t\notin N$. By density of $\{b_i\}_{i=1}^\infty$ in the unit ball in X and Lipschitz continuity of $x\to\langle x,\check{p}(t)\rangle$ on the set $\{b_i\}_{i=1}^\infty$, $\check{p}(t)$ has an extension to all X, such that $|\langle x,\check{p}(t)\rangle|\le\gamma(t)|x|$ for all $x\in X, t\notin N$.

Let $\hat{p}^k(t)\in X^*$ be defined by $\langle b_i,\hat{p}^k(t)\rangle=\gamma_{i,k}(t)$, (the extension to X is again trivial). Then $|\langle x,\hat{p}^k(t)\rangle|\le\gamma(t)|x|$ for all $x\in X, t\notin N^*$. Evidently, by density of the b_i's, $\check{p}(t)$ is a weak* limit point of the sequence $\hat{p}_j(t):=\sum_{n=1}^{n_j}\theta_n\hat{p}^{k^j_{m_n}}(t), t\notin N$. Moreover, by density of the b_i's, $p_{k_m}(T)$ converges weakly* to some limit p_T and, for any $t, p(t):=p_T+\int_T^t\check{p}(s)ds$ is a weak* limit of $p_{k_m}(t)$, (by weak convergence of $\{\gamma_{i,k_m}(.)\}_m$). All the properties in Remark 1 hold for $p(.)$. Moreover, by applying Lemma 1 to ϕ, (with $Z=\mathbb{R},\hat{z}^*=1$), it is obtained that p_T belongs to $\partial^{M_\phi}(\phi,x^*(T),1)=\partial\phi(x^*(T))$.

Now, a subsequence $u_{k_m}(.)$ of $u_k(.)$ satisfies $\sigma_T(u_{k_m},u^*)\le 1/2^{m+1}$. Then $C_j=\cup_{m\ge j}\{s:u_{k_m}(s)\ne u^*(s)\}$ satisfies $\text{meas}(C_j)\le 1/2^j$, by definition of σ_T. Let $C=\cap_j C_j$. For simplicity write $u_m(.)=u_{k_m}(.), x_m(.)=x_{k_m}(.), P_m(.)=P_{k_m}(.)$, and $p_m(.)=p_{k_m}(.)$. Observe that, for any t, for any $\varepsilon'>0$,

$$|g^{k_m}(t,x^{u_m}(t),u(t))-g(t,x^*(t),u(t))|\le\varepsilon' \qquad (34)$$

when m is large, uniformly in $u(.)$. To see this, note that, for any t,

$|g^{k_m}(t, x^{u_m}(t), u(t)) - g(t, x^*(t), u^*(t))| \leq$
$|g^{k_m}(t, x^{u_m}(t), u(t)) - g^{k_m}(t, x^*(t), u(t))| + |g^{k_m}(t, x^*(t), u(t)) - g(t, x^*(t), u(t))|.$

The first term is small by Lipschitz continuity and the fact that $x^{u_m}(t) \to x^*(t)$, (see (28)), and the second term is small due to the construction of the mollifier, (see Lemma 1). Moreover, for any $\varepsilon'' > 0$ and for any t, for m large enough,

$$(1 + M_*(t))(\xi/k_m)^{1/2} \leq \varepsilon'' \text{ , so } |P_m(t)|(1 + M_*(t)) \leq \varepsilon'', \tag{35}$$

by (33). Let $\varepsilon > 0, \varepsilon$ arbitrary. For any $t \notin N_{u(.)}, N_{u(.)} := \cup_k N_{u(.),u_k(.)}$, for m large, the following inequalities can be shown:

$$\begin{aligned}&\langle g(t, x^*(t), u(t)), p_m(t)\rangle - 2\varepsilon \leq \langle g^{k_m}(t, x^{u_m}(t), u(t)), p_m(t)\rangle - \varepsilon \leq\\ &\langle g^{k_m}(t, x^{u_m}(t), u(t)), p_m(t)\rangle + (1 + M_*(t))1_{\{\tau:u(\tau)\neq u_m(\tau)\}}(t)P_m(t) \leq\\ &\langle (g^{k_m}(t, x^{u_m}(t), u_m(t)), p_m(t)\rangle \leq \langle g(t, x^*(t), u_m(t)), p_m(t)\rangle + \varepsilon.\end{aligned}$$

The first inequality follows from (34), the second one from (35), the third one from
optimality of $u_m(.), ((31))$, and the fourth one from (34). (Also the bound $\beta(t)$ on all functions $p_m(t)$ has been used.)

Now, for any $t \notin C \cup N_{u(.)}$, for some j_t, $t \notin C_{j_t}$, and $u_m(t) = u^*(t)$ for

$m \geq j_t$, so

$$\langle g(t, x^*(t), u(t)), p_m(t)\rangle - 2\varepsilon \leq \langle g(t, x^*(t), u^*(t)), p_m(t)\rangle + \varepsilon \text{ for } m \geq j_t, \tag{36}$$

m large, and, by (32),

$$\hat{p}^{k_m}(t) = -[\nabla g^{k_m}(t, x^{u_m}(t), u^*(t))]^* p_m(t) \text{ for } m \geq j_t. \tag{37}$$

From (36), it follows that the cluster point $p(t)$ must also satisfy $\langle g(t, x^*(t), u(t)), p(t)\rangle - 2\varepsilon \leq \langle (g(t, x^*(t), u^*(t)), p(t)\rangle + \varepsilon$, for $t \notin C \cup N_{u(.)}$. By the arbitraryness of ε, (7) holds for all $u(.) \in D_T$, for $t \notin C \cup N_{u(.)}$. But then, (7) holds a.e. for any $\hat{u}(.)$ in $\mathcal{U}$, since it holds for $u(.) = \hat{u}(.)1_I + u^*(.)(1 - 1_I)$, where I is any interval so small that $u(.) \in D_T$. By (37) and the last part of Lemma 1, the limit $\check{p}(t)$ satisfies (20') for a.e. t, which implies (11).

Let us now remove the assumption of separability of X. Let U' be a finite set of controls, let $U'' := \{\sum_j 1_{M_j} u_j(.) : u_j(.) \in U'$, $\{M_j\}$ is a finite measurable partition of $J\}$, (U'' is the "switching closure" of U'), and let U''' be the subset of U'' obtained by requiring the M_j's to be intervals with rational end points. (Then U''' is countable.) Let U^* be the $\sigma-$closure of the set U''. Then there exists a countable subset $X^{U'}$ of X, containing x_0, such that, a.e., $\dot{x}^u(t)$ belongs to $\mathrm{cl}X^{U'}$ for all $u := u(.) \in D_T \cap U'''$. We shall assume that the subset $X^{U'}$ is so chosen that $\mathrm{linspan}X^{U'} \subset \mathrm{cl}X^{U'}$. This can always be arranged by replacing $X^{U'}$ by the countable set of finite sums $\mathcal{Q}(X^{U'}) := \{\sum \lambda_i x_i : \lambda_i$ rational, $x_i \in X^{U'}\}$. By (28), Lipschitz continuity of g in x, and $\sigma-$density of U''' in U^* , for all $u \in U^* \cap D_T$, $\dot{x}^u(t) \in \mathrm{cl}X^{U'}$ for a.e. t. Then also $x^u(t) \in \mathrm{cl}X^{U'}$ for such u. For simplicity, assume that an open set A exists such that $B(x^*(t), \varsigma/4) \subset A \subset B(x^*(t), \varsigma)$ for all t. (At least, there exist a finite number of points t_i, $i = 1, ..., i^*$, increasing in i, $t_0 = 0, t_{i^*+1} = T$, such that $B(x^*(t), \varsigma/4) \subset A_i := B(x^*(t_i), \varsigma/2) \subset B(x^*(t), \varsigma)$, for $t \in [t_i, t_{i+1}] =: J_i$. The below construction can then be carried out on each J_i.) For each $x \in X^{U'} \cap A$, for a.e. t, for all $u(.) \in U''', g(t, x, u(t))$ takes values in the closure $\mathrm{cl}X_x$, of some countable set X_x. Let $\check{X}^1_{U'} := \mathcal{Q}(X^{U'} \cup \{\cup_{x \in X^{U'} \cap A} X_x\})$. Assuming the countable set $\check{X}^k_{U'}$ defined, then, by induction, let for each $x \in \check{X}^k_{U'} \cap A$, for a.e. t, for all $u(.) \in U''', g(t, x, u(.))$ take values in the closure $\mathrm{cl}X^k_x$, where X^k_x is countable, and let $\check{X}^{k+1}_{U'} = \mathcal{Q}(\check{X}^k_{U'} \cup \{\cup_{x \in \check{X}^k_{U'} \cap A} X_x\})$. Finally, let $\check{X}_{U'} := \cup_{k=1}^{\infty} \check{X}^k_{U'}$. Note that $\mathcal{Q}(\check{X}_{U'}) = \check{X}_{U'}$, ($\check{X}^k_{U'}$ is increasing in k), and note that for any x in the countable set $\check{X}_{U'} \cap A$, for a.e. t, for all $u(.) \in U''', g(t, x, u(t))$ takes values in $\mathrm{cl}\check{X}_{U'}$. By continuity in x, for t not in a null set $N_{U'}$, ($N_{U'}$ not dependent on x), for all $x \in (\mathrm{cl}\check{X}_{U'}) \cap B(x^*(t), \varsigma/4)$, for all $u(.) \in U''', g(t, x, u(t))$ takes values in $\mathrm{cl}\check{X}_{U'}$. (We in this case say that $\mathrm{cl}\check{X}_{U'}$ is g, U'-invariant.) This even holds for all $u(.) \in U^*$, for t not in a null set $N_{U',u(.)}$ independent of x.

Let $\mathcal{U}'$ be the family of finite sets U' that contain $u^*(.)$. For any such finite set U', and any countable subset V of X, let $X^{U',V}$ be the set $\mathrm{cl}\check{X}_{U'}$ obtained by including V in the set $X^{U'}$ with which we started the above construction. If, in the definition of $d_x^{\beta(t)}(g(t, x, u), p)(w)$, see (10), y is restricted to belong to a subset X' of X, we write $d_x^{\beta(t),X'}(g(t, x, u), p)(w)$. The latter expression is written $d^{X'}(t, p(t))(w)$ when $p = p(t), u = u^*(t), x = x^*(t)$. Applying the necessary conditions for the separable case to U^* instead of $\mathcal{U}$, with $X^{U',V}$ as state space, and using the notation of Remark 1, we obtain a $p^{T,U',V} \in (X^{U',V})^*$ and a $\dot{p}^{U',V}(.) \in L_2^s(J, (X^{U',V})^*)$, the topscript s indicating scalarwise integrability, such that (39) below holds, such that (7) holds for $p(.)$ replaced by $p^{U',V}(.)$ (given in (41) below), for all $u(.) \in U^*$, a.e., and such that, for a.e. t, $|\dot{p}^{U',V}(t)| \leq \gamma(t)$. Thus, we have,

$$\int_J \langle x(t), \dot{p}^{U',V}(t)\rangle dt \leq \int_J \gamma(t)|x(t)|dt \text{ for all } x(.) \in L_\infty(J, X^{U',V}), \quad (38)$$

for all $w \in X^{U',V}$, for a.e. t, $-\langle w, \dot{p}^{U',V}(t)\rangle \leq d^{X^{U',V}}(t, p^{U',V}(t))(w)$,
(39a)

$$\langle w, p^{T,U',V}\rangle \leq d^{0,X^{U',V}}\phi(x^*(T))(w) \text{ for all } w \in X^{U',V}, \quad (39b)$$

where $d^{0,X^{U',V}}\varphi(x^*(T))(w)$ is the generalized directional derivative at $x^*(T)$ of ϕ restricted to $X^{U',V}$. Below, we will apply the integrated version of the maximum condition, namely:

$$\int_J \langle g(t, x^*(t), u(t)), p^{U',V}(t)\rangle dt \leq \int_J \langle g(t, x^*(t), u^*(t)), p^{U',V}(t)\rangle dt, u(.) \in U^*. \quad (40)$$

Of course,

$$p^{U',V}(t) = p^{T,U',V} + \int_T^t \dot{p}^{U',V}(s)ds. \quad (41)$$

The function $\dot{p}^{U',V}(.)$ represents a continuous linear functional on $L_2(J, X^{U',V})$. By the Hahn-Banach Theorem, this functional has an extension to all $L_2(J, X)$, preserving its norm $|\gamma(.)|_2$. By a general representation theorem (Dunford-Pettis theorem), the extended functional can be represented by a function $\dot{p}_{U',V}(.) \in L_2^s(J, X^*)$, with $|\dot{p}_{U',V}(.)|_2 \leq |\gamma(.)|_2$, see e.g. Fattorini (1999), p.668. Similarly, the functional $p^{T,U',V}$ has an extension $p_{T,U',V}$ to X, preserving its norm M_ϕ. Let Γ be the directed set of pairs (U', V), $U' \in \mathcal{U}'$, V a countable set in X, ordered by the relation $(\check{U}', \check{V}) \succeq (U', V)$ iff $\check{U}' \supset U', \check{V} \supset V$. To each $\gamma := (U', V)$, there corresponds a pair $(p_\gamma^T, \dot{p}_\gamma(.)) := (p_{T,U',V}, \dot{p}_{U',V}(.))$. Since $\{(p_\gamma^T, \dot{p}_\gamma(.)) : \gamma \in \Gamma\}$ is bounded in $X^* \times L_2^s(J, X^*)$, the generalized sequence $(p_\gamma^T, \dot{p}_\gamma(.)), \gamma \in \Gamma$, has a weak* cluster point $(p^T, \dot{p}(.)) \in X^* \times L_2^s(J, X^*)$. Let $p(t) := p^T + \int_T^t \dot{p}(s)ds$. Then, for each t, the generalized sequence $p_\gamma(t) := p_\gamma^T + \int_T^t \dot{p}_\gamma(s)ds$, has $p(t)$ as a weak* cluster point, so (38),(39b) and (41) immediately imply two of the three properties below, namely (42) and (44).

$$\int_J \langle x(t), \dot{p}(t)\rangle dt \leq \int_J \gamma(t)|x(t)|dt \text{ for all } x(.) \in L_\infty(J, X). \quad (42)$$

For all $w \in X, -\langle w, \dot{p}(t)\rangle \leq d^X(t, p(t))(w)$ for t not in a null set N_w.
(43)

$$p(t) = p^T + \int_T^t \dot{p}(s)ds,\ \langle w, p^T\rangle \leq d^{0,X}\phi(x^*(T))(w) \text{ for all } w \in X. \quad (44)$$

(The inequality in (42) follows from the fact that , for any $x(.) \in L_\infty(J, X)$, there exists a countable set V with the property that $x(t) \in \mathrm{cl}V$ a.e., so for such a V, $x(.) \in L_\infty(J, X^{U',V})$ and (38) holds.) When $p(.)$ is suitably chosen, $|\dot{p}(.)|$ can be assumed to be measurable, in fact all properties in Remark 1 hold. For any $u(.) \in \mathcal{U}$, (40) holds for any $\gamma \succeq (\{u(.), u^*(.)\}, \emptyset)$, so (40) holds for any such $u(.)$, for the cluster point $p(t)$. Hence, (7) holds.

Finally, let us prove (43). (We don't want to address the question if $d^X(t, p(t))(w)$ is measurable, this explains part of the route of proof taken). At this point we need the property that if $\check{U}' \subset U'$, $\check{V} \subset V$, then $X^{\check{U}',\check{V}} \subset X^{U',V}$. This can be assumed to hold: We can assume that when we chose $X^{U'}$, for all subsets $\check{U}' \subset U'$, we arranged it so that $X^{\check{U}'} \subset X^{U'}$.

Let $w \in X, X_0 = X^{\{u^*\},\{w\}}$, let $\{x_n^0(.)\}_{n=0}^\infty$ be dense in $L_2(J, X_0)$ and let $\{x_n^0\}_{n=0}^\infty$ be dense in X_0. There exists a pair $(p_1^T, \dot{p}_1(.)) := (p_{\gamma_1}^T, \dot{p}_{\gamma_1}(.))$, $\gamma_1 \succeq (\{u^*\}, \{w\})$, such that the following condition holds for $k = 1$: For $i, n \in \{0, ..., k-1\}$,

$$|\int \langle x_n^i(t), \dot{p}_k(t)\rangle dt - \int \langle x_n^i(t), \dot{p}(t)\rangle dt| \leq 1/k, |\langle x_n^i, p_1^T\ \rangle - \langle x_n^i, p^T\ \rangle| \leq 1/k, \quad (45)$$

Define by induction elements $\gamma_{k''} = (U'_{k''}, V_{k''}) \in \Gamma, k'' \in \{1, ..., k\}$, $\gamma_{k''+1} \succeq \gamma_{k''}$, closed separable $g, U'_{k''-1}$-invariant linear subspaces $X_{k''-1} := X^{U'_{k''-1}, V_{k''-1}}$ increasing in k'', and sequences $\{x_n^{k''-1}(.)\}_{n=0}^\infty, \{x_n^{k''-1}\}_{n=0}^\infty$ dense in $L_2(J, X_{k''-1})$ and $X_{k''-1}$, respectively, such that (45) holds for k replaced by any $k'' \in \{1, ..., k\}$, with $(p_{k''}^T, \dot{p}_{k''}(.)) := (p_{\gamma_{k''}}^T, \dot{p}_{\gamma_{k''}}(.))$. Thus, given that (45) holds for $k'' \leq k$, then, let $X_k = X^{\gamma_k}$, let $\{x_n^k(.)\}_{n=0}^\infty$ be dense in $L_2(J, X_k)$ and $\{x_n^k\}_{n=0}^\infty$ be dense in X_k, choose a $\gamma_{k+1} \succeq \gamma_k$ such that (45) holds for $k+1$, for $\dot{p}_{k+1}(.) := \dot{p}_{\gamma_{k+1}}(.), p_k^T = p_{\gamma_{k+1}}^T$.

Let $Q := |\gamma(.)|_2$ and $X^0 := \mathrm{cl}\cup_k X_k$, and observe that for a.e. t, $x \in (\cup_k X_k) \cap B(x^*(t), \varsigma/4) \Rightarrow g(t, x, u^*(t)) \in \cup_k X_k \subset X^0$, so for a.e. t, $x \in X^0 \cap B(x^*(t), \varsigma/4) \Rightarrow g(t, x, u^*(t)) \in X^0$. Note that for any $x(.) \in L_2(J, X^0)$, for any $\varepsilon > 0$, there exists a step function $\hat{x}(.)$ with values in $\cup_k X_k$, such that $|x(.) - \hat{x}(.)|_2 \leq \varepsilon/6Q$, in fact $\hat{x}(J) \subset X_k$ for some $k = k'$. By density of $\{x_n^{k'}(.)\}_n$ in $L_2(J, X_{k'})$, for some $n = n'$, $|\hat{x}(.) - x_{n'}^{k'}(.)|_2 \leq \varepsilon/6Q$, so $|x(.) - x_{n'}^{k'}(.)|_2 \leq \varepsilon/3Q$. For $k > \max\{k', n'\}$, so large that $1/k \leq \varepsilon/3$, by (45), $|\int_J \langle x_{n'}^{k'}(t), \dot{p}_k(t)\rangle dt - \int_J \langle x_{n'}^{k'}(t), \dot{p}(t)\rangle dt| \leq \varepsilon/3$. Evidently,

$|\int_J\langle x(t),\dot p_k(t)\rangle dt - \int_J\langle x(t),\dot p(t)\rangle dt| = |\int_J\langle x(t),\dot p_k(t)\rangle dt - \int_J\langle x^{k'}_{n'}(t),\dot p_k(t)\rangle dt| +$
$|\int_J\langle x^{k'}_{n'}(t),\dot p_k(t)\rangle dt - \int_J\langle x^{k'}_{n'}(t),\dot p(t)\rangle dt| + |\int_J\langle x^{k'}_{n'}(t),\dot p(t)\rangle dt - \int_J\langle x(t),\dot p(t)\rangle dt| \le$
$\varepsilon/3+\varepsilon/3+\varepsilon/3$

$=\varepsilon$. Hence, $\{\dot p_k(.)\}$, for each t restricted to X^0, is weakly* convergent in $L_2^s(J,(X^0)^*)$, i.e. for the duality $L_2(J,X^0), L_2^s(J,(X^0)^*)$. Denote the restrictions $\{\dot p'_k(.)\}$, and let $\dot p'(t)$ having a corresponding meaning. Similarly, $\{p_k^T\}$, restricted to X^0 is weakly* convergent, denote the restrictions $\{p'^T_k\}$. This implies that for each t, the sequence of corresponding restrictions $\{p'_k(t)\}$ is weakly* convergent in $(X^0)^*$, (i.e. for the duality $X^0,(X^0)^*$). Now, for any $k,m,\ k\ge m$, we have $-\langle w,\dot p'_k(t)\rangle \le \sup_{n\ge m} -\langle w,\dot p'_n(t)\rangle := \alpha_m(t)$, and so $-\int_J\langle w,\dot p'_k(t)\rangle\chi(t)dt \le \int_J\alpha_m(t)\chi(t)dt$, for any $\chi(t)\in L_\infty(J,\mathbb{R}), \chi(t)\ge 0$ a.e. The next to last inequality also holds in the limit, which entails $-\langle w,\dot p'(t)\rangle \le \alpha_m(t)$ a.e. Since this holds for all m, then for a.e. $t,-\langle w,\dot p(t)\rangle = -\langle w,\dot p'(t)\rangle$ $\le \lim_m\alpha_m(t) \le \limsup_n\{d^{X_n}(t,p'_n(t))(w)\} \le \limsup_n\{d^{X^0}(t,p'_n(t))(w)\} \le d^{X^0}(t,p'(t))(w) = d^{X^0}(t,p(t))(w) \le d^X(t,p(t))(w)$. The last equality follows from the fact that, a.e., $g(t,x,u^*(t))\in X^0$ when $x\in B(x^*(t),\varsigma/4)\cap X^0$, the next to last inequality follows from $p'(t)=\lim_n p'_n(t)$ (weak*) and upper semicontinuity, (see Lemma 2), and the next to first one from (39a). Hence, (43), i.e. (11), holds. □

For later use, let us make the observation that, from the arguments above, it follows that for any $w\in X$, there exists a closed separable subspace X_w containing w, and a null set N_w such that the following properties hold. For all $t\notin N_w$,

$$-\langle w,\dot p(t)\rangle \le d_x^{\beta(t),X_w}(g(t,x^*(t),u^*(t)),p(t))(w) \text{ and } g(t,X_w\cap B(x^*(t),\varsigma/4),u^*(t))\subset X_w \tag{46}$$

Proof of Theorem 3, (the end constrained case)
It is possible to reduce the proof to the case where $G=I$, (the identity map). The general case is then obtained by applying the results for this case to a control problem where J is replaced by $[0,T+1]$, and in which an auxiliary control variable z is introduced, governed by $\dot z = G(x)1_{(T,T+1]}(t)$, $z(0)=0$, with $g=0$ on $(T,T+1]$, and with end condition $(x(T+1),z(T+1))\in X\times C$.

Thus, assume that $G=I,(Y=X)$. For any given $\varepsilon>0$, define

$$\Psi(x) := \max(0,\phi(x^*(T))-\phi(x)+\varepsilon^2)+\text{dist}(x,C) \tag{47}$$

Now, for any system pair $(x(.), u(.))$ with $u(.) \in D_T$, we have $\Psi(x(T)) > 0$, (otherwise, if $\Psi(x(T)) \leq 0$ then $x(T)$ satisfies $x(T) \in C$, and $\phi(x^*(T))+ \varepsilon^2 \leq \phi(x(T))$, which contradicts the optimality of $x^*(.)$). Since $\Psi(x^*(T)) = \varepsilon^2$, then $\Psi(x^*(T)) \leq \Psi(x(T)) + \varepsilon^2$ for all system pairs $(x(.), u(.))$. By Ekeland's theorem, there exists a system pair $(x_\varepsilon(.), u_\varepsilon(.))$, $u_\varepsilon(.) \in D_T$, such that $\sigma_T(u_\varepsilon(.), u^*(.)) \leq \varepsilon$ and $u_\varepsilon(.)$ minimizes

$$\Psi^*(u(.)) := \Psi(x^u(T)) + \varepsilon\sigma_T(u(.), u_\varepsilon(.)) \tag{48}$$

for $u(.)$ in D_T. Thus, $u_\varepsilon(.)$ gives minimum in a free end problem with criterion Ψ^*. Let $\varepsilon = 2^{-m}$, and then write $u_\varepsilon(.) = u_m(.)$.

For a free end problem we have already found necessary conditions. To apply them, the auxiliary state y has to be reintroduced. Let $y^u(t)$ be the solution of

$$dy/dt = (1 + M_*(t))1_{\{\tau:u(\tau)\neq u_m(\tau)\}}(t), \text{ a.e. } , y(0) = 0,$$

and write $\phi^*(x, y) := \Psi(x) + y2^{-m}$. Then u_m minimizes $\phi^*(x^u(T), y^u(T))$ for $u \in D_T$. Now, $x^{u_m}(t) \in B(x^*(t), \varsigma/2)$, by definition of D_T. Applying Theorem 2 to $u_m(.), \varsigma/2$ instead of $u^*(.), \varsigma$, with (46), yields an adjoint function $\tilde{p}_m(t) := (p_m(t), P_m(t))$ such that, for any $u(.) \in D_T$, for a.e. t,

$$\begin{aligned} &\langle (g(t, x^{u_m}(t), u(t)), (1 + M_*(t))1_{\{\tau:u(\tau)\neq u_m(\tau)\}}(t)), (p_m(t), P_m(t)\rangle \leq \\ &\qquad \langle (g(t, x^{u_m}(t), u_m(t)), 0), (p_m(t), P_m(t)\rangle. \end{aligned} \tag{49}$$

Here $p_m(t)$ is an absolutely continuous function from J into X^*, with scalarwise derivative $\dot{p}(t)$, such that $p_m(t) = p_m(T) + \int_T^t \dot{p}_m(s)ds$, such that

$$\textstyle\int_J \langle x(t), \dot{p}_m(t)\rangle dt \leq \int_J \gamma(t)|x(t)|dt \text{ for all } x(.) \in L_\infty(J, X), \tag{50}$$

and such that, for all $w \in X$, a closed separable subspace X_w^m exists, such that, for t not in a null set N_w^m,

$$\begin{aligned} &-\langle w, \dot{p}_m(t)\rangle \leq d_x^{\beta(t),X_w^m}(g(t, x^{u_m}(t), u_m(t)), p_m(t))(w), \\ &g(t, X_w^m \cap B(x^*(t), \varsigma/8), u^*(t)) \subset X_w^m. \end{aligned} \tag{51}$$

Moreover, $dP_m/dt = 0$ a.e., and $(p_m(T), P_m(T)) = -(\partial\Psi(x^{u_m}(T)), 2^{-m})$.

(Note that Ψ is Lipschitz continuous of rank $\leq M_\phi + 1$.)

Now, at any $x = \hat{x} \in B(x^*(T), \varsigma), \partial\Psi(\hat{x}) \subset \partial \max(0, \phi(x^*(T)) - \phi(\hat{x}) + \varepsilon^2) + \partial\text{dist}(\hat{x}, C)$. Hence, any $x^* \in \partial\Psi(\hat{x})$ can be written $x^* := \hat{x}^* + \check{x}^*$, where $\hat{x}^* \in y^0\partial(-\phi(\hat{x}))$, $y^0 \in [0,1]$, and $\check{x}^* \in X^*$ satisfies $|\check{x}^*| \leq 1$, and $\langle c, \check{x}^*\rangle \leq \langle \hat{x}, \check{x}^*\rangle$ for all $c \in C$.

Thus, write $p_m(T) = y_m^0 \hat{p}_m^T + \check{p}_m^T$, where $\hat{p}_m^T \in \partial\phi(x^{u_m}(T)), y_m^0 \in [0,1]$, $|\check{p}_m^T| \leq 1, \check{p}_m^T \in X^*$, and $\langle c, \check{p}_m^T\rangle \geq \langle x^{u_m}(T), \check{p}_m^T\rangle$ for all $c \in C$. Since $\Psi(x^{u_m}(T)) > 0$, either $\phi(x^*(T)) - \phi(x^{u_m}(T)) + \varepsilon^2 > 0$, in which case $y_m^0 = 1$, or $\text{dist}(x^{u_m}(T), C) > 0$, in which case $\check{p}_m^T \neq 0$. Normalizing, we can assume $y_m^0 + |\check{p}_m^T| = 1$ for all m.

Choose a convergent subsequence $y_{m_j}^0$ of $y_m^0, j = 1, 2, ...,$ and let $(y^0, \hat{p}^T, \check{p}^T, \dot{p}(t))$ be a weak* cluster point of $(y_{m_j}^0, \hat{p}_{m_j}^T, \check{p}_{m_j}^T, \dot{p}_{m_j}(t))$ in $\mathbb{R} \times X^* \times X^* \times L_2^s(J, X^*)$. Then $y^0 \in [0,1], \hat{p}^T \in \partial\phi(x^*(T))$, and $\check{p}^T$ satisfies $\langle c, \check{p}^T\rangle \geq \langle x^*(T), \check{p}^T\rangle$ for all $c \in C$. Moreover, if $y^0 = 0$, then for all large j, $|\check{p}_{m_j}^T| > 1/2$. Now, (13)-(16) are satisfied for $y^* = \check{p}_m^T, x(.) = x_{m_j}(.), \hat{u}(.) = u_{m_j}(.), p(.) = p_{m_j}(.)$ for j large, in particular, $y_{m_j}^0|\hat{p}_{m_j}^T| < \varepsilon'$ for j large, so (16) holds. Then, by (17), for all large j, $\langle y, \check{p}_{m_j}^T\rangle \geq \varepsilon'$, hence $\langle y, \check{p}^T\rangle \geq \varepsilon'$. Hence, in any case, $(y^0, \check{p}^T) \neq 0$.

Evidently, for each t, $p(t) := y^0\hat{p}^T + \check{p}^T + \int_T^t \dot{p}(s)ds$ is a weak* cluster point of $p_{m_j}(t)$. It is easily seen that the integrated version of the maximum condition holds for the cluster point $p(t)$, so the pointwise maximum condition (7) is satisfied by $p(t)$. Furthermore, (50) entails the inequality $\int\langle x(t), \dot{p}(t)\rangle dt \leq \int \gamma(t)|x(t)|dt$ for all $x(.) \in L_\infty(J, X)$ for the cluster point $p(.)$ so $|\dot{p}(t)| \leq \gamma(t)$ a.e. holds. (Again $\dot{p}(.)$ can be chosen such that $|\dot{p}(.)|$ is measurable, in fact all properties of Remark 1 then hold.) Finally, consider the "adjoint inequality" (11). Let $w \in X$ be given. For each $j = 1, 2, ...,$ we can imagine that a null set N_j and a closed separable subspace X_j^0 containing w is constructed such that for all $t \notin N_j$, $-\langle w, \dot{p}_{m_j}(t)\rangle \leq d_x^{\beta(t), X_j^0}(g(t, x_{m_j}(t), u_{m_j}(t)), p_{m_j})(w)$ and $g(t, X_j^0 \cap B(x^{u_{m_j}}(t), \varsigma/8), u_{m_j}(t)) \subset X_j^0$, see (51). We can even assume that X_j^0 is increasing in j. To see this, if X_j^0 is defined, let X_{j+1}^0 be the closed separable subspace $X^{\{u_{m_{j+1}}\}, \tilde{X}_j^0}$ where $\tilde{X}_j^0$ is a countable and dense set in $X_j^0 \cup X_w^{j+1}, X_1^0 = X_w^{m_1}$. Define $\hat{X} := \text{cl}\cup_j X_j^0$, and note that, for a.e. t, for any $x \in (\cup_j X_j^0) \cap B(x^*(t), \varsigma/16)$, for j large enough, $g(t, x, u_{m_j}(t)) \subset X_j^0 \subset \hat{X}$. Restricting $\dot{p}_{m_j}(t)$ to the separable space $\hat{X}$, a subsequence $\dot{p}_{m_{j_i}}(.)$ of $\dot{p}_{m_j}(.)$ is weakly* convergent to $\dot{p}(.)$ (restricted to $\hat{X}$) in $L_2^s(J, \hat{X}^*)$. We also assume that $p_{m_{j_i}}(T)$, when restricted to $\hat{X}$, is weakly* convergent to $p(T)$. Then, for each t, also $p_{m_{j_i}}(t)$, when restricted to $\hat{X}$, is weakly* convergent to $p(t)$. Now, for any i, n, $i \geq n$, we have $-\langle w, \dot{p}_{m_{j_i}}(t)\rangle \leq \sup_{i\geq n} -\langle w, \dot{p}_{m_{j_i}}(t)\rangle := \alpha_n(t)$, so $\int_J \langle -w, \dot{p}_{m_{j_i}}(t)\rangle \chi(t)dt \leq \int_J \alpha_n(t)\chi(t)dt$, for any $\chi(t) \in L_\infty(J, \mathbb{R}), \chi(t) \geq 0$ a.e. The next to last inequality also holds in the limit $\dot{p}(.)$, which entails $-\langle w, \dot{p}(t)\rangle \leq \alpha_n(t)$ a.e. For any t not in the null set $C = \cap_k C_k, C_k = \cup_{i\geq k}\{t : u_{m_{j_i}} \neq u^*(t)\}$, we have that $t \notin C_{k_t}$ for some k_t. Thus, for $i \geq k_t$, $u^*(t) = u_{m_{j_i}}(t)$ and $d_x^{\beta(t), X^0_{j_i}}(g(t, x_{m_{j_i}}(t), u_{m_{j_i}}(t)), p_{m_{j_i}}(t))(w) = d_x^{\beta(t), X^0_{j_i}}(g(t, x_{m_{j_i}}(t), u^*(t)), p_{m_{j_i}}(t))(w)$. Hence, for a.e. t, we have that

$$-\langle w, \dot{p}(t)\rangle \leq \lim_n \alpha_n(t) \leq \limsup_i d_x^{\beta(t), X^0_{j_i}}(g(t, x_{m_{j_i}}(t), u^*(t)), p_{m_{j_i}}(t))(w) \leq$$
$$\limsup_i d_x^{\beta(t), \hat{X}}(g(t, x_{m_{j_i}}(t), u^*(t)), p_{m_{j_i}}(t))(w) \leq d^{\hat{X}}(t, p(t))(w) \leq d^X(t, p(t))(w).$$

The second inequality follows from (51) and $X^0_{j_i} \supset X_w^{m_{j_i}}$ and the fourth one from upper semicontinuity (Lemma 2) and the implication: For a.e. $t \notin C_{k_t}$, $x \in (\cup_j X^0_j) \cap B(x^*(t), \varsigma/16) \Rightarrow g(t, x, u^*(t)) \subset \hat{X}$. □

Proof of Remark 2 *Proof of* $-\dot{p}(t) \in \tilde{\partial}_x^{\beta(t)}(g(t, x^*(t), u^*(t)), p(t))$ a.e **.** It is claimed that this property holds when X is reflexive. Evidently, $\delta^{-1}\int_{t-\delta}^t \langle -w, \dot{p}(s)\rangle ds \leq \delta^{-1}\int_{t-\delta}^t d^X(s, p(s))(w)ds$. When X is reflexive, $\dot{p}(.)$ can be assumed to belong to $L_2(J, X^*)$, hence $\dot{p}(.)$ has a Lebesgue point a.e. Let $t > 0$ be such a point. Then, for any w,

$$-\langle w, \dot{p}(t)\rangle \leq \liminf_{\delta\searrow 0} \delta^{-1}\int_{t-\delta}^t d^X(s, p(s))(w)ds.$$

Proof of Remark 3 A proof of the assertion in Remark 3 is obtained by choosing a convergent subsequence $(y^0_{m_j}, y'_{m_j})$ of (y^0_m, y'_m) in the proof of Theorem 3, where y'_m is the restriction of $\check{p}^T_m$ to Y', and considering the two cases: $\lim_j(y^0_{m_j}, y'_{m_j}) = 0$, $\lim_j(y^0_{m_j}, y'_{m_j}) \neq 0$, (norm limits). In the former case, the existence of y again yields a nonvanishing limit $\check{p}^T$, in the latter case, trivially, $(y^0, \check{p}^T) \neq 0$.

Proof of Remark 5 For any given w, we have shown that there exists a closed separable space $\hat{X} = X_w$ containing w and $\dot{x}^*(t)$ for a.e. t, which, for a.e. t, is g, u^*-invariant, and for which (46) holds. The lemma below shows that $d^{\hat{X}}(t, p(t))(w) = d^{0,\hat{X}}\langle g(t, x^*(t), u^*(t)), p(t)\rangle(w) \leq$
$d^{0,X}\langle g(t, x^*(t), u^*(t)), p(t)\rangle(w)$, so the assertion in Remark 5 follows.

Lemma 3 Let X be separable. Let h be Lipschitz continuous in $B(x, \xi)$

of rank K, with values in X. For any $w \in X$, assume that for any sequence x_k converging to x, and for any sequence $\lambda_k \searrow 0$, the sequence $\Delta h(x_k)(w, \lambda_k) := [h(x_k + \lambda_k w) - h(x_k)]/\lambda_k$ contains a norm-convergent subsequence. Then, for any $\beta > 0$, $d^\beta(h, x, x^*)(w) = d^0\langle h(x), x^*\rangle(w)$.

Proof: Given $w \in X$. For some sequence (x_k, x_k^*, λ_k) converging to $(x, x^*, 0)$, $(x_k \to x, x_k^* \to^* x^*)$, $x_k^* \in$cl$B(0, \beta)$, we have $d^\beta(h, x, x^*)(w) = \lim_{k\to\infty}\langle \Delta h(x_k)(w, \lambda_k), x_k^*\rangle$. We may assume that $\Delta h(x_k)(w, \lambda_k)$ is norm-convergent, with limit x'. Then $\langle \Delta h(x_k)(w, \lambda_k), \hat{x}^*\rangle - \langle x', \hat{x}^*\rangle$ is small, uniformly in $\hat{x}^* \in$cl$B(0, \beta)$, so $d^0\langle h(x), x^*\rangle(w) \geq \lim_k\langle \Delta h(x_k)(w, \lambda_k), x^*\rangle = \langle x', x^*\rangle = \lim_k\langle x', x_k^*\rangle = \lim_k\langle \Delta h(x_k)(w, \lambda_k), x_k^*\rangle$.□

Proof of Remark 7 We shall merely indicate proofs of the assertions in Remark 7. The first assertion, (i.e. $M^{u(.)}(.), M^*(.) \in L_1(J, \mathbb{R})$), follows from this idea: Let $\gamma_n(t) = \gamma(t)1_{\{t:n\leq\gamma(t)<n+1\}}(t)$. Then each $\gamma_n(.)\dot{p}^{U',V}(.)$ represents a continuous linear functional on $L_2(J, X^{U',V})$, with extension $\gamma_n(.)\dot{p}_{U',V}(.)$ to $L_2(J, X)$. The corresponding cluster point is denoted $\dot{p}_n(.)$, it vanishes outside $\{t : n \leq \gamma(t) < n+1\}$ and satisfies (42), and so $|\dot{p}_n(.)| \leq \gamma(.)$. Evidently, $\dot{p}(t) = \sum_n \dot{p}_n(t)$ (a L_1^s-limit) also satisfies the last inequality, and in fact is a weak cluster point of the generalized sequence $\sum_n \gamma_n(.)\dot{p}_{U',V}(.)$ for the duality $L_\infty(J, X), L_1^s(J, X^*)$. In the proof of (43), as well as in the proof of Theorem 3, the terms weak* limits and weak* cluster points of functions in $L_2(J, X^*)$ used can be replaced by corresponding weak terms referring to the duality $L_\infty(J, X), L_1^s(J, X^*)$.

The second assertion follows from the fact that it suffices to prove the necessary conditions for $u(.) \in \mathcal{U}_\varepsilon := \{u(.) \in \mathcal{U} : \int_0^T M^*(s)1_{\{s:u(s)\neq u^*(s)\}}(s)ds \leq \varepsilon\}$, for any given $\varepsilon > 0$ so small that $\mathcal{U}_\varepsilon \subset D_T$. But then, in fact, $|\nabla g^k(t, x^{u_k}(t), u_k(t))| \leq M^{**}(s) := M^*(s)1_{\{s:u_k(s)\neq u^*(s)\}}(s) + M^{u^*(.)}(s)1_{\{s:u_k(s)=u^*(s)\}}(s)$, so from (32) and (33), $|p_k(t)| \leq M_\phi \exp(\int_t^T M^{**}(s)ds) \leq M_\phi \exp(\varepsilon + \int_t^T M^{u^*(.)}(s)ds)$. Hence, (11) holds for $\beta(t)$ replaced by $(M_\phi + M_G)\exp(\varepsilon + \int_t^T M^{u^*(.)}(s)ds)$, (with $M_G = 0$, in the free end case). Since ε is arbitrary, (11) holds even for $\varepsilon = 0$. I.e., (11) holds for $\beta(t) = (M_\phi + M_G)\exp(\int_t^T M^{u^*(.)}(s)ds)$.

The third assertion follows from the fact that for any finite set $U' \in \mathcal{U}'$, functions $M^*(.)$ and $M_*(.)$ exist in $L_1(J, \mathbb{R})$, bounding the functions $M^{u(.)}(.)$ and $M_{u(.)}(.)$ for all $u(.) \in U' \cup \hat{\mathcal{U}}$. Thus, Theorem 3 holds for $\mathcal{U}$ replaced by the switching- and $\sigma-$closure U^* of $U' \cup \hat{\mathcal{U}}$, for $p(.) := p_{U'}(.)$, for $\beta(t)$ as just defined. A cluster point $p(.)$ of the generalized sequence $p_{U'}(.)$ (the sets U' $\in \mathcal{U}'$ directed by inclusion), then yields that Theorem 3 also holds for $\mathcal{U}$, (for

$\beta(t)$ as just defined). □

A few proofs of assertions contained in the Remarks, can be found in Appendix.

Appendix

Lemma A If $h : X \to Y$ is Lipschitz continuous in $B(x', \delta')$ and has a Gâteaux derivative at x' when restricted to a dense linear subspace X' in X containing x', then h has a (bounded) Gâteaux derivative at x' with respect to X.

Proof: Let K be the Lipschitz rank of h. Then, $|\nabla h(x')[x]| \le K|x|$, for all $x \in X'$, so, evidently, $\nabla h(x')[x]$ has an extension to all X. Next, let $w \in X$, and choose $w' \in X'$ such that $|w - w'| \le \varepsilon/3K$. Next choose δ so small that $|\lambda^{-1}\{h(x' + \lambda w') - h(x')\} - \nabla h(x')w'| \le \varepsilon/3K$ when $\lambda \in (0, \delta)$. Now, $|\lambda^{-1}\{h(x' + \lambda w') - h(x')\} - \lambda^{-1}\{h(x' + \lambda w) - h(x')\}| \le K|w' - w| \le \varepsilon/3$ and $|\nabla h(x')w' - \nabla h(x')w| \le \varepsilon/3$. Hence, $|\lambda^{-1}\{h(x' + \lambda w) - h(x')\} - \nabla h(x')w| \le \varepsilon, \lambda \in (0, \delta)$.

Let $g : J \times X \to X$ be separately measurable in t, and Lipschitz continuous in $B(x^*(t), \varsigma)$ of rank $\kappa(t)$, ($\kappa(t)$ integrable), where $x^*(t)$ is an antidifferentiable function satisfying $\dot{x}^*(t) = g(t, x^*(t))$ a.e., $x^*(0) = v^*$. Assume that the Gâteaux derivative $\nabla_2 g(t, x^*(t))$ exists for all t. Consider the equation

$$dx/dt = g(t, x(t)),\ x(0) = v, \tag{52}$$

and the corresponding variational equation

$$dq/dt = \nabla_2 g(t, x^*(t))(q(t)), q(0) = v. \tag{53}$$

By standard theory, the unique solution $q(t) := q(t, v)$ to (53) can be written $q(t,v) = C(t, 0)v$, where $C(t, 0)$ is a bounded linear operator with $C(0, 0) = I$, continuous in t in operator norm. In fact, $|C(t, 0)| \le \exp(\int_0^t \kappa(s)ds) \le e^{\kappa^*}$, where $\kappa^* := \int_0^T \kappa(s)ds$.

Lemma B For $\gamma' = \varsigma e^{-\kappa^*}/2$, (52) has a solution $x(t) := x(t, v)$ in $\mathrm{cl}B(x^*(t), \varsigma/2)$ for all $v \in B(v^*, \gamma')$. Moreover, $v \to x(t, v)$ has a bounded linear Gâteaux derivative $\nabla_2 x(t, v^*)$, and $q(t, v) = \nabla_2 x(t, v^*)v$. The resolvent $C(t, 0)$ of (53) equals $\nabla_2 x(t, v^*)$.

Proof Write $x(t) = x^*(t) = x(t, v^*)$. For simplicity, assume $v^* = 0$. By Gronwall's inequality, a local existence and continuation argument, a solution $x(s, v)$ exists and belongs to $clB(x(s), \varsigma/2)$, for $|v| \le \gamma'$. Let $v \in X$ be arbitrary, and, below, let $\lambda \in (0, \gamma'/|v|]$. Define

$$z(v, t, \lambda) := [x(t, \lambda v) - x(t)]/\lambda = \int_0^t \{[g(s, x(s) + \lambda z(v, s, \lambda)) - g(s, x(s))]/\lambda\} ds,$$

the norm of the integrand being $\le \kappa(s)|z(v, s, \lambda)|$. Then, by Gronwall's inequality, $|z(v, t, \lambda)| \le |v|e^{\kappa^*}$, so $|\partial z(v, s, \lambda)/\partial s| \le \kappa(s)|v|e^{\kappa^*}$. Note that $|\partial q(s, v)/\partial s| \le \kappa(s)|q(s, v)|$. By Gronwall's inequality, $|q(t, v)| \le |v|e^{\kappa^*}$, so $|\partial q(s, v)/\partial s| \le \kappa(s)|v|e^{\kappa^*}$.

Let $\varepsilon > 0$ (ε arbitrary), and $\alpha(v, s, \lambda) := |z(v, s, \lambda) - q(s, v)|$. Then $\alpha(v, 0, \lambda) = 0$. Define $\gamma(v, s, \lambda) := |a(s, \lambda) - \nabla_2 g(s, x(s))q(s, v)|$, where $a(s, \lambda) := [g(s, x(s) + \lambda q(s, v)) - g(s, x(s))]/\lambda$. Note that $\lim_{\lambda \searrow 0} \gamma(v, s, \lambda) = 0$. Because $|a(s, \lambda)| \le \kappa(s)|q(s, v)| \le \kappa(s)|v|e^{\kappa^*}$, $|\nabla_2 g(s, x(.))(q(s, v))| \le \kappa(s)|q(s, v)| \le \kappa(s)|v|e^{\kappa^*} \le$ and $\gamma(v, s, \lambda) \le 2\kappa(s)|v|e^{\kappa^*}$, then, by dominated convergence, for some $r > 0$, $\int_J \gamma(v, s, \lambda) ds < \varepsilon$ when $\lambda \le r$. Evidently,

$$|[g(s, x(s) + \lambda z(v, ., \lambda)) - g(s, x(s))]/\lambda - [g(s, x(s) + \lambda q(s, v)) - g(s, x(s))]/\lambda|$$

$\le \kappa(s)\alpha(v, s, \lambda)$.
Now,

$$|[g(s, x(s) + \lambda q(s, v)) - g(s, x(s))]/\lambda - \partial q(s, v)/\partial s| \le \gamma(v, s, \lambda),$$

so

$$|\partial z(v, s, \lambda)/\partial s - \partial q(s, v)/\partial s| = |[g(s, x(s) + \lambda z(v, s, \lambda)) - g(s, x(s))]/\lambda - \partial q(s, v)/\partial s| \le \gamma(v, s, \lambda) + \kappa(s)\alpha(v, s, \lambda),$$

Hence, by Gronwall's inequality, $\alpha(v, s, \lambda) \le \varepsilon \exp(\kappa^*)$, when $\lambda \in (0, r)$. □

Proof of Theorem 1 Let $u(.) \in \mathcal{U}$, and let t be a Lebesgue point (see Lebesgue set in Dunford and Schwartz, (1967)) of both $g(., x(.), u(.))$ and $g(., x^*(.), u^*(.))$, and let $u_\delta(.) = u(.)1_{[t-\delta,t]} + u^*(.)(1 - 1_{[t-\delta,t]})$, $\delta > 0$. For δ small enough, the solution $x_\delta(t)$ of $dx/ds = g(s, x(s), u_\delta(s))$ a.e., $x(0) = x_0$ exists in $\text{cl}B(x^*(t), \varsigma/2)$. Moreover, by Gronwall's inequality, for some constant C, $|x_\delta(s) - x^*(s)| \le C\delta$, (see (28)). Let $a = g(t, x^*(t), u_\delta(t)) - g(t, x^*(t), u^*(t))$. Then, $|x_\delta(t) - x^*(t) - \delta a| =$

$$|\int_{[t-\delta,t]} g(s, x_\delta(s), u_\delta(s)) ds - \int_{[t-\delta,t]} g(s, x^*(s), u^*(s)) ds - \delta a|$$
$$\le |\int_{[t-\delta,t]} g(s, x_\delta(s), u_\delta(s)) ds - \int_{[t-\delta,t]} g(s, x^*(s), u_\delta(s)) ds +$$

$\int_{[t-\delta,t]} g(s,x^*(s),u_\delta(s))ds - \int_{[t-\delta,t]} g(s,x^*(s),u^*(s))ds - \delta a| \leq$
$\int_{[t-\delta,t]} M^{u(.)}(s)C\delta ds + |\int_{[t-\delta,t]} g(s,x^*(s),u_\delta(s))ds - \int_{[t-\delta,t]} g(s,x^*(s),u^*(s))ds - \delta a|$

Using these inequalities and that t is a Lebesgue point, it follows that $\lim \delta^{-1}|x_\delta(t) - x^*(t) - a| = 0$, so $\lim \delta^{-1}(x_\delta(t) - x^*(t)) = a$. By Lemma B, $x_\delta(T) = C(T,t)a$, where $C(T,t)$ is the resolvent of $\dot{q} = g_x(s,x^*(s),u^*(s))q$, $(C(t,t) = I)$.By optimality, at $\delta = 0, d\phi(x_\delta(T))/d\delta \leq 0$, hence $(d\phi(x^*(T))/dx)C(T,t)a \leq 0$. Defining $p(T) = (d\phi(x^*(T))/dx)$ and $p(t) = p(T)C(T,t)$, the maximum condition (7) in Theorem 1 follows. Standard theory gives that $p(t)$ is a weak* solution of the adjoint equation.

Comment on Remark 1 The existence of $\dot{p}(.)$ comes out of the proof of Theorems 2 and 3. However, it is also an easy consequence of absolute continuity and (the stronger property) $|p(t) - p(s)| \leq \int_s^t \gamma(\rho)d\rho, t > s$. Let J' be the Lebesgue set of $\gamma(.)$. For each $w \in X, \langle w, p(t)\rangle$ is absolutely continuous. Then, for all $t \in J'$ not in a null set N_w, there exists a real number $\check{p}_w(t)$ such that $(d/dt)\langle w, p(t)\rangle = \check{p}_w(t)$, where $|\check{p}_w(t)| \leq \gamma(t)|w|$. Assume that $t \in J'$ also belongs to $J \backslash N_{w'}$. Then, $(d/dt)\langle(\alpha w + \beta w'), p(t)\rangle$ $=\alpha\check{p}_w(t) + \beta\check{p}_{w'}(t)$. In fact, $w \to \check{p}_w(t)$ is linear on the set (in fact linear subspace) W_t for which $(d/dt)\langle w, p(t)\rangle$ exists, $t \in J'$. By the Hahn-Banach theorem, for each $t \in J', \check{p}_w(t)$ has an extension to all X, denoted $\check{p}(t)$, satisfying $|\check{p}(t)| \leq \gamma(t)$. Trivially, $(d/dt)\langle w, p(t)\rangle = \langle w, \check{p}(t)\rangle$ is measurable, and $\int_0^s \langle w, \check{p}(t)\rangle = \int_0^s (d/dt)\langle w, p(t)\rangle = \langle w, p(t) - p(0)\rangle$. The last equality determines $\check{p}(t)$ uniquely in the sense that the continuous linear functional it represents is unique: If, for all s, $\int_0^s \langle \hat{x}, \check{p}(t)\rangle dt = \int_0^s \langle \hat{x}, \hat{p}(t)\rangle dt$, $\check{p}(.), \hat{p}(.) \in L_2^s(J, X^*)$, then for any countable set $X' \subset X$, for some set $J_{X'}$ of full measure, for all $t \in J_{X'}, \hat{x} \in X', \langle \hat{x}, \check{p}(t)\rangle = \langle \hat{x}, \hat{p}(t)\rangle$. This also holds for $\hat{x} \in clX'$. Now, any $x(.) \in L_2(J, X)$ can be assumed to take values in a set of the form $\text{cl}X'$, so $\langle x(t), \check{p}(t)\rangle = \langle \hat{x}(t), \hat{p}(t)\rangle$, for $t \in J_{X'}$, which shows the uniqueness claimed.

Proof of $y \notin C \Rightarrow 0 \notin \partial\text{dist}(y, C)$. Choose $c \in C$, such that $|y - c| \leq 3\text{dist}(y, C)/2$. Write $y_\lambda = \lambda y + (1 - \lambda)c$. Evidently, $\text{dist}(y_{1/2}, C) \leq |y_{1/2} - c| = |y - c|/2 \leq 3\text{dist}(y, C)/4$. Since $\lambda \to \text{dist}(y_\lambda, C) =: \alpha(\lambda)$ is convex, $d^0\alpha(y_1)(-1) = d'\alpha(y_1)(-1) \leq \{\alpha(1/2) - \alpha(1)\}/(1/2) \leq \{3\text{dist}(y, C)/4 - \alpha(1)\}/(1/2) = -\text{dist}(y, C)/2 < 0$, ($y \to \text{dist}(y, C)$ is convex, $d' = d^0$, see Clarke 1983 p. 53, p. 40). Hence, $[d^0\text{dist}(x, C)]_{x=y}(c - y) < 0$, so $0 \notin [\partial\text{dist}(x, C)]_{x=y}$.

Proof of measurability of $d_x^{\beta(t)}(g(t,x^*(t),u^*(t)),p(t))(w)$. This function is surely measurable if g is simultaneously continuous in (t,u) and X is separable. To see this, it suffices to note that by Lusin's theorem, there exists an increasing squence of closed sets J_k, $\mathrm{meas}(J\backslash J_k) < 1/k$, such that $u^*(t)$ is continuous on J_k. Then, for all $x \in B(0,\varsigma)$, $g(t,x^*(t)+x,u^*(t))$ is continuous on J_k, and for any $\delta > 0$ and weak* neigbourhood W in $\mathrm{cl}B(0,\beta(t))$, $\Theta(\delta,W) := \sup_{x\in B(0,\delta),x^*\in W,\lambda\in(0,\delta)}\langle\lambda^{-1}[g(t,x^*(t)+x+\lambda w,u^*(t)) - g(t,x^*(t)+x,u^*(t)),x^*)],x^*\rangle$ is lower semicontinuous on each J_k. The weak* topology on $\mathrm{cl}B(0,\beta(t))$ is metric when X is separable, so when taking $\lim\Theta(\delta,W)$, we can confine ourselves to a sequence, δ_n,W_n.

The function $d_x^{\beta(t)}(g(t,x^*(t),u^*(t)),p(t)\rangle)(w)$ is convex in w, (the proof is similar to the proof of convexity of $d_x^0\langle g(t,x^*(t),u^*(t)),p(t)\rangle(w)$, see Clarke (1983), p.26.).

Proof of the last assertion in Remark 3. If $Y = Y'+Y''$, $Y'\cap Y'' = \{0\}$, and if for some $z \in Y$, some $\varepsilon \in (0,1], \Pi''B(z,\varepsilon) \subset \Pi''[(C - G(x^*(T)))\cap B(0,1)]$, chooce $\varepsilon' > 0$ so small that $\varepsilon' \leq \langle\varepsilon/16K'\rangle$, where $K' \in [1,\infty)$ is greater or equal to the operator norm on Π', the projection onto Y'. Let $y \in clB(0,1)$ be such that $1/2 \leq \langle y,y^*\rangle$. Then, by the additional condition in Remark 3, $|\langle\Pi'\hat{y},y^*\rangle| \leq \varepsilon'|\Pi'\hat{y}| \leq K\varepsilon' \leq 1/4$, for all $\hat{y} \in clB(0,1) \subset Y$. Thus, $1/2 \leq \langle y,y^*\rangle = \langle\Pi'y,y^*\rangle + \langle\Pi''y,y^*\rangle$, so $1/4 \leq \langle\Pi''y,y^*\rangle$. Now, by (16) $0 \geq \langle(C - G(x^*(T)))\cap B(0,1),y^*\rangle = \langle(\Pi'+\Pi'')[(C - G(x^*(T)))\cap B(0,1)],y^*\rangle$, so $K'\varepsilon' \geq \langle\Pi''[(C - G(x^*(T)))\cap B(0,1)],y^*\rangle \geq \langle\Pi''B(z,\varepsilon),y^*\rangle$. Now, $z+\varepsilon y/2 \in B(z,\varepsilon)$, so $\varepsilon/16 \geq K'\varepsilon' \geq \langle\Pi''(z+\varepsilon y/2),y^*\rangle \geq \langle\Pi''z,y^*\rangle + \langle\Pi''\varepsilon y/2,y^*\rangle$, or $-\varepsilon/16 \geq \varepsilon/16 - \langle\Pi''\varepsilon y/2,y^*\rangle \geq \langle\Pi''z,y^*\rangle$. Hence, $\varepsilon' \leq \langle -\Pi''z,y^*\rangle$.

Proof of Remark 6 Assume $M_{u(.)}(t) \equiv M_{u^*(.)} \equiv M_*$, $M^{u(.)} \equiv M^{u^*(.)}(.) \equiv M^*$. Let us prove that the property $B(z,\varepsilon) \subset cl\{G'(x^*(T))q_u(T) - c + G(x^*(T)) : u \in \mathcal{U}, c \in C\}$ implies (17): Note that for $\varepsilon' \in (0,\varepsilon/64(1+M^{**})]$, where $M^{**} := 2M_*\exp(M^*T)$, ε' small enough, for any given set M with $\mathrm{meas}(M) < \varepsilon'$, $B(z,\varepsilon/2) \subset \mathrm{cl}\{G'(x^*(T))q_u(T) - c + G(x^*(T)) : c \in C, u \in \mathcal{U}, u = u^* \text{ on } M\}$. Moreover, $B(z,\varepsilon/4) \subset \mathrm{cl}\{G'(x^*(T))q_u(T) - c + G(x^*(T)) : c \in C, u \in \mathcal{U}, u = \hat{u} \text{ on } M\}$ even if q_u is redefined to satisfy $dq_u/dt = g(t,x^*(t),u(t)) - g(t,x^*(t),\hat{u}(t)) + g'_x(t,x^*(t),\hat{u}(t)), q(0) = 0$, when $\{s : \hat{u}(s) \neq u^*(s)\} \subset M$, when ε' is small enough. Finally, for any continuous function $x(.)$ close enough to $x^*(.)$, $B(z,\varepsilon/8) \subset \mathrm{cl}\{G'(x(T))q_u(T) - c + G(x(T)) : c \in C, u \in \mathcal{U}, u = \hat{u} \text{ on } M\}$ even if q_u is redefined to satisfy $dq_u/dt = g(t,x(t),u(t)) - g(t,x(t),\hat{u}(t)) + g_x(t,x(t),\hat{u}(t))q, q(0) = 0$.

Thus, the last inclusion holds for ε' small enough, for $(x(.), \hat{u}(.))$ being a system pair, with $\sigma(\hat{u}, u^*) < \varepsilon'$. We then write $q_u(t) = \hat{q}_u(t)$. Given any quintuple $(x(.), \hat{u}(.), p(.), \hat{p}, y^*)$ satisfying (13)-(16). Now, (14) is equivalent to $\langle \hat{q}_u(T), p(T)\rangle \leq \varepsilon'$. By (16), $[G'(x(T))]^* y^* = p(T) - \hat{p}$, so, $\langle \hat{q}_u(T), [G'(x(T))]^* y^*\rangle - \langle \hat{q}_u(T), p(T)\rangle \leq \varepsilon' |\hat{q}_u(T)|$, by $|\hat{p}| \leq \varepsilon'$ in (16). Thus, $\langle \hat{q}_u(T), [G'(x(T))]^* y^*\rangle \leq \langle \hat{q}_u(T), [G'(x(T))]^* y^*\rangle - \langle \hat{q}_u(T), p(T)\rangle$
$+\langle \hat{q}_u(T), p(T)\rangle \leq \varepsilon'(1 + M^{**})$. Since, by (16), $\langle -C + G(x(T)), y^*\rangle \leq 0$, then $\langle G'(x(T))\hat{q}_u(T) - C + G(x(T)), y^*\rangle \leq \varepsilon'(1 + M^{**})$, so $\langle B(z, \varepsilon/8), y^*\rangle \leq \varepsilon'(1 + M^{**})$. As $\langle y', y^*\rangle \geq 1/2$ for some $y' \in \mathrm{cl}B(0,1)$ and $z + \varepsilon y'/16 \in B(z, \varepsilon/8)$, then $\langle z + \varepsilon y'/16, y^*\rangle \leq \varepsilon'(1 + M^{**})$. Hence, $\langle z, y^*\rangle \leq -\langle \varepsilon y'/16, y^*\rangle + \varepsilon'(1 + M^{**}) \leq -\varepsilon/32 + \varepsilon'(1 + M^{**}) \leq -\varepsilon/64$. I.e., $\langle -z, y^*\rangle \geq \varepsilon'$.

References

Ahmed, N.U. and K.L.Teo, (1981), *Optimal Control of Distributed Parameter Systems*, North Holland, New York.

Barbu, V., (1984), *Optimal Control of Variational Inequalities*, Pitman, Boston.

Barbu,V., (1993), *Analysis and Control of Nonlinear Infinite Dimensional Systems*, Academic Press, Boston .

Basile, N. and M. Ninimmi, (1992), "Proximal normal analysis applied to optimal control problems in infinite dimensional spaces", *J. Optim. Theory Appl.* 73, 121-147.

Clarke, F.H., (1983), *Optimization and Nonsmooth Analysis*, J. Wiley, New York.

Corduneanu,C., (1990), "Some control problems for abstract Volterra functional-differential equations", *Prog. Syst. Control Theory 5*, 331-338.

Dunford, N. and J.T. Schwarz (1967), *Linear Operators, Part I,* Interscience, New York.

Fattorini, H.O., (1987), "A unified theory of necessary conditions for nonlinear, nonconvex control systems", *Applied Math. Optim.*, 15, 141-185.

Fattorini, H.O., (1993), "Optimal control problems for distributed parameter systems in Banach spaces", *Applied Math.Optim.* 28, 225-257.

Fattorini, H.O. and H. Frankowska, (1991), "Necessary conditions for infinite dimensional control problems", *Math. Control, Signals and Systems*, 4, 41-67.

Fattorini, H.O,. (1999), *Infinite Dimensional Optimization and Control Theory*, Cambridge University Press, Cambridge, England.

Ledzewicz, U. and A. Nowakowski, (1997), "Optimality conditions for problems governed by abstract semilinear differential equations in complex

Banach spaces", *J. of Applied Analysis* 3, 67-91.

Kerbal, S., N.U. Ahmed, (1996), "Necessary conditions of optimality for systems governed by B-evolutions", in Ladde, G.S.(ed.) et al., *Dynamic systems and applications, Vol.2, Proceedings of the 2nd international conference*, Morehouse College, Atlanta, GA, USA, May 24-27, (1995), Dynamic Publishers 2030-300, Atlanta, GA,USA.

Li, Xunjing, (1996), "Optimality conditions of infinite dimensional systems", in Lakshmikantham,V.(ed.) *World congress of nonlinear analysis '92. Proceedings from the first world congress*, Tampa, FL, USA, August 19-26,(1992), de Gruyter,Berlin 25492556.

Li, Xunjing and Jiongmin Yong, (1991), "Necessary conditions for optimal control of distributed parameter systems", *SIAM J.Control Optimization* 29, 895-908.

Mansimov, K.B., and M.G. Mustafaev, (1985), "Some necessary optimality conditions in a control problem described by a system of Volterra type integral equations", *Izv.Akad.Nauk Az. SSR, Ser. Fiz,-Tekh. Mat. Nauk* 5, 118-122.

Pallu de la Barriere, R (1980), *Optimal Control Theory*, Dover Publications, New York.

Raymond, J. and H. Zidani, (1999), "Hamiltonian Pontryagin's principles for control problems governed by Semilinear Parabolic Equations", *Appl. Math. and Optim.* 39, 143-177.

Seierstad, A. (1975), "An extension to Banach space of Pontryagin's maximum principle", *J. of Optim. Theory and Appl.,* 17, 293-335.

Seierstad, A. (1996), "Constrained retarded nonlinear optimal control problems in Banach state space", *Appl. Math Optim.* 34, 191-230.

Seierstad, A. (1997), "Nonsmooth control problems in Banach state space", *Optimization*, 41, 303-319.

Tiba, D., (1990), *Optimal control of nonsmooth distributed parameter systems*, Lecture Notes in Mathematics 1459, Springer Verlag, Berlin.

Sumin,V.I. (1989), "Volterra functional-operator equations in the theory of optimal control of distributed systems", *Sov. Math.,Dokl.* 39, 374-378.

Yong, Jiongmin (1990a), "Optimal controls for distributed parameter systems with mixed constraints", *Colloq. Math.* 60/61, 35-48.

Yong, Jiongmin (1990b), "Maximum principle of optimal controls for a nonsmooth semilinear evolution system"*. Analysis and optimization of systems, Proc. 9th Int Conf., Antibes/Fr, Lect. Notes Control Inf. Sci.* 144, 559-569.