MEMORANDUM

No 23/99

Necessary conditions involving generalized directional derivatives for optimal control of nonsmooth retarded Volterra integral equations in Banach 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.sv.uio.no/sosoek/
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 22	By Karl Ove Moene and Michael Wallerstein: Inequality, Social Insurance and Redistribution. 40 p.
No 21	By Kai Leitemo: Inflation Targeting Strategies in small open Economies. 44 p.
No 20	By Erik Biørn: Estimating Regression Systems from unbalanced Panel Data: A Stepwise Maximum Likelihood Procedure. 18 p.
No 19	By Knut Røed and Steinar Strøm: Progresssive Taxes and the Labour Market – Is the Trade-Off between Equality and Efficiency Inevitable? 39p.
No 18	By Erik Hernæs: Fewer in Number but Harder to Employ: Incidence and Duration of Unemployment in an Economic Upswing. 28 p.
No 17	By Erik Hernæs, Marte Sollie and Steinar Strøm: Early retirement and Economic Incentives. 27 p.
No 16	By Oddbjørn Raaum and Tom Erik Aabø: The effect of schooling on earnings: The role of family background studied by a large sample of Norwegian twins. 31 p.
No 15	By T. Hægeland, T.J. Klette and K.G. Salvanes: Declining returns to education in Norway? Comparing estimates across chorts, sectors and overtime. 21 p.
No 14	By Knut Røed and Tao Zhang:Unemployment Duration in a Non-Stationary Macroeconomic Environment. 32 p.
No 13	By Tor Jakob Klette, Jarle Møen and Zvi Griliches: Do subsidies to commercial R&D reduce market failures? Microeconometric evaluation studies. 36 p.

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

Necessary conditions involving generalized directional derivatives for optimal control of nonsmooth retarded Volterra integral equations in Banach space.

by
Atle Seierstad, University of Oslo.

Abstract. Necessary conditions for the optimal control of solutions to nonsmooth nonlinear retarded differential equations and Volterra integral equations in Banach state space are proved. The results also apply to problems of control of mild solutions to abstract weakly nonlinear evolution equations. Terminal condition are imposed, and certain linearized controllability properties are required for the necessary conditions to hold. A special type of generalized directional derivatives are used in the proof and to express the necessary conditions.

Content

1. Introduction

The purpose of this paper is to give necessary conditions for the optimal control of continuous solutions to nonsmooth nonlinear retarded differential and integral equations in Banach state space. The theory below is general enough to include problems involving Volterra retarded equations of the form

$$x(t) = x_0(t) + \int_0^t g(t, s, x(\to s), u(s))ds, t \in [0, T], \tag{1}$$

the symbol $x(\to s)$ denoting dependence on the values $x(.)$ takes on $[0, s]$, ("retarded dependence"). The constraints in the problem are:

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

The continuous functions $x_0(t)$ and $x(t)$ take values in a Banach space X, the measurable controls $u(t)$ take values in some given topological space U. Furthermore, G is a given function from X into a Banach space Y, C is a fixed subset in Y, g is a fixed function with values in X, $x_0(.)$ is fixed, and T is a fixed number > 0.

The criterion to be maximized is

$$\phi(x(T)), \phi \text{ a given function from } X \text{ into } \mathbb{R}. \tag{3}$$

The necessary conditions will be expressed by means of a certain type of generalized directional (semi-) derivatives. The results in particular apply to problem where g, (see (1)), has ordinary directional (semi-) derivatives with respect to $x(.)$. The results apply both to problems of control of ordinary differential equations, and to problems of control of mild solutions to abstract weakly nonlinear evolution equations of the form $dx/dt = Ax + h(t, x(t), u(t))$, where A is a generator of a strongly continuous semigroup in a Banach space, (A a closed, linear operator). Such mild solutions can be expressed by an equation of the form (1), see e.g. Fattorini (1999). A selection of recent references that are mainly concerned with applications to partial differential equations are included. For the most part, these works discuss abstract semilinear (weakly nonlinear) evolution equations, some of which are functional ones. Nonsmooth problems are treated in Yong (1990a,1990b). Fattorini (1987) also considers other systems within a general framework. The present paper is based on certain abstract results from nonsmooth analysis. These results can be compared to those used in Fattorini and Frankowska (1991) and Fattorini (1993). The abstract results used in the present paper require similar, but somewhat more demanding "linearized" controllability properties, but do not require stability of the set of variations with respect to perturbations of the optimal control for useful necessary conditions to be formulated. For references to finite dimensional results concerned with ordinary retarded equations, see Clarke and Wolenski (1996) and Neustadt (1976). The latter one also contains results on Volterra equations. Volterra equations are also treated in Burnap and Kazemi (1999), Sumin (1989), Mansimov and Mustafaev (1985), Corduneanu (1990).

2. Notation and terminology

A number $T > 0$ is given. Let $J := [0, T]$, let X and Y be real normed spaces, and let $C(J, X)$ be the space of continuous functions from J into X, furnished with the sup-norm. Let clA be the norm closure of A, (A a subset of Y) and let coA be the convex hull of A. Let E be a locally Lipschitz continuous function on an open subset U' of X into Y. A directional derivative container (abbreviated d.d. container) $DE(x_0)(v)$ at $x_0 \in U'$, in direction $v \in X$, ($v \neq 0$) is a set in Y such that for all $\varepsilon > 0$, an $r > 0$ exists, such that $\Delta^*E(x_0)(v, r) := \{(E(x_0 + \lambda v) - E(x_0))/\lambda : \lambda \in (0, r]\}$ is contained in $DE(x_0)(v) + B(0, \varepsilon)$. If r can be chosen such that this inclusion holds for all v in a given set Q, the d.d. container is said to be uniform in $v \in Q$. The set $D^*E(x_0)(v) := \cap_{r>0} cl\Delta^*E(x_0)(v, r)$ is called the contingent derivative in direction v, (it has another definition in the non-Lipschitzian case). The entity $\Delta E(x_0)(v, \lambda) := (E(x_0 + \lambda v) - E(x_0))/\lambda$ is called a difference quotient at x_0 in direction v. For any set Q, if $DE(x_0)(v)$ is defined for all $v \in Q$, then $DE(x_0)(Q) := \cup_{v \in Q} DE(x_0)(v)$. The d.d. containers occuring in this paper, except in Section 8, are, for each v, assumed to be bounded sets, and, as functions of v, (positively) linearly homogeneous (i.e. $DE(x_0)(\lambda v) = \lambda DE(x_0)(v), \lambda \geq 0$; it is assumed that if $DE(x_0)(v)$ is defined for a given v, it is also defined for any vector $\lambda v, \lambda \geq 0$). If $DE(x_0)(v)$ reduces to a single point, it means that a directional (semi-) derivative exists in direction v. (See further comments and results on d.d. containers in Section 8 and in Appendix B.)

A family F of functions $f : J \times J \times C(J, X) \to X$, is closed (convex) under switching, if for each measurable set $M \subset J$, and $f, f' \in F$, we have $f(t, s, x(.))1_M(s) + f'(t, s, x(.))(1 - 1_M(s)) \in F$, ($1_M$ is the indicator function of M). Such families F are furnished with the pseudometric $\sigma(f, g) = \inf\{\text{meas}(M) : M \text{ is a measurable set containing } \{s : f(t, s, x(.)) \neq g(t, s, x(.))\} \text{ for all } (t, x(.))\}$. Throughout this paper, $B(f, r)$ and $B(f^*, r)$, $f, f^* \in F$, are σ-balls in F.

Vector valued functions are called measurable and integrable in the sense of Dunford and Schwartz (1967), often called strong or Bochner measurability (integrability) in the literature. A function $x(.) : J \to X$ is antidifferentiable if it is absolutely continuous and has a derivative a.e. which is integrable. A set-valued function $A(t)$ from J into the set of all subsets of Y is measurable if $\{t : A(t) \cap \check{U} \neq \varnothing\}$ is measurable for each open set $\check{U}$.

3. The control system

Let X be a real Banach space, $x_0(.)$ a given element in $C(J, X)$, G a function from X into a real Banach space Y, and ϕ a real-valued function on X. Let F be a family of functions $f(t, s, x(.))$ from $J \times J \times C(J, X)$ into X, exhibiting retarded dependence on $x(.)$, (precisely defined in (6) below). This family may be generated by a family of control functions $u(.)$, see Remark 1 below. The following control problem will be discussed. The state equation is a Volterra retarded integral equation of the form:

$$x(t) = x_0(t) + \textstyle\int_0^t f(t, s, x(\to s))ds, t \in [0, T], \tag{4}$$

where the symbol $x(\to s)$ indicates the retarded dependence.

The maximization problem is

$$\max_{x(.), f} \phi(x(T)), \text{ subject to (4)}, G(x(T)) \in C \text{ and } f \in F, \tag{5}$$

where $x_0(.), \phi, T, G, F$ and C, (a closed convex set in Y) are given entities.

A "system pair" means a pair $(x(.), f)$, such that $f \in F$ and $x(.)$ is a continuous function satisfying (4). If $x(.)$ is unique, for a given f, we often write $x(.) = x^f(.)$, (below, conditions will be imposed securing uniqueness for f's close to the optimal f^*). If the system pair also satisfies $G(x(T)) \in C$, it is called admissible. The control problem (5) amounts to maximizing $\phi(x(T))$ in the class of admissible pairs $(x(.), f)$.

Remark 1. The standard example of a family of functions f of the type above is $F := \{g(t, s, x(\to s), u(s)) : u(s) \in U$ for all s, $u(.)$ strongly measurable$\}$, where U is a given topological space and $g : J \times J \times C(J, X) \times U \to X$ is a fixed function, separately, continuous in t, measurable in s, continuous in $x(.)$, and continuous in u. (Strong measurability of $u(.)$ is taken to mean that $u(.)$ is a limit a.e. of a sequence of measurable step functions.) See a discussion of a similar family arising in the control of ordinary differential equations in Seierstad (1975). Note that it can be proved that F becomes closed under switching and essentially σ-closed, because $\{u(.) : u(t) \in U, u(.)$ measurable$\}$ has both properties. The values of f for $t < s$ are "irrelevant" as far as the Volterra equation (4) is concerned, and for families F generated by controls $u(.)$, if the f 's do not satisfy $f(t, s, x) = f(s, s, x), t < s$, we may redefine the f 's so that this property holds. No further comment on this particular example will be given, except one in Remark 4. $\square$

It is assumed that

F is essentially closed in the pseudometric σ, and each $f \in F$ is, separately, measurable in s, and continuous in t for each $x(.) \in C(J, X)$ and satisfies $f(t, s, x(.)) = f(t, s, \check{x}(.))$ if $x(\tau) = \check{x}(\tau)$ for $\tau \geq s$, $x(.), \check{x}(.) \in C(J, X)$. (6)

The last property in (6), retarded dependence, means that, whatever t is, $f(t, s, x(.))$ depends only on "past values" of $x(.)$, i.e. on the values $x(s')$, for $s' \leq s$.

Assume that there exists a system pair $(x^*(.), f^*)$, such that for all $f \in F$, there exist positive numbers M, M^f, and ς, ς and M independent of f, such that, for all f and all t, s,

$\sup_{x(.)\in B(x^*(.),\varsigma)} |f(t, s, x(\to s))| \leq M$, and $x(.) \to f(t, s, x(\to s))$ is Lipschitz continuous of rank $\leq M^f$ in $B(x^*(.), \varsigma)$, (7)

where $B(x^*(.), \varsigma)$ is a ball around $x^*(.)$ in $C(J, X)$ of radius ς.

The existence of $(x^*(.), f^*)$ and conditions (6)-(7) are assumed throughout this paper. Frequent use is also made of one or more of the conditions (8)-(11):

$G(x) : X \to Y$ and $\phi(x) : X \to \mathbb{R}$ are Lipschitz continuous in $B(x^*(T), \varsigma)$ with ranks M_G and M_ϕ, respectively. (8)

F is closed under switching. For any $x(.)$ in $C(J, X)$ the function $s \to f(., s, x(.)) : J \to C(J, X)$ is measurable. (9)

Let $f \in F, x(.) \in B(x^*(.), \varsigma)$, $f, x(.)$ arbitrary. For each s, $\tilde{x}(.) \to f(., s, \tilde{x}(\to s)) : B(x^*(.), \varsigma) \to C(J, X)$ has a closed d.d. container $D_3 f(., s, x(.))(q(.))$ at $x(.)$, for each $q(.) \in C(J, X)$. Furthermore, $q(.) \to D_3 f(., s, x(.))(q(.))$, for each s, depends only on past values of $q(.)$, (values $q(s'), s' \leq s$). Given any $\tilde{f}, \tilde{f}' \in coF$, let $\hat{Q}(\tilde{f}, \tilde{f}', f, x(.))$ be the set of anti-differentiable functions $y(.) : J \to C(J, X)$ satisfying

$$(*) \quad dy(s)ds \in [\tilde{f}(., s, x(\to s)) - \tilde{f}'(., s, x(\to s))] + D_3 f(., s, x(.))(x^y(\to s)) \text{ a.e., } y(0) = 0$$

where $x^y(.)$ is the function $s \to y(s)(s)$. For each $\tilde{f}$ and each s, $D_3 f(., s, x(.))(x^y(\to s))$ is uniform in $y(.)$ in the set $\hat{Q}(\tilde{f}, f, f, x(.))$. Moreover, $q(.) \to D_3 f(., s, x(.))(q(\to s))$ is Lipschitz continuous with rank $\leq M^f$ on $C(J, X)$ and, as a function of s, $D_3 f(., s, x(.))(q(\to s))$ is measurable and essentially separably valued, (i.e. for some separable set $X_{f,x(.),q(.)} \subset C(J, X)$, $D_3 f(., s, x(.))(q(\to s)) \subset X_{f,x(.),q(.)}$ for a.e. s). (10)

For each $f \in F$, $\tilde{f} \in coF$, $x(.) \in B(x^*(.), \varsigma)$, at $x(T)$, $x \to G(x)$ and $x \to \phi(x)$ are assumed to have d.d. containers in all directions v, being uniform in $v \in \check{Q}(T, \tilde{f}, f, x(.)) := \{y(T)(T) : y(.) \in \hat{Q}(\tilde{f}, f, f, x(.))\}$,(for $\hat{Q}(., ., .)$, see (10)). (11)

4. Necessary conditions for optimality

To establish necessary conditions for nonsmooth as well as smooth problems with end targets in a Banach state space, some controllability properties of the first order variations are needed. (Using generalized directional derivatives, in the approach of this paper, such controllability properties are even needed in finite dimensions, see Seierstad (1994).) In the present nonsmooth case, (15) and (19) below represent such conditions.

A set of necessary conditions for the optimality of $(x^*(.), f^*)$ are provided in the following two theorems. The conclusion of Theorem 1 is a consequence of an exact penalization result. In case of "enough differentiability", for the point target $C = \{0\}$, the conclusion reads roughly as follows: There exists a non-negative constant Λ such that for $\phi^*(f) := \phi(x^f(T)) - \Lambda|G(x^f(T))|$ and $f_\delta = \hat{f}1_{[t,t+\delta]} + f^*(1 - 1_{[t,t+\delta]})$, ($\hat{f}$ arbitrary in F), the inequality $[d\phi^*(f_\delta)/d\delta \le 0]_{\delta=0}$ holds, (here a right derivative is taken). The first order condition in Theorem 1 in certain cases entails a maximum principle, which is the content of Corollary 1.

By the assumptions on the f 's in F, it can be shown that there exists a $\zeta > 0$, (namely the one defined in (14) below), such that the state equation has a unique system solution $x^f(.)$ with $x^f(t) \in clB(x^*(t), \varsigma/2)$, for all $f \in clB(f^*, \zeta) \subset F$, (see (24)(i) below).

For $\tilde{f} \in coF, f \in clB(f^*, \zeta) \subset F$, let

$$Q(\tilde{f}, f) = \hat{Q}(\tilde{f}, f, f, x^f(.)), \tag{12}$$

(for $\hat{Q}(.,.,)$ see (10)), and let

$$Q(t, \tilde{f}, f) : \{q(t)(t) : q(.) \in Q(\tilde{f}, f)\}. \tag{13}$$

Define ζ by the equality

$$2M\zeta \exp(TM^{f^*}) = \varsigma/2; \text{ for } \varsigma, M, M^{f^*}, \text{ see (7).} \tag{14}$$

Theorem 1. Consider problem (4)-(11). Assume that $(x^*(.), f^*)$ is an optimal admissible pair in the problem. Assume that there exists a triple $(\mu, \hat{\mu}, \mu'), \mu \in (0, 1), \hat{\mu} \in (0, \zeta], \mu' \in (0, \infty)$, with the property that for any $f \in clB(f^*, \hat{\mu}) \subset F$, for any $v \in Y$ with $|v| = \mu'$, there exists a triple $(\tilde{f}, c'', \tilde{\delta}) \in coF \times (C \cap clB(G(x^*(T)), 1)) \times (0, \infty)$ such that $\tilde{f} \in coB(f, \tilde{\delta})$, $(B(f, \tilde{\delta}) \subset F)$, and

$$\sup\{|\tilde{\delta}v - z + \tilde{\delta}[c'' - G(x^*(T))]| : z \in DG(x^f(T))(Q(T, \tilde{f}, f))\} \\ \le (1 - \mu)\tilde{\delta}\mu'. \tag{15}$$

Then the following necessary condition holds: For any $\tilde{f} \in coF$, and any $c \in C$, the inequality $0 \ge \inf \Omega_{c,\tilde{f}}$ holds, where

$\Omega_{c,\tilde{f}} := \{w - \Lambda|v| : (w, v)$ belongs to a triple (w, v, z) satisfying $z \in$

$$Q(T,\tilde{f},f^*),\ w \in D\phi(x^*(T))(z), v \in DG(x^*(T))(z) - c + G(x^*(T))\}, \quad (16)$$

and where

$$\Lambda := 64M_\phi M \exp(M^{f^*}T)\mu^{-1}\max\{1/\mu', 1 + 1/\check{\mu}\},$$
$$\check{\mu} := \min\{\hat{\mu}, \mu\mu'/8\}. \quad (17)$$

□

The condition $0 \geq \inf \Omega_{c,\tilde{f}}$ arises as a necessary conditions for optimality in a free end problem, where the end condition $G(x(T)) \in C$ is replaced by a penalization term. This result gives rise to necessary conditions based on separating linear functionals, in case linear Gâteaux derivatives exist, see Corollary 1 below.

For applications of necessary conditions of the above type to simple examples of finite dimensional problems involving ordinary differential equation, see Seierstad (1994). It is shown there that the conditions are sharper than those arising from a nonsmooth maximum principle of Clarke's type.

Corollary 1. Let $(x^*(.), f^*)$ be optimal in problem (4)-(9). Assume that, for all t, s, and $f \in F$, the map $x(.) \to f(., s, x(.)) : B(x^*(.), \varsigma) \to C(J, X)$ has a bounded linear Gâteaux derivative $f_x(., s, x(\to s))$ at all points in $B(x^*(.), \varsigma)$, (i.e. directional derivatives exists, and are linear in the "direction"). Assume that ϕ and G have bounded linear Gâteaux derivatives $\phi_x(x)$ and $G_x(x)$ in $B(x^*(T), \varsigma)$. Assume also that $x(.) \to f_x^*(., s, x(\to s))q(\to s)$ is continuous in $B(x^*(.), \varsigma)$, for each $q(.) \in C(J, X)$, and similarly, that $x \to G_x(x)q'$ is continuous in $B(x^*(T), \varsigma) \subset X$ for each $q' \in X$. Moreover, assume that

$$\text{all } M^f = M^* \text{ for some } M^* > 0, \text{ (see (7))}. \quad (18)$$

Finally, assume that for some $z^* \in Y$, some $\varepsilon > 0$, some $\hat{\mu} \in (0, \zeta]$, for all $f \in clB(f^*, \hat{\mu})$,

$$clB(z^*, \varepsilon) \subset cl\{G_x(x^f(T))q^{\tilde{f},f}(T)(T) - c + G(x^*(T)) :$$
$$c \in C \cap clB(G(x^*(T)), 1), \tilde{f} \in coF\}, \quad (19)$$

where $q^{\tilde{f},f}(.)$ is the unique element making up $\hat{Q}(\tilde{f}, f, f^*, x^f(.))$ in this case. Then the following maximum principle holds: For some non-negative number λ_0, and some bounded linear functional $\lambda^* \in Y^*$, (Y^* the topological dual of Y), $(\lambda^*, \lambda_0) \neq 0$, for any $\tilde{f} \in coF$ and $c \in C$,

$$\langle G_x(x^*(T))q^{\tilde{f},f^*}(T)(T) - c + G(x^*(T)), \lambda^*\rangle + \lambda_0\phi_x(x^*(T))q^{\tilde{f},f^*}(T)(T)$$

$$\leq 0. \tag{20}$$

□

In many cases, the condition (20) can be rewritten using co-state (adjoint) variables.

Remark 2. **a.** When $M^f = M^{f^*}$for all $f \in F$, then in condition (15), the set $Q(T, \tilde{f},\, f)$ can be replaced by $\{q(T)(T) : q(.) \in \hat{Q}(\tilde{f}, f, f^*, x^f(.))\}$, provided Λ and $\breve{\mu}$ are suitably modified, (still functions only of T, M_ϕ, M, M^{f^*}, μ,$\mu', \hat{\mu}$).

b. The condition (19) can be weakened as follows: Replace the inclusion in (19) by $clB(z^*, \varepsilon) \subset clco\{\hat{\delta}^{-1} G_x(x^f(T)) q^{\hat{f},f}(T)(T) - c + G(x^*(T)) :$ $\hat{\delta} > 0, \hat{f} \in B(f^*, \hat{\delta}), c \in C \cap clB(G(x^*(T)), 1)\}$.

c. If $z^* = 0$ in (19) (or in the condition in **b**.), then in Corollary 1, (18) and the continuity of $x(.) \to (f^*_x(., s, x(.))q(.)$ and $x \to G_x(x)q'$ can be deleted.

Remark 3. In the situation of Corollary 1, in the case where $x \to G(x)$ and $x(.) \to f^*(., s, x(.))$ have Frechét derivatives that are continuous in $B(x^*(T), \varsigma) \subset X$ and $B(x^*(.), \varsigma) \subset C(J, X)$, respectively, then (19) need only hold for $f = f^*$. (This comment does not pertain to the weakened condition in Remark 2 b.)

Proofs of the above results are given in Sections 5-11.

Remark 4. In nonsmooth problems, in the setting of Remark 1, to obtain better tools, it is often needed to use strong variations in $u(.)$ (switching) in combination with weak "variations" of the type $(1 - \lambda)u(t) + \lambda v(t) \in U$, $\lambda \in [0, 1]$, (U for the moment assumed to be a convex set in a normed space). Theorem 1 utilizes only strong variations. By rewriting the system using an auxiliary state, (actually λ), it is possible to obtain weak variations as strong ones, (see Seierstad (1994)). Hence, when Theorem 1 is applied to the rewritten system, in effect also weak variations occur.

5. More general necessary conditions.

Necessary conditions, more general than those in the preceding section, are presented in Lemma 1 and Theorems 2 and 3 below.

In the results above, "approximate" controllability properties related to linearized variations are used. ("Approximate" here means a quite rough approximation: The v's in (15) are only roughly approximated by the variations.) The next lemma utilizes approximate controllability by exact variations of the system. Then no assumptions concerning existence of d.d. containers are needed. To indicate the content of the lemma, note that it is related to the necessary condition that if a point c^{**} in $C - G(x^*(T))$ is attainable by $G(x^{\hat{f}}(T)) - G(x^*(T))$, for a perturbation $\hat{f}$ of f^*, then, by optimality, the corresponding change in ϕ, $\phi(x^{\hat{f}}(T)) - \phi(x^*(T))$, must be nonpositive. If c^{**}, and points nearby, are only known to be roughly attainable (see the first inequality in (21) below), the corresponding change in the criterion cannot be two large, (see the last inequality in (21)). It is here needed to include perturbations not only of f^*, but also of f's near f^*.

Lemma 1. (No d.d. containers, no switching) Let $(x^*(.), f^*)$ be optimal in problem (4)-(7), and let ζ be as defined in (14). Assume that $G(x)$ and $\phi(x)$ are continuous in $B(x^*(T), \varsigma)$. Then there exists no quintuple $(K, c^{**}, \mu, \hat{\mu}, \mu') \in (0, \infty) \times (C - G(x^*(T))) \times (0, 1) \times (0, \zeta] \times (0, \infty)$ with the property that, for all $f \in clB(f^*, \hat{\mu}) \subset F$, all $c \in C \cap clB(G(x^*(T), \hat{\mu})$, all $\hat{v} := (v, \omega) \in Y \times (-\infty, 0]$ with $|\hat{v} - (c^{**}, 0)| = \mu'$, all $r \in (0, 1]$, there exists a triple $(\hat{f}, c'', \delta)$, $\hat{f} \in clB(f^*, \zeta) \subset F$, $c'' \in C$, and $\delta \in (0, r]$, such that

$$\begin{aligned}
&|G(x^{\hat{f}}(T)) - G(x^f(T)) - \delta v - \delta(c'' - c)| \leq (1 - \mu)\delta\mu'|\hat{v}|/(|c^{**}| + \mu'),\\
&\sigma(\hat{f}, f) \leq \delta K|\hat{v}|,\ |c'' - c| \leq K|\hat{v}|, \text{ and}\\
&\phi(x^{\hat{f}}(T)) - \phi(x^f(T)) + \delta\omega \geq -(1 - \mu)\delta\mu'|\hat{v}|/(|c^{**}| + \mu').
\end{aligned} \tag{21}$$

Proof. Let us first prove (22),(23), and (24) below, which partly contain results only needed later on. If $\tilde{f} = \sum \lambda_i f_i \in coF$, ($f_i \in F$), let $M^{\tilde{f}} := \sum_i \lambda_i M^{f_i}$,(see (7)).

> Let $\tilde{f}, f \in coF$ be given. Assume that there exists an integrable function $\alpha(.)$ such that $x(.) \to f(., s, x(.))$ has Lipschitz rank $\alpha(s) \leq M^f$ in $B(x^*(.), \varsigma)$. Then, $|x^{\tilde{f}}(t) - x^f(t)|$
> $\leq \exp(\int_0^t \alpha(s)ds) \sup_{t' \leq t} |\int_0^{t'} \{\tilde{f}(t', s, x^{\tilde{f}}(\to s)) - f(t', s, x^{\tilde{f}}(\to s))\}ds|$
> for $t \in [0, T']$, $(T' \leq T)$, whenever $x^{\tilde{f}}(t)$ and $x^f(t)$ are solutions that exist on $[0, T']$ and belong to $B(x^*(t), \varsigma)$ for all t. (22)

Hence, if $\tilde{f}, f \in coF$, and if $x^{\tilde{f}}(t)$ and $x^f(t)$ are solutions that exist on $[0, T']$ and belong to $B(x^*(t), \varsigma)$ for all t, then, by (7), for all $t \in [0, T']$

$$\begin{aligned}&\text{(i) } |x^{\tilde f}(t)-x^f(t)| \le \exp(\textstyle\int_0^t \alpha(s)ds)2M\sigma(\tilde f, f),\\ &\text{(ii) } |x^{\tilde f}(t)-x^f(t)| \le \exp(\textstyle\int_0^t \alpha(s)ds)T'||\tilde f - f||,\end{aligned} \tag{23}$$

where $||\tilde f - f|| := \sup_{t,s,x(.)\in B(x^*(.),\varsigma)} |\tilde f(t,s,x(.)) - f(t,s,x(.))|$. Finally,

(i) For all $\tilde f, f$ in $clB(f^*, \varrho\zeta) \subset F$, $\varrho \in (0,2)$, solutions $x^{\tilde f}(s)$ and $x^f(s)$ exist on all $[0,T]$, belong to $clB(x^*(s), \varrho\varsigma/2)$, and $|x^{\tilde f}(t) - x^f(t)| \le \exp(M^f t)2M\sigma(\tilde f, f)$.

(ii) Let $f \in clB(f^*, \zeta)$, $\tilde f \in coF$, with $||\tilde f - f|| \le \varsigma/4T\exp(M^f T)$. Then, a solution $x^{\tilde f}(s)$ exists on all $[0,T]$, belongs to $clB(x^f(s), \varsigma/4) \subset clB(x^*(s), 3\varsigma/4)$ and $|x^{\tilde f}(t) - x^f(t)| \le T||f - \tilde f|| \exp(M^f t)$. (24)

To prove (22), define $|x^{\tilde f}(.) - x^f(.)|_t := \sup_{s\le t} |x^{\tilde f}(s) - x^f(s)|$. If $x^{\tilde f}(s)$ and $x^f(s)$ belong to $B(x^*(s), \varsigma)$ for all s, then $|x^{\tilde f}(t) - x^f(t)| \le$

$$\begin{aligned}&|\textstyle\int_0^t \{\tilde f(t,s,x^{\tilde f}(.)) - f(t,s,x^{\tilde f}(.))\}ds| + |\int_0^t \{f(t,s,x^{\tilde f}(.)) - f(t,s,x^f(.))\}ds| \le\\ &|\textstyle\int_0^t \{\tilde f(t,s,x^{\tilde f}(.)) - f(t,s,x^{\tilde f}(.))\}ds| + \int_0^t \alpha(s)|x^{\tilde f}(.) - x^f(.)|_s ds,\end{aligned}$$

so $|x^{\tilde f}(.) - x^f(.)|_t \le$
$$\sup_{t'\le t} |\textstyle\int_0^{t'} \{\tilde f(t',s,x^{\tilde f}(.)) - f(t',s,x^{\tilde f}(.))\}ds| + \int_0^t \alpha(s)|x^{\tilde f}(.) - x^f(.)|_s ds.$$

By a version of Gronwall's inequality, evidently (22) (and then also (23)) holds.

To prove (24)(i), apply a suitable local existence theorem to obtain that, for any $\tilde f \in clB(f^*, \varrho\zeta)$, a solution $x^{\tilde f}(t)$ of (4) exists in $B(x^*(t), \varrho\varsigma)$ on some interval $[0, T']$. (In fact, the standard local existence theorems for Volterra nonretarded equations in finite (even one) dimension in the Lipschitzian case, see e.g. Pogorzelski (1966), p.191, carry over virtually without change to the present case.) By (23)(i) and $\alpha(.) \le M^f$, it follows that $|x^{\tilde f}(t) - x^{f^*}(t)|$ $\le \exp(M^{f^*}t)2M\sigma(\tilde f, f^*)$, so, in fact $x^{\tilde f}(t)$ belongs to $clB(x^*(t), \varrho\varsigma/2)$, by (14). A continuation argument yields that the solution $x^{\tilde f}(.)$ exists on all $[0,T]$ and belongs to $clB(x^*(t), \varrho\varsigma/2)$ for all t. Hence, (24)(i) follows. The property (24)(ii) has a similar proof.

To finish the proof of Lemma 1, note that for $A = clB(f^*, \zeta) \subset F, H(f) = G(x^f(T))$, and $\eta(f) = -\phi(x^f(T))$, $a^* = f^*, \mu'' = \mu$, Corollary 5 in Section

10 below yields the desired conclusion. (The existence of $x^f(.)$ and the continuity required in (46) is provided by (24)(i).) $\square$

When (8) and (9) hold, then, no greater generality is gained by allowing $c^{**} \neq 0$ in Lemma 1. So in this case, it suffices to note that no quadruple $(K, \mu, \hat{\mu}, \mu')$ exists, such that (21) holds with $c^{**} = 0$. Note that when (25) below fails (i.e. no quadruple exists, with the asserted property), then trivially no quintuple with $c^{**} = 0$ exists with the property described in Lemma 1, (it suffices to note that then (21) fails for $\omega = 0$). In Lemma 2, (25) yields an exact penalization result, when combined with slightly stronger continuity assumptions.

Lemma 2. (Exact penalization. No d.d. containers. No switching) Let $(x^*(.), f^*)$ be optimal in problem (4)-(7), and let ζ be as defined in (14). Assume that $G(x)$ is continuous - and $\phi(x)$ Lipschitz continuous - in $B(x^*(T), \varsigma)$, the Lipschitz rank of ϕ being M_ϕ. Assume, moreover, that there exists a quadruple $(K', \mu, \hat{\mu}, \mu') \in (0, \infty) \times (0, 1) \times (0, \zeta] \times (0, \infty)$ with the property that for all $f \in clB(f^*, \hat{\mu}) \subset F$, all $v \in Y$ with $|v| = \mu'$, all $r \in (0, 1]$, there exists a triple $(\hat{f}, c'', \hat{\delta})$, $\hat{f} \in clB(f^*, \zeta) \subset F$, $c'' \in C$, and $\hat{\delta} \in (0, r]$, such that

$$\begin{aligned} &|G(x^{\hat{f}}(T)) - G(x^f(T)) - \hat{\delta}v - \hat{\delta}(c'' - c^*)| \leq (1 - \mu)\hat{\delta}\mu', \\ &\sigma(\hat{f}, f) \leq \hat{\delta}K'\mu', \text{ and } |c'' - c^*| \leq K'\mu', \end{aligned} \tag{25}$$

where $c^* = G(x^*(T))$. Then, (f^*, c^*) maximizes

$$\iota^*(f, c) := \phi(x^f(T)) - 16\mu^{-1}KMM_\phi|G(x^f(T)) - c| \exp(\beta^f) \tag{26}$$

for (f, c) in $clB(f^*, \check{\mu}/2) \times (C \cap clB(c^*, \check{\mu}/2))$, where $\check{\mu} = \min\{\hat{\mu}, \mu\mu'/2\}$, $K := \max\{K', 1 + K'\mu'/\check{\mu}\}$, and $\beta^f := M^f\sigma(f, f^*) + M^{f^*}(T - \sigma(f, f^*))$.

Proof: For each f in F, there exists a set C_f such that $\sigma(f, f^*) = \text{meas}(C_f)$ and $C_f \supset \{s : f(t, s, x(.)) \neq f^*(t, s, x(.))\}$ for all $t \in J, x(.) \in C(J, X)$. Define $\alpha^f(.) = M^f 1_{C_f} + M^{f^*}(1 - 1_{C_f})$ and $c^* := G(x^*(T))$. For $A = clB(f^*, \zeta)$, $H(f) = G(x^f(T))$, $\eta(f) = -\phi(x^f(T))$, $a^* = f^*, \mu^* = \hat{\mu}$, and $\mu'' = \mu$, Corollary 7 in Section 10 below yields the conclusion in the lemma. (For $a = f$, let W_a in the Corollary equal $M_\phi 2M \exp(\int_0^T \alpha^f(s)ds) = M_\phi 2M \exp(\beta^f)$, see (23)(i).) $\square$

In the next result, (8) and (9) are added to the set of properties postulated, in particular, switching closedness is utilized.

Theorem 2. (Exact penalization, no d.d. containers, switching) Let $(x^*(.), f^*)$ be optimal in problem (4)-(9) and let ζ be as defined in (14). Assume that there exists a triple $(\mu'', \hat{\mu}, \mu') \in (0,1) \times (0, \zeta] \times (0, \infty)$ such that for all $f \in clB(f^*, \hat{\mu}) \subset F$, and for all $v \in Y$ with $|v| = \mu'$, there exists a triple $(\tilde{f}, c'', \tilde{\delta}), \tilde{\delta} > 0, \tilde{f} \in coB(f, \tilde{\delta}), (B(f, \tilde{\delta}) \subset F), c'' \in C \cap clB(G(x^*(T)), 1)$, such that for all $r \in (0, \min\{1, \varsigma/8TM \exp(TM^f)\}]$, a $\delta \in (0, r]$ exists, such that:

$$|G(x^{\tilde{f}^\delta}(T)) - G(x^f(T)) - \delta\tilde{\delta}v - \delta\tilde{\delta}(c'' - G(x^*(T)))| \leq (1 - \mu'')\delta\tilde{\delta}\mu' \quad (27)$$

where

$$\tilde{f}^\delta := \delta\tilde{f} + (1 - \delta)f. \quad (28)$$

Then, for any $\tilde{f} \in coF$, for any $c \in C$, for $f^\delta := \delta\tilde{f} + (1 - \delta)f^*, \breve{\mu}'' := \min\{\hat{\mu}, \mu''\mu'/4\}, K := \max\{1/\mu', 1 + 1/\breve{\mu}''\}$, and $\iota_{c,\tilde{f}}(\delta) :=$

$$\phi(x^{f^\delta}(T)) - 32KMM_\phi\mu''^{-1}\exp(TM^{f^*})|G(x^{f^\delta}(T)) - G(x^*(T)) - \delta c + \delta G(x^*(T))|,$$

the inequality $\limsup_{\delta\searrow 0}[\iota_{c,\tilde{f}}(\delta) - \iota_{c,\tilde{f}}(0)]/\delta \leq 0$ holds. □

Note that when $\delta \in (0, \varsigma/8MT\exp(TM^f)]$, $x^{\tilde{f}^\delta}(t)$ exists on $[0, T]$ and belongs to $clB(x^*(t), 3\varsigma/4)$, (see (24)(ii) and note that $||\tilde{f}^\delta - f|| = \delta||\tilde{f} - f|| \leq \delta 2M$). A similar remark pertains to $x^{f^\delta}(.)$. A proof of this result is given in the next section.

6 Reduction to the case of ordinary retarded equations

It is possible to reduce the proofs of Theorems 1 and 2 to the case of ordinary retarded equation. This will now be proved. Let $\pi_s : C(J, C(J, X)) \to X$ be defined by $\pi_s z(.) = z(s)(s)$, $(z(.) \in C(J, C(J, X))$, and let $\pi :$ $C(J, C(J, X))$ into $C(J, X)$ be defined by $\pi(z(.)) := s \to \pi_s z(.) = z(s)(s)$. To each $f \in F$, there corresponds a function $\check{f}(s, z(.)) := f(., s, \pi z(.))$, defined on $J \times C(J, C(J, X))$, with values in $C(J, X)$. Let $\check{F} = \{\ \check{f} : f \in F\}$. Consider the ordinary retarded differential equation

$$dz/ds = \check{f}(s, z(\to s)), z(0) = x_0,\ \check{f}(s, z(.)) := f(., s, \pi z(.)), \tag{29}$$

with state space $C(J, X)$, $(x_0 = x_0(.))$. By retarded dependence in f, $\check{f}(s, z(.))$ depends only on the values $z(s'), s' \le s$. By definition, a solution $z(.)$ to (29) is absolutely continuous and has an integrable derivative existing a.e.

Lemma 3. Assume (6),(7),(9). To each solution $z(t) \in C(J, C(J, X))$ of (29) for which $z(t)(t) \in B(x^*(t), \varsigma)$ for all t, there corresponds a solution of (4) that belongs to $B(x^*(t), \varsigma)$, and vice versa.

Proof: Let $x(t)$ be a solution of (4) in $B(x^*(t), \varsigma)$ for some given f, and define

$$z_{x(.)}(\tau, t) := x_0(t) + \textstyle\int_0^\tau f(t, s, x(\to s))ds,\ \tau, t \in J, \tag{30}$$

(by (7), the integrand is bounded). Then $z(\tau) := t \to z_{x(.)}(\tau, t)$ satisfies (29), $(z(\tau) = x_0 + \int_0^\tau f(., s, x(\to s))ds$, the integral exists in $C(J, X)$, by (9)). On the other hand, let $z(\tau) := t \to z(\tau)(t)$ be a solution of (29) with $z(t)(t) \in B(x^*(t), \varsigma)$, for some f. (Note that then the right hand side is bounded by M). Then, define $x(t) :=$

$$x_{z(.)}(t) := z(t)(t). \tag{31}$$

Evidently, $(x(t), f)$ then satisfies (4). □

Note that when $f, \hat{f} \in B(f^*, \zeta)$, the corresponding solutions $z^f(t)$ and $z^{\hat{f}}(t)$ of (29) exist in $clB(z^{f^*}(t), \varsigma/2)$, for all t in [0,T] and satisfy

$$|z^f(t) - z^{\hat{f}}(t)| \le \exp(M^f t)2M\sigma(f, \hat{f}), \tag{32}$$

by arguments analogous to those leading to (24).

From these observations, it follows that necessary conditions for optimality for systems of the type (4)-(9) automatically follow from necessary conditions for optimality for systems governed by ordinary retarded differential equations. In this case, the functions f do not depend on t, closedness under switching is defined in the same manner as before, the σ-metric $\sigma(f, \hat{f})$ is now $\sigma(f, \hat{f}) := \inf\{\text{meas}(M') : M'$ is measurable and contains $\{s : f(s, x(.)) \neq \hat{f}(s, x(.))\}$ for all $x(.) \in C(J, X)\}$. Further properties utilized in this case (specializations of (6), (7), (9)) are the following:.

The set F of functions f is assumed to be essentially σ-closed, and any $f = f(s, x(.))$ in F is separately measurable in s and satisfies the retardedness condition: $f(s, x(.)) = f(s, \check{x}(.))$ if $x(\tau) = \check{x}(\tau)$ for $\tau \geq s, x(.), \check{x}(.) \in C(J, X)$. (33)

F is closed under switching. (34)

Moreover, for some $\varsigma > 0$, some $M > 0$, for all f in F, for some M^f, for all s,

$\sup_{x(.)\in B(x^*(.),\varsigma)} |f(s, x(\to s))| \leq M$ and $x(.) \to f(s, x(\to s))$ is Lipschitz continuous of rank $\leq M^f$ in $B(x^*(.), \varsigma) \subset C(J, X)$. (35)

Sometimes, we will also need to refer (8), which is repeated here:

$G(x) : X \to Y$ and $\phi(x) : X \to \mathbb{R}$ are assumed to be Lipschitz continuous in $B(x^*(T), \varsigma)$ with ranks M_G and M_ϕ, resp. (36)

In this case, (4) reduces to

$$dx/dt = f(s, x(\to s)) \text{ a.e.}, x(0) = x_0 \in X, \tag{37}$$

where a solution $x(s)$, by definition, is antidifferentiable.

The optimization problem is again: For the given entities x_0, ϕ, G, F, and C (a closed convex set in Y),

$$\max_{x(.),f} \phi(x(T)), \text{ subject to } (37),\ G(x(T)) \in C \text{ and } f \in F. \tag{38}$$

From now on, we shall only give proofs for the problem (37), (38), (these proofs automatically also provide proofs in case of system (4),(5)).

Evidently, Lemma 1 and Lemma 2 in particular hold for the present type of systems, and this is formally stated below.

Lemma 4. (Ordinary retarded equation, no d.d. container, no switching). In Lemma 1 the reference to problem (4)-(7) can be replaced by a reference to problem (33), (35),(37) (38).

Lemma 5. (Ordinary retarded equation, no d.d. container, no switching). In Lemma 2 the reference to problem (4)-(7) can be replaced by a reference to problem (33),(35),(37) (38).

Lemma 6. (Ordinary retarded equations, no d.d. containers, switching). Let $(x^*(.), f^*)$ be optimal in problem (27), (33)-(35),(37), (38) and assume that $G(x)$ is continuous, and $\phi(x)$ is Lipschitz continuous, in $B(x^*(T), \varsigma)$, the Lipschitz rank of ϕ being M_ϕ. Then $(f^*, c^*), (c^* := G(x^*(T)))$, maximizes $\tilde{\iota}(f, c)$ for (f, c) in $clB(f^*, \check{\mu}/2) \times (C \cap clB(c^*, \check{\mu}/2))$, where $\tilde{\iota}(f, c) := \phi(x^f(T)) - 32KMM_\phi\mu''^{-1}\exp(\beta^f)|G(x^f(T)) - c|$, $\check{\mu} := \min\{\hat{\mu}, \mu''\mu'/4\}$, $K = \max\{1/\mu', 1 + 1/\check{\mu}\}$, and $\beta^f := M^f\sigma(f, f^*) + M^{f^*}(T - \sigma(f, f^*))$.

This lemma and the next one is given a joint proof.

Lemma 7. (Ordinary retarded equation, no d.d. containers, switching) Theorem 2 holds for (4)-(9) replaced by (33)-(38).

Proof: Let $\hat{r} \in (0, 1]$. It is first proved that (27) implies (25) for $\mu = \mu''/2$ and $K' := 1/\mu'$. Assume that (27) holds for $(f, v, \tilde{f}, c'', \tilde{\delta}, r, \delta)$, $r \le \min\{\hat{r}/\tilde{\delta}, \zeta/\tilde{\delta}\}]$, $\tilde{f} := \sum_i \lambda_i f_i \in coB(f, \tilde{\delta}), (f_i \in B(f, \tilde{\delta}) \subset F)$, and $\tilde{f}^\delta = \delta\tilde{f} + (1 - \delta)f$. By continuity of G, there exists an $\varepsilon > 0$, such that $|G(x) - G(x^{\tilde{f}^\delta}(T))| \le \delta\tilde{\delta}\mu'\mu''/2$, when $|x - x^{\tilde{f}^\delta}(T)| < \varepsilon$. Choose sets C^i such that $C^i \supset \{s : f_i(s, x(.)) \ne f(s, x(.))\}$ for all continuous $x(.)$ and $meas(C^i) = \sigma(f_i, f)$. For $\alpha := \max\{M^f, \max_i M^{f_i}\}$, the inequality $\sup_{t' \le T} |\int_0^{t'} \{\tilde{f}^\delta(s, x^{\tilde{f}^\delta}(s)) - \hat{f}(s, x^{\tilde{f}^\delta}(s))\}ds| < \varepsilon\exp(-T\alpha)$ can be obtained for a function $\hat{f} \in F$ constructed by "rapid switching" between f and the f_i's with "weights" $1 - \sum_i \delta\lambda_i$ and $\delta\lambda_i$. More precisely, $\hat{f}$ equals $\sum_i f_i 1_{C_i} + (1 - \sum_i 1_{C_i})f$, where meas$(C_i) = \delta\lambda_i T$, C_i measurable and disjoint, see e.g. Seierstad (1975, Remark 10.1). Note that $\int_J \sum_i \lambda_i 1_{C^i} ds < \tilde{\delta}$. By "rapid switching", the sets C_i can be chosen such that also $|\int_J \{\sum_i (1_{C^i} 1_{C_i} - \delta\lambda_i 1_{C^i})\}ds| < \delta\tilde{\delta} - \int_J \sum_i \delta\lambda_i 1_{C^i} ds$, which yields $\int_J \sum_i 1_{C^i} 1_{C_i} ds < \delta\tilde{\delta}$. Evidently, for all continuous $x(.)$, $\{s : \hat{f}(s, x(.)) \ne f(s, x(.))\} \subset \cup_i\{C_i \cap C^i\}$, so $\sigma(\hat{f}, f) < \delta\tilde{\delta} \le \zeta$. Existence of $x^{\hat{f}}(.)$ follows from (24)(i). By (22) and the above inequality involving α, it follows that $|x^{\tilde{f}^\delta}(T) - x^{\hat{f}}(T)| < \varepsilon$, so $|G(x^{\tilde{f}^\delta}(T)) - G(x^{\hat{f}}(T))| \le \delta\tilde{\delta}\mu'\mu''/2$. The latter inequality and (27) yield (25) for $\mu = \mu''/2, \hat{\delta} = \delta\tilde{\delta} \in (0, \hat{r}], c' = \hat{\delta}c'' + (1 - \hat{\delta})c$. As $K' = 1/\mu'$, $|c' - G(x^*(T))| \le \delta\tilde{\delta}K'\mu'$ and $\sigma(\hat{f}, f) \le \delta\tilde{\delta}K'\mu'$. Hence, Lemma 6 is valid.

To prove Lemma 7, from Lemma 6 we know that (f^*, c^*) ($c^* := G(x^*(T))$) maximizes $\tilde{\imath}(f, c)$ in $clB(f^*, \check{\mu}/2) \times (C \cap clB(c^*, \check{\mu}/2))$. Fix any $c \in C$ and $\tilde{f} := \sum_i \lambda_i f_i \in coF$, let f_δ be any switching combination of the f_i's and f^* with weights $\delta\lambda_i$ and $(1 - \delta \sum_i \lambda_i), \delta \in (0, \zeta]$, and let $\hat{\imath}(f_\delta, c) := \phi(x^{f_\delta}(T)) - \tilde{\Lambda} \exp(M^{f^*}T)|G(x^{f_\delta}(T)) - c|$, where $\tilde{\Lambda} := 32KMM_\phi/\mu''$. We can take $M^{f_\delta} := \max\{f^*, \max_i\{M^{f_i}\}\}$, $\beta^{f_\delta} := M^{f_\delta}\sigma(f_\delta, f^*) + M^{f^*}(T - \sigma(f_\delta, f^*))$. Let $\Lambda' := M_G \exp(TM^{f^*})2MT + |c - c^*|$ and let ε be an arbitrary positive number. As $\text{meas}(C_{f_\delta}) \le \delta T$, then for $\delta' \in (0, \zeta]$, small enough, the inequality $\varepsilon' := |\exp(\beta^{f_\delta}) - \exp(TM^{f^*})| \le \varepsilon/3\tilde{\Lambda}\Lambda'$, is valid for all $\delta \in (0, \delta']$, uniformly in the set S_δ of switching combinations f_δ with the prescribed weights. Furthermore, note that (24)(i) gives, for $\delta \in (0, \delta']$, that $|G(x^{f_\delta}(T)) - G(x^*(T))| \le M_G \exp(TM^{f^*})2M\delta T$, so $\varepsilon'' := |G(x^{f_\delta}(T)) - c^* - \delta(c - c^*)|/\delta \le \Lambda'$. Now, Lemma 6 provides the inequality $0 \ge [\tilde{\imath}(f_\delta, \delta c + (1 - \delta)c^*) - \tilde{\imath}(f^*, c^*)]/\delta$, $\delta \in (0, \check{\mu}/2]$. Replacing the term β^{f_δ} occurring in $\tilde{\imath}(f_\delta, \delta c + (1 - \delta)c^*)$ by TM^{f^*} changes the right hand side by an amount $= \tilde{\Lambda}\varepsilon'\varepsilon'' \le \tilde{\Lambda}\varepsilon'\Lambda' \le \varepsilon/3$, which yields $\varepsilon/3 \ge [\hat{\imath}(f_\delta, \delta c + (1 - \delta)c^*) - \hat{\imath}(f^*, c^*)]/\delta$ for all switching combinations f_δ in $S_\delta, \delta \in (0, \min\{\delta', \check{\mu}/2\}]$. Finally, let $f^\delta := \delta\tilde{f} + (1 - \delta)f^*$, and note that for each $\delta \in (0, \delta''], \delta'' := \min\{\delta', \check{\mu}/2, \varsigma/8TM \exp(TM^{f^*})\}]$, by (22), (24) (ii), a switching combination $f_\delta \in S_\delta$ exists such that $|\phi(x^{f^\delta}(T)) - \phi(x^{f_\delta}(T))|/\delta < \varepsilon/3$ and $|G(x^{f^\delta}(T)) - G(x^{f_\delta}(T))|/\delta < \varepsilon/3\tilde{\Lambda} \exp(TM^{f^*})$. The three inequalities involving $\varepsilon/3$ yield $\varepsilon \ge [\hat{\imath}(f^\delta, \delta c + (1 - \delta)c^*) - \hat{\imath}(f^*, c^*)]/\delta$, $\delta \in (0, \delta'']$, and the conclusion of Lemma 7 (Theorem 2) follows.

7. Proof of necessary conditions involving d.d. containers

When assumption (10) holds, then $\check{f}$ (see 29) has the d.d. container $D_2\check{f}(s, x(.))(q(.)) := D_3 f(., s, \pi x(.))(\pi q(.)), q(.) \in C(J, C(J, X))$. In fact, the following properties then hold for the functions $\check{f}$ in $\check{F}$: Given any $s \in J$,

> at each $\check{x}(.) \in B(z^{\check{f}^*}(.), \varsigma) \subset C(J, C(J, X))$, the map
> $x(.) \to \check{f}(s, x(.)) : B(z^{\check{f}^*}(s), \varsigma) \to C(J, X)$, has a closed d.d.
> container $D_2\check{f}(s, \check{x}(.))(q(.))$ in all directions $q(.) \in C(J, C(J, X))$,
> which is uniform in $q(.) \in \tilde{Q}(\check{f}^+, \check{f}, \check{x}(.))$ for any $\check{f}^+ \in co\check{F}$, (39)

where the set $\tilde{Q}(\check{f}^+, \check{f}, \check{x}(.))$ consists of all antidifferentiable functions $\check{q}(.)$ on J with values in $C(J, X)$ that satisfy

$$d\check{q}(s)/ds \in \check{f}^+(s, \check{x}(.)) - \check{f}(s, \check{x}(.)) + D_2\check{f}(s, \check{x}(.))(\check{q}(.)) \text{ a.e.}$$

$$\check{q}(0) = 0. \tag{40}$$

Furthermore, $D_2\check{f}(s, \check{x}(.))(\check{q}(.))$ exhibits retarded dependence on $\check{q}(.)$, is Lipschitz continuous in $\check{q}(.)$ of rank $\leq M^f$, and, as function of s, is measurable and essentially separably valued. These observations make it possible to confine again the discussion to the case where the f's are independent of t. Hence, from now on in this section, F consists of functions $f(s, x(\to s)) : J \times C(J, X) \to X$. The following assumption is needed, (a specialization of (10) to the present case, with (11) added):

> Let $f \in F, x(.) \in B(x^*(.), \varsigma) \subset C(J, X), f, x(.)$ arbitrary. The map $\tilde{x}(.) \to f(s, \tilde{x}(.)), B(x^*(.), \varsigma) \to X$ has a closed d.d. container $D_2 f(s, x(.))(q(.))$ at $x(.)$, for each $q(.) \in C(J, X)$. Furthermore, $q(.) \to D_2 f(s, x(.))(q(.))$, for each s, depends only on past values of $q(.)$. Let $\tilde{f} \in coF$, $\tilde{f}$ arbitrary, and let $Q'(\tilde{f}, f, x(.))$ be the set of antidifferentiable functions $q(.) : J \to X$ satisfying $dq(s)/ds \in \tilde{f}(s, x(\to s)) - f(s, x(\to s)) + D_2 f(s, x(.))(q(\to s))$ a.e. For each s, $D_2 f(s, x(.))(q(\to s))$ is uniform in $q(.)$ in the set $Q'(\tilde{f}, f, x(.))$. Moreover, $q(.) \to D_2 f(s, x(.))(q(\to s))$ is Lipschitz continuous with rank $\leq M^f$ on $C(J, X)$ and, as a function of s, $D_2 f(s, x(.))(q(\to s))$ is measurable and essentially separably valued, (i.e. for some separable set $X_{f,x(.),q(.)} \subset X, D_2 f(s, x(.))(q(\to s)) \subset X_{f,x(.),q(.)}$ for a.e. s). At $x(T), x \to G(x)$ and $x \to \phi(x)$ are assumed to have d.d. containers in all directions v, being uniform in $v \in \check{Q}(T, \tilde{f}, f, x(.)) := \{q(T) : q(.) \in Q'(\tilde{f}, f, x(.))\}$. (41)

A joint proof will be given for the following two special cases of Theorem 2 and Theorem 1, respectively.

Lemma 8. (Ordinary retarded equation, d.d. containers, switching) Consider problem (33) - (38), (41). Assume that $(x^*(.), f^*)$ is an optimal admissible pair. Assume that the following condition holds: There exists a triple $(\mu, \hat{\mu}, \mu')$, $\mu \in (0, 1), \hat{\mu} \in (0, \zeta], \mu' \in (0, \infty)$, with the property that for any $f \in clB(f^*, \hat{\mu})$, for any $v \in Y$ with $|v| = \mu'$, there exists a triple $(\tilde{f}, c, \tilde{\delta}) \in coF \times (C \cap clB(G(x^*(T)), 1)) \times (0, \infty)$ such that $\tilde{f} \in coB(f, \tilde{\delta})$ and

$$\begin{aligned}&\sup\{|\tilde{\delta}v - z + \tilde{\delta}(c'' - G(x^*(T)))| : z \in DG(x^f(T))(Q(T, \tilde{f}, f))\} \\ &\leq (1 - \mu)\tilde{\delta}\mu',\end{aligned} \tag{42}$$

where $Q(T,\tilde f,f) := \check Q(T,\tilde f,f,x^f(.))$, (for $\check Q$, see (41)). Then the following necessary condition holds: For each $\tilde f \in coF, c \in C$, the inequality $\limsup_{\delta\searrow 0}[\check\iota_{c,\tilde f}(\delta) - \check\iota_{c,\tilde f}(0)]/\delta \leq 0$ holds, where $\check\iota_{c,\tilde f}(\delta) :=$

$\phi(x^{f^\delta}(T)) - \Lambda''|G(x^{f^\delta}(T)) - G(x^*(T)) - \delta c + \delta G(x^*(T))|$,

$K'' := \max\{1/\mu', 1 + 1/\min\{\hat\mu, \mu\mu'/8\}\}, \Lambda'' := 64K''MM_\phi\mu^{-1}\exp(TM^{f^*})$, and $f^\delta = \delta f + (1-\delta)f^*$.

Lemma 9. (Ordinary retarded equation, d.d. containers, switching) Consider problem (33) - (38), (41). Assume that $(x^*(.), f^*)$ is an optimal admissible pair and that (42) holds. Then, for any $\tilde f \in coF, c \in C$, the inequality $0 \geq \inf \Omega_{c,\tilde f}$ holds. (For $\Omega_{c,\tilde f}$, see (16).)

Proofs. We anticipate further results by claiming the validity of the following property, (it is proved subsequent to Lemma 10 in Section 9, after the development of the necessary amount of d.d. container calculus in Sections 8 and 9):

> For any $f \in clB(f^*, \zeta)$, for any $\tilde f \in coF$, for all $\varepsilon > 0$, there exists a $\delta^f \in (0, \varsigma/8TM\exp(TM^f)]$, such that for all $\delta \in (0, \delta^f]$, $\delta^{-1}(G(x^{\tilde f^\delta}(T)) - G(x^f(T)), \phi(x^{\tilde f^\delta}(T)) - \phi(x^f(T))) \in \Psi + B(0,\varepsilon)$, where $\Psi := \cup_{y\in Q(T,\tilde f,f)} DG(x^f(T))(y) \times D\phi(x^f(T))(y)$, (43)

and where $\tilde f^\delta := \delta\tilde f + (1-\delta)f$, ($x^{\tilde f^\delta}(t)$ exists in $clB(x^*(t), 3\varsigma/4)$, by(24)(ii)).

Lemma 8 follows from Lemma 7, with $\mu'' = \mu/2$, once (27) is proved: Choose δ^f so small that (43) holds for $\varepsilon = \tilde\delta\mu'\mu/2$. Then, by (42), for $\Delta^* := \{[G(x^{\tilde f^\delta}(T)) - G(x^f(T))]/\delta : \delta \in (0, \delta^f]\}$,
$\sup\{|\tilde\delta v - z + \tilde\delta(c'' - G(x^*(T)))| : z \in \Delta^*\} \leq (1-\mu'')\tilde\delta\mu'$. As (27) then holds, the conclusion of Lemma 7 holds: For any $\tilde f \in coF, c \in C, \limsup_{\delta\searrow 0}[\iota_{c,\tilde f}(\delta) - \iota_{c,\tilde f}(0)]/\delta \leq 0$, which yields the conclusion in Lemma 8. This conclusion says that for any $\varepsilon' > 0$, for all $\delta > 0$, small enough, $[\check\iota_{c,\tilde f}(\delta) - \check\iota_{c,\tilde f}(0)]/\delta \leq \varepsilon'/3$. By (43), for $f = f^*, \varepsilon = \varepsilon'/3, f^\delta = \delta\tilde f + (1-\delta)f^*$, for any δ small enough, there exist $z \in Q(T,\tilde f,f^*), v' \in DG(x^*(T))(z)$ and $w \in D\phi(x^*(T))(z)$, such that the inequalities $\Lambda''|v' - [G(x^{f^\delta}(T)) - G(x^*(T))]/\delta| \leq \varepsilon'/3$ and $|w - [\phi(x^{f^\delta}(T)) - \phi(x^*(T))]/\delta)| \leq \varepsilon'/3$ hold. The three inequalities involving $\varepsilon'/3$ yield $w - \Lambda''|v' - c + G(x^*(T))| \leq \varepsilon'$, and the conclusion of Lemma 9 follows.

8. Properties of d.d. containers

If g is a function of n variables x_i, the difference quotient of the function $x_i \to g(x_1, ..., x_n)$ in direction v_i is written $\Delta_i g(x_1, ..., x_n)(v_i, \lambda)$, and $\Delta_i^* g(x_1, ..., x_n)(v_i, r) := \{\Delta_i g(x_1, ..., x_n)(v_i, \lambda) : \lambda \in (0, r]\}$. A d.d. container of $x_i \to g(x_1, ..., x_n)$ is written $D_i g(x_1, ..., x_n)(v_i)$. The definition of d.d. container does not require the corresponding function to be locally Lipschitz continuous, so the latter property is explicitely mentioned in the results below when it is needed.

A few calculus rules for d.d. containers will now be presented. Let $X, \tilde{X}$, $Y, \tilde{Y}$, and Z be real normed spaces, and let $E : X \to Y, F : X \to Y, \tilde{F} : X \to \tilde{Y}, G : Y \to Z$ be given functions.

A). If $DF(x_0)(v)$ is a d.d. container of F at x_0 in direction v, then $\mu DF(x_0)(v)$ is a d.d. container of μF at x_0 in direction v.

B). Let $H = G(F(x))$. Let $F(x)$ have a d.d. container $DF(x_0)(v)$ at x_0, uniform in $v \in V$, V a given set, and let G have a d.d. container $DG(F(x_0))(\tilde{v})$ at $F(x_0)$, uniform in $\tilde{v} \in V' = DF(x_0)(V)$. Assume also that G is locally Lipschitz continuous near $F(x_0)$, and that V' is bounded. Then $DG(F(x_0))(DF(x_0)(v))$ is a d.d. container of H at x_0, uniform in $v \in V$.

Proof. Observe that $\Delta H(x_0)(v, \lambda) = \Delta G(F(x_0))(\tilde{v}, \lambda)$, where $\tilde{v} := \Delta F(x_0)(v, \lambda)$, so $\Delta^* H(x_0)(v, r) \subset$

$$\Delta^* G(F(x_0))(\Delta^* F(x_0)(v, r), r) := \cup_{\tilde{v} \in \Delta^* F(x_0)(v,r)} \Delta^* G(F(x_0))(\tilde{v}, r).$$

Now, G is Lipschitz continuous of rank $\leq M_G \geq 1$ in some ball $B(F(x_0), \varepsilon_G)$. Let ε be any given number in $(0, \varepsilon_G)$. Then, for some $r \in (0, 1)$, $\Delta^* F(x_0)(v, r) \subset DF(x_0)(v) + B(0, \varepsilon/2M_G)$ and $\Delta^* G(F(x_0))(w, r) \subset DG(F(x_0))(w) + B(0, \varepsilon/2)$ for all $v \in V, w \in V'$. Here, r can be chosen so small that $r < \varepsilon_G/2|V'| (\Rightarrow |V'| < \varepsilon_G/2r)$. It remains to prove that any $z \in \Delta^* G(F(x_0))(\Delta^* F(x_0)(v, r), r)$ belongs to $DG(F(x_0))(DG(x_0)(v)) + B(0, \varepsilon)$. Now, $z \in \Delta^* G(F(x_0))(w', r)$ for some $w' \in \Delta^* F(x_0)(v, r)$. There exists a $w \in DF(x_0)(v)$, such that $|w' - w| \leq \varepsilon/2M_G \leq \varepsilon_G/2M_G$. Now, $|w| < \varepsilon_G/2r$ and (hence) $|w'| < |w' - w| + \varepsilon_G/2r < \varepsilon_G/2M_G + \varepsilon_G/2r \leq \varepsilon_G/2r + \varepsilon_G/2 \leq \varepsilon_G/r, (r < 1)$. Using the last inclusion and Lipschitz continuity of $w'' \to \Delta^* G(F(x_0))(w'', r)$ of rank M_G in $B(F(x_0), \varepsilon_G/r)$ yield $z \in \Delta^* G(F(x_0))(w, r) + B(0, M_G\varepsilon/2M_G) \subset DG(F(x_0))(w) + B(0, \varepsilon)$. $\square$

Let $F \times \tilde{F}$ be the map $x \to (F(x), \tilde{F}(x))$. Because

$\Delta^*(F \times \tilde{F})(x_0)(v,r) \subset \Delta^* F(x_0)(v,r) \times \Delta^* \tilde{F}(x_0)(v,r)$, the proof of the following result is immediate:

C). Assume that $DF(x_0)(v)$ and $D\tilde{F}(x_0)(v)$ are d.d. containers of F and $\tilde{F}$ at x_0 , uniform in directions $v \in V$. Then, $DF(x_0)(v) \times D\tilde{F}(x_0)(v)$ is a d.d. container of $F \times \tilde{F}$, uniform in V.

Because $\Delta^*(E+F)(x_0)(v,r) \subset \Delta^* E(x_0)(v,r) + \Delta^* F(x_0)(v,r)$, we have:

D). Let $DE(x_0)(v)$ and $DF(x_0)(v)$ be d.d. containers of E and F at x_0, both uniform in $v \in V$. Then $DE(x_0)(v) + DF(x_0)(v)$ is a d.d. container of $E+F$ at x_0, uniform in $v \in V$.

E). Let $K(x,\tilde{x}) : X \times \tilde{X} \to Y$, and let $x \to K(x,\tilde{x})$ have a Frechét derivative $K_1'(x,\tilde{x})$ at all points $(x,\tilde{x})$ in a neighborhood of a given point $(x_0,\tilde{x}_0)$, $K_1'(x,\tilde{x})$ continuous at $(x_0,\tilde{x}_0)$. Assume also that $\tilde{x} \to K(x_0,\tilde{x})$ has a d.d. container $D_2K(x_0,\tilde{x}_0)(\tilde{v})$, uniform in $\tilde{v} \in \tilde{V}$, $\tilde{V}$ a given subset of $\tilde{X}$. Then $DK(x_0,\tilde{x}_0)(v,\tilde{v}) := K_1'(x_0,\tilde{x}_0)(v) + D_2K(x_0,\tilde{x}_0)(\tilde{v})$ is a d.d. container of $(x,\tilde{x}) \to K(x,\tilde{x})$ at $(x_0,\tilde{x}_0)$, uniform in $B(0,1) \times \tilde{V}$.

The proof follows easily from the equality $[K(x_0+\lambda v,\tilde{x}_0+\lambda\tilde{v}) - K(x_0,\tilde{x}_0)]/\lambda$ $= [K(x_0+\lambda v,\tilde{x}_0+\lambda\tilde{v}) - K(x_0,\tilde{x}_0+\lambda\tilde{x}_0)]/\lambda + [K(x_0,\tilde{x}_0+\lambda\tilde{v}) - K(x_0,\tilde{x}_0)]/\lambda$ and the inequality $|[K(x_0+\lambda v,\tilde{x}_0+\lambda\tilde{v}) - K(x_0,\tilde{x}_0+\lambda\tilde{x}_0)]/\lambda - K_1'(x_0,\tilde{x}_0)(v)| \leq$ $\sup_{\gamma\in[0,\lambda]} |K_1'(x_0+\gamma v,\tilde{x}_0+\lambda\tilde{x}_0) - K_1'(x_0,\tilde{x}_0)||v|$.

F). Let $g^*(s,x) : I \times X \to Y$, ($Y$ a Banach space, I a bounded interval $\subset \mathbb{R}$), be Lipschitz continuous in $x \in B(x_0,\gamma)$ of rank $\leq \kappa(s)$, for some given $\gamma > 0$ (independent of s), $x_0 \in X$, $\kappa(s)$ integrable. Assume that for each $x \in B(x_0,\gamma), s \to g^*(s,x)$ is integrable and define $\tilde{f}(x) = \int_I g^*(s,x)dt$. Assume also that, for each s, $x \to g^*(s,x)$ has a closed d.d. container $D_2g^*(s,x_0)(v)$ at x_0, which is uniform in $v \in V$, where V is a bounded set. Assume that $|D_2g^*(s,x_0)(V)| \leq \kappa'(t), \kappa'(t)$ integrable, and that, for each $v \in V, s \to D_2g^*(s,x_0)(v)$ is measurable and essentially separably valued. Then $D\tilde{f}(x_0)(v) := \int_I D_2g^*(t,x_0)(v)dt := \{\int_I w(t)dt : w(t) \in D_2g^*(t,x_0)(v)$ a.e., $w(.)$ integrable $\}$ is a d.d. container of $\tilde{f}(x)$ at x_0, uniform in $v \in V$.

A proof of this result is given in B. in Appendix.

As an example of a function with d.d. containers, let Ω be any measure space, let $g(x,\omega) : \mathbb{R}^n \times \Omega \to \mathbb{R}$ be Lipschitz continuous in x, uniformly in ω, and measurable in ω, with $g(0,.) \in L_2(\Omega,\mathbb{R})$, and let $f : L_2(\Omega,\mathbb{R}^n) \to L_2(\Omega,\mathbb{R})$ be defined by $f(x(.)) := \omega \to g(x(\omega),\omega), x(.) \in L_2(\Omega,\mathbb{R}^n)$. Then, for any $x(.), z(.) \in L_2(\Omega,\mathbb{R}^n)$, the following set is a d.d. container: $Df(x(.))(z(.)) := \{c(.) \in L_2(\Omega,\mathbb{R}) : \hat{\alpha}(\omega) \leq c(\omega) \leq \hat{\beta}(\omega)\}$, where $\hat{\alpha}(\omega) := \mathrm{liminf}_{\lambda\searrow 0} h(\lambda,\omega), \hat{\beta}(\omega) := \mathrm{limsup}_{\lambda\searrow 0} h(\lambda,\omega)$, and $h(\lambda,\omega) := \lambda^{-1}[g(x(\omega)+\lambda z(\omega),\omega) - g(x(\omega),\omega)]$.

9. Generalized variational equations

The following result has a proof similar to the proof of the selection result 3.17 in Clarke (1983). In the proof, Kuratowskii's measurability selection theorem, is needed, see Aubin and Frankowska (1990) for that theorem. In this section, X is a Banach space.

Lemma 10. (Selection of "neighbour") . Let H be an open subset of $C(J,X)$. Let $A(t,q(.))$ be a multifunction on $J \times H$, with values being nonempty closed sets in X, such that $A(t,q(.)) = A(t,q'(.))$ if $q(s) = q'(s)$ for $s > t$, $q(.), q'(.) \in H$, (retarded dependence). Assume that for each $q(.) \in H$, there exists a separable subset $X_{q(.)}$ of X such that $A(t,q(.)) \subset X_{q(.)}$ for a.e. t. Assume also that for each $q(.) \in H$, $t \to A(t,q(.))$ is measurable in t, and that, for each t, $q(.) \to A(t,q(.))$ is Lipschitz continuous of rank $\kappa(t)$ in H, $\kappa(.)$, integrable. Let $q_0(.)$ be a given antidifferentiable function in H, let $\lambda_0(.)$ be an integrable function such that $\lambda_0(t) \geq \mathrm{dist}(dq_0(t)/dt, A(t,q_0(.)))$ a.e., let $B(q_0(.),\varepsilon) \subset H$, $\varepsilon > 0$, and assume that, for all t, $\xi(t) := (2\int_0^t \lambda_0(s)ds)\exp(\int_0^t 2\kappa(s)ds) < \varepsilon$. Then there exists an antidifferentiable function $q(.)$, with $q(0) = q_0(0)$ such that $dq(t)/dt \in A(t,q(.))$ for a.e. t and such that $|q_0(t) - q(t)| \leq \xi(t)$.

See A. in Appendix for a proof.

Below, d.d. containers with respect to perturbations in the initial state of solutions to retarded differential equations will be considered. Consider the equation

$$dx(s)/ds = g(s, x(\to s)) \text{ a.e. } , x(0) = v, \tag{44}$$

with corresponding "variational inclusion"

$$dq(s)/ds \in D_2 g(s, x(.))(q(\to s)) \text{ a.e. } , q(0) = \hat{v}. \tag{45}$$

Denote the solution of (44) by $x(t, v)$, and let $Q(\hat{v}, x(.)) \subset C(J, X)$ be the set of antidifferentiable solutions $q(.) : J \to X$ of (45). For the conditions on g in the next lemma, if (44) has a solution $x(t, v_0)$ for $v = v_0$, it will also have a solution for v close to v_0. A d.d. container of $v \to x(t, v)$ at v_0 is presented in what follows:

Lemma 11. (Variational inclusion.) Let $g(s, y(.)) : J \times C(J, X) \to X$ have retarded dependence on $y(.) \in C(J, X)$, and assume that $g(s, y(\to s))$ is measurable in s for each $y(.)$. Let V be a given bounded set in X. Assume that (44) has an antidifferentiable solution $x(t) := x(t, v_0)$ on $[0, T]$ for a given element v_0 in V, (so $s \to g(s, x(\to s,\ v_0))$ is integrable by assumption). Assume that $\hat{x}(.) \to g(s, \hat{x}(\to s))$ is Lipschitz continuous in $B(x(., v_0), \varsigma) \subset C(J, X)$ of rank $\leq \kappa(s)$, $\kappa(s)$ integrable. Assume, moreover that, for each s, $\hat{x}(.) \to g(s, x(\to s, v_0) + \hat{x}(\to s))$ has a closed d.d. container $D_2 g(s, x(.,\ v_0))(q^*(.))$ at $\hat{x}(.) = 0$ in all directions $q^*(.) \in C(J, X)$, which is uniform in $q^*(.) \in Q(V_0, x(., v_0)), V_0 := V - v_0$. Assume also that $D_2 g(s, x(., v_0))(q^*(.))$ is Lipschitz continuous in $q^*(.) \in C(J, X)$ of rank $\leq \kappa(s)$, with retarded dependence on $q^*(.)$. Assume finally that $s \to D_2 g(s, x(.,\ v_0))(q^*(.))$ is measurable and essentially separably valued for each $q^*(.)$. Then, for some $\gamma' > 0$, (44) has a solution $x(t, v)$ for all $v \in B(v_0, \gamma')$. Moreover, $Q(t, \hat{v}) := \{q(t) : q(.) \in Q(\hat{v}, x(., v_0))\}$ is a d.d. container of $v \to x(t, v)$ at v_0 in direction $\hat{v}$, uniformly in $\hat{v} \in V_0$.

Proof. Write $x(t) = x(t, v_0)$ and $Q(\hat{v}) = Q(\hat{v}, x(., v_0))(\neq \emptyset$, by Lemma 10). For simplicity, assume $v_0 = 0$. 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 $|v| \leq \varsigma/2e^{\kappa^*} =: \gamma'$, where $\kappa^* := \int_J \kappa(s)ds$. Below, let $\lambda \in (0, \gamma'/|V|]$. Define $z(v, t, \lambda) :=$

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

the norm of the integrand being $\leq \kappa(s)|z(v, ., \lambda)|_s$, where $|y(.)|_s := \sup_{t \leq s} |y(t)|$. Then, by Gronwall's inequality, $|z(v, t, \lambda)| \leq |v|e^{\kappa^*}$, so $|\partial z(v, s, \lambda)/\partial s| \leq \kappa(s)|V|e^{\kappa^*}$. Note that for any solution $q(s, \hat{v}) \in Q(\hat{v})$, $\hat{v} \in V$, the inequality $|dq(s)/ds| \leq \kappa(s)|q(.)|_s$ holds. By Gronwall's inequality, $|q(t)| \leq |\hat{v}|e^{\kappa^*}$, so $|dq(s)/ds| \leq \kappa(s)|V|e^{\kappa^*}$.

Let v be an arbitrary element in V, and let $\varepsilon > 0$. Let $S := \{s_n\}$ be a countable dense set in $\mathbb{R}$, and let $\breve{\alpha}(v, s, \lambda) := \inf\{|z(v, ., \lambda) - q(.)|_s : q(.) \in Q(v)\}$ and $\alpha(v, s, r) := \sup_{\lambda \in S \cap (0,r]} \breve{\alpha}(v, s, \lambda) = \sup_{\lambda \in (0,r]} \breve{\alpha}(v, s, \lambda)$, (the last equality by continuity of $\lambda \to z(v, s, \lambda)$). Then $\alpha(v, 0, r) = 0$. Choose for each n, functions $q_n := q_{n,v,\lambda}(.) \in Q(v)$, such that $|z(v, ., \lambda) - q_n|_{s_n} \leq \breve{\alpha}(v, s_n, \lambda) + \varepsilon/2$. By the bounds on $\partial z(v, s, \lambda)/\partial s$ and $dq_n(s)/ds$, for any s, for some $n =$

n_s, $|z(v,.,\lambda) - q_n|_s \leq \breve{\alpha}(v,s,\lambda) + \varepsilon$, moreover $s \to \breve{\alpha}(v,s,\lambda)$ is measurable, (in fact continuous). Define $\gamma(v,s,\lambda) := \sup_n \mathrm{dist}(a_n(s,\lambda), D_2 g(s,x(.))(q_n))$, where $a_n(s,\lambda) := [g(s,x(.)+\lambda q_n) - g(s,x(.))]/\lambda$. Note that $\lim_{\lambda \searrow 0} \gamma(v,s,\lambda) = 0$, by uniformity of the d.d. container in $Q(v)$. As $D_2 g(s,x(.))(q_n)$ is essentially separably valued, $s \to \gamma(v,s,\lambda)$ is measurable. Because $|a_n(s,\lambda)| \leq \kappa(s)|q_n| \leq \kappa(s)|v|e^{\kappa^*}$, $|D_2 g(s,x(.))(q_n)| \leq \kappa(s)|q_n| \leq \kappa(s)|v|e^{\kappa^*} \leq \kappa(s)|V|e^{\kappa^*}$, and $\gamma(v,s,\lambda) \leq 2\kappa(s)|V|e^{\kappa^*}$, then, by dominated convergence, for some $r > 0$, $\int_J \gamma(v,s,\lambda)ds < \varepsilon$ when $\lambda \leq r$. Below, let $\lambda \leq r$. Then, evidently,
$|[g(s,x(.) + \lambda z(v,.,\lambda)) - g(s,x(.))]/\lambda - [g(s,x(.) + \lambda q_{n_s}) - g(s,x(.))]/\lambda|$
$\leq \kappa(s)(\alpha(v,s,r) + \varepsilon)$. Now,

$$[g(s,x(.) + \lambda q_{n_s}) - g(s,x(.))]/\lambda \in D_2 g(s,x(.))(q_{n_s}) + clB(0,\gamma(v,s,\lambda)),$$

so

$$\partial z(v,s,\lambda)/\partial s = [g(s,x(.) + \lambda z(v,.,\lambda)) - g(s,x(.))]/\lambda$$
$$\in D_2 g(s,x(.))(q_{n_s}) + clB(0,\gamma(v,s,\lambda) + \kappa(s)(\alpha(v,s,r) + \varepsilon)),$$

and $\partial z(v,s,\lambda)/\partial s \in$

$$D_2 g(s,x(.))(z(v,.,\lambda)) + clB(0,\gamma(v,s,\lambda) + 2\kappa(s)(\alpha(v,s,r) + \varepsilon)),$$

by Lipschitz continuity of rank $\kappa(s)$. By Lemma 10, there exists a $q_\lambda(.) \in Q(v)$ such that $|z(v,t,\lambda) - q_\lambda(t)|$
$\leq (\int_0^t 2\{\gamma(v,s,\lambda) + 2\kappa(s)(\alpha(v,s,r) + \varepsilon)\}ds) \exp(\int_0^t 2\kappa(s)ds \leq$
$2K'(\varepsilon + 2\kappa^*\varepsilon) + \int_0^t 4K'\kappa(s)\alpha(v,s,r)ds$,
where $K' = e^{2\kappa^*}$. As $\breve{\alpha}(v,t,\lambda) \leq |z(v,.,\lambda) - q_\lambda(.)|_t$, then $\alpha(v,t,r) \leq 2K'\varepsilon(1 + 2\kappa^*) + \int_0^t 4K'\kappa(s)\alpha(v,s,r)ds$,
and by Gronwall's inequality, $\alpha(v,t,r) \leq \varepsilon K''$ for $K'' :=$
$2K'(1 + 2\kappa^*)\exp(\int_0^T 4K'\kappa(s)ds)$. Hence, $z(v,t,\lambda) \in Q(t,v) + clB(0,\varepsilon K'')$, for $\lambda \in (0,r]$.

A slight extension of this argument yields that an r exists, such that $z(v,t,\lambda) \in Q(t,v) + clB(0,\varepsilon K'')$, uniformly in $v \in V$, when $\lambda \leq r$: By contradiction, assume that, for some $\varepsilon > 0$, for some t, for all natural numbers m, there exist $\lambda_m \in (0,1/m]$ and $v_m \in V$, such that $z(v_m,t,\lambda_m) \notin Q(t,v_m) + clB(0,\varepsilon K'')$. Then, for each m, choose a sequence $q_{n,m} = q_{n,m,\lambda}(.)$, $n = 1,2,...$, such that $|z(v_m,.,\lambda) - q_{n,m}|_{s_n} \leq \breve{\alpha}(v_m,s_n,\lambda) + \varepsilon/2$. Then, define $\gamma(s,\lambda) = \sup_{m,n}\mathrm{dist}([g(s,x(.) + \lambda q_{n,m}) - g(s,x(.))]/\lambda, D_2 g(s,x(s))(q_{m,n}))$. For any $\varepsilon > 0$, for some $r > 0$, dominated convergence and uniformity of the d.d. container in $Q(V)$ yield $\int_J \gamma(s,\lambda)ds \leq \varepsilon$ for $\lambda \leq r$, which implies that for any m, $z(v_m,t,\lambda) \in Q(t,v_m) + clB(0,\varepsilon K'')$ for $\lambda \in (0,r]$, contradicting $z(v_m,t,\lambda_m) \notin Q(t,v_m) + clB(0,\varepsilon K'')$, for m such that $1/m \leq r$. □

The next lemma gives a variational inclusion for the perturbation of a parameter v in a differential equation of the form $dx/ds = g(s,x(.),v)$. By rewriting this system, using an auxiliary state z, governed by $dz/ds =$

$0, z(0) = v$, the next lemma follows immediately from Lemma 11.

Lemma 12. (Variational inclusion.) Let Z be a Banach space, let V be a bounded subset of $Z, x_0 \in X, v_0 \in V$, and let $g(s, x(.), z) : J \times C(J, X) \times Z \to X$ be a given function with retarded dependence on $x(.)$. Denote the solution of the equation $dx/ds = g(s, x(.), v)$ by $x(t, v)$. Assume, for each s, that $(x(.), z) \to g(s, x(.), z)$ has a d.d. container at $(x(., v_0), v_0)$, $D_{2,3}g(s, x(., v_0), v_0)(q^*(.), v^*)$, in all directions $(q^*(.), v^*)$. Write $\check{X} := X \times Z, \check{x} := (x, z) \in \check{X}$, $\check{q}^*(.) = (q^*(.), v^*(.)) \in C(J, \check{X})$, $\check{g}(s, \check{x}(.)) := (g(s, x(\to s), z(s)), 0) : J \times C(J, \check{X}) \to \check{X}$, and $\check{V} := \{x_0\} \times V$, and assume that $\check{g}$, (with corresponding equation $d\check{x}/ds = \check{g}(s, \check{x}(.))), \check{x}(0) = (x_0, v) = \check{v} \in \check{V}$, satisfies the conditions in Lemma 11, with $\check{X}, \check{x}, \check{v}_0 := (x_0, v_0), \check{x}(t, \check{v}) = (x(t, v), v), D_2\check{g}(s, \check{x}(t, \check{v}_0))(\check{q}^*(.)) := (D_{2,3}g(s, x(., v_0), v_0))(q^*(.), v^*(s)), 0)$, and $\check{V}$ playing the roles of $X, x, v_0, x(t, v), D_2g(s, x(., v_0))(q^*(.))$, and V. Then $\hat{Q}(t, \tilde{v})$ is a d.d. container of $v \to x(t, v)$ at v_0, in direction $\tilde{v}$, uniform in $\tilde{v} \in V - v_0$, where $\hat{Q}(t, \tilde{v}) = \{q(t) : q(.)$ is an antidifferentiable solution of $dq/ds \in D_{2,3}g(s, x(., v_0), v_0)(q(.), \tilde{v})$ a.e. $, q(0) = 0\}$. □

A proof of (43), (which was postponed above), can now be given: Let $f, \tilde{f}$,and $\tilde{f}^\delta$ be as in (43). Let $g(s, x(.), v) := v\tilde{f}(s, x(.)) + (1-v)f(s, x(.)), Z := \mathbb{R}$, $V := [0, 1], v_0 = 0$. By E. in Section 8,
$D_{2,3}g(s, x^f(.), 0)(v(.), \hat{v}) = (\tilde{f}(s, x^f(.)) - f(s, x^f(.)))\hat{v} + D_2f(s, x^f(.))(v(.))$.
By Lemma 12, $\delta \to x^{\tilde{f}^\delta}(T)$ has the d.d. container $D_\delta x^{\tilde{f}^0}(T)(1) = Q(T, \tilde{f}, f)$. Then, by C. and B. in Section 8, (43) follows.

Abstract attainability results, abstract necessary conditions.

Below, on product spaces, maximum norms (=maximum of norms) and maximum metrics are used. In the sequel, the following entities are used:

a) Y is a normed space, and C is a nonempty complete subset of Y. A is a complete pseudometric space with pseudometric ρ, and a^* is a given element in A. The function $H(a) : A \to Y$ is continuous.

b) The function $\eta(a) : A \to \mathbb{R}$ is continuous. (46)

Theorem 3. (Attainability) Let the entities in (46) be given, (C will not be used). Let positive numbers $K, \hat{\mu}, \mu', \mu$, $\mu \in (0, 1)$, and an element $\hat{z}^*:= (z^*, \omega^*)$ in $Y \times \mathbb{R}$ be given. Assume that the following properties hold for all $a \in clB(a^*, \hat{\mu})$: For all $\hat{v} := (v, \omega) \in Y \times (-\infty, \omega^*]$ with $|\hat{v} - \hat{z}^*| = \mu'$, for all $r > 0$, a pair $(a', \delta) \in A \times (0, r]$ exists, such that,

$$\begin{aligned} &|H(a') - H(a) - \delta v| \leq (1-\mu)\delta\mu'|\hat{v}|/(|\hat{z}^*| + \mu'), \\ &\eta(a') - \eta(a) - \delta\omega \leq (1-\mu)\delta\mu'|\hat{v}|/(|\hat{z}^*| + \mu'), \text{ and} \\ &\rho(a', a) \leq \delta K|\hat{v}|. \end{aligned} \tag{47}$$

Then, for all $z \in clB(H(a^*), \mu\mu'\hat{\mu}/4K(|\hat{z}^*| + \mu'))$, $\omega \in [\eta(a^*) - \mu\mu'\hat{\mu}/4K(|\hat{z}^*| + \mu'), \eta(a^*)]$, there exists a pair $(a, \alpha) \in clB(a^*, \hat{\mu}\gamma/2) \times [0, \hat{\mu}\gamma/2K(|\hat{z}^*| + \mu')]$, such that $z + \alpha z^* = H(a)$ and $\omega + \alpha\omega^* \geq \eta(a)$, where $\gamma := 4K(|\hat{z}^*| + \mu')\max\{|H(a^*) - z|, \eta(a^*) - \omega\}/\mu\mu'\hat{\mu} \leq 1$.

Proof. The property (47) also holds for $\hat{v}$ in the set $B^* := \{\lambda\tilde{v} : \lambda > 0, \tilde{v} = (v, \omega) \in Y \times \mathbb{R}, |\tilde{v} - \hat{z}^*| = \mu', \omega \leq \omega^*\}$. To see this, let $\hat{v}' := (v', \omega') \in B^*$, $r' > 0$. Then $\hat{v}' = \lambda\hat{v}$ for some $\lambda > 0, \hat{v} = (v, \omega), |\hat{v} - \hat{z}^*| = \mu', \omega \leq \omega^*$. Define $r := \lambda r'$. Now, for all $a \in clB(a^*, \hat{\mu})$, there exists a pair $(a', \delta), \delta \in (0, r]$, such that the inequalities in (47) hold. From these inequalities, for $\delta' := \delta/\lambda \in (0, r/\lambda] = (0, r']$, using $\delta'\hat{v}' = \delta\hat{v}$, it follows that $|H(a') - H(a) - \delta' v'| \leq (1-\mu)\delta'\mu'|\hat{v}'|/(|\hat{z}^*| + \mu'), \eta(a') - \eta(a) - \delta'\omega' \leq (1-\mu)\delta'\mu'|\hat{v}'|/(|\hat{z}^*| + \mu')$, and $\rho(a', a) \leq \delta' K|\hat{v}'|$. Hence, (47) holds for $\hat{v}' \in B^*$.

Below, write $|\hat{z}^*| + \mu' =: \kappa$. The following lemma is needed in the proof:

Lemma 13. Let $z \in clB(H(a^*), \mu\mu'\hat{\mu}/4K\kappa)$, $\omega \in [\eta(a^*) - \mu\mu'\hat{\mu}/4K\kappa, \eta(a^*)]$. Assume that the pair $(a_1, \lambda_1), a_1 \in clB(a^*, \hat{\mu}/2), \lambda_1 \in [-\hat{\mu}/2K\kappa, 0]$ minimizes $(a, \lambda) \to$

$$\max\{|H(a) + \lambda z^* - z|, \eta(a) + \lambda\omega^* - \omega\} + (\mu\mu'/2K\kappa)\max\{\rho(a, a_1), |\lambda - \lambda_1|K\kappa\}$$

in $clB(a^*, \hat{\mu}) \times [-\hat{\mu}/K\kappa, 0]$. Then, $\max\{|H(a_1) + \lambda_1 z^* - z|, \eta(a_1) + \lambda_1\omega^* - \omega\} = \max\{|H(a_1) + \lambda_1 z^* - z|, \max\{0, \eta(a_1) + \lambda_1\omega^* - \omega\}\} = 0$.

Proof of Lemma 13. By contradiction, assume $|\hat{z}| > 0$, $\hat{z} := (z', \omega') := -(H(a_1) + \lambda_1 z^* - z, \max\{0, \eta(a_1) + \lambda_1\omega^* - \omega\})$. The vector $\hat{v} := \hat{z}^* + \mu'\hat{z}/|\hat{z}|$ satisfies $|\hat{v} - \hat{z}^*| = \mu'$, and its second component belongs to $(-\infty, \omega^*]$, so $|\hat{z}|\hat{v} = |\hat{z}|\hat{z}^* + \mu'\hat{z}$ belongs to B^*. Hence, there exist an $a' \in A$, and a $\delta \leq \hat{\mu}/(2K\kappa|\hat{z}|), \delta \in (0, 1/\mu']$, such that $|H(a') - H(a_1) - \delta(|\hat{z}|z^* + \mu' z')| \leq (1-\mu)\delta\mu'|(|\hat{z}|\hat{z}^* + \mu'\hat{z})|/\kappa \leq (1-\mu)\delta\mu'|\hat{z}|$ and $\eta(a') - \eta(a_1) - \delta(|\hat{z}|\omega^* + \mu'\omega') \leq (1-\mu)\delta\mu'|(|\hat{z}|\hat{z}^* + \mu'\hat{z})|/\kappa \leq (1-\mu)\delta\mu'|\hat{z}|$. Moreover, $\rho(a', a_1) \leq \delta K|(|\hat{z}|\hat{z}^* + \mu'\hat{z})| \leq \hat{\mu}/2$, (use the inequality for δ), which implies $a' \in clB(a^*, \hat{\mu})$. Define $\lambda' = \lambda_1 - \delta|\hat{z}| \in [-\hat{\mu}/K\kappa, 0]$, $(\delta|\hat{z}| \leq \hat{\mu}/2K\kappa)$. Then

$|H(a') + \lambda' z^* - z| = | - z + H(a_1) + H(a') - H(a_1) + \lambda' z^*|$
$\leq | - z + H(a_1) + \delta(|\hat{z}|z^* + \mu' z') + \lambda' z^*| + (1-\mu)\delta\mu'|\hat{z}|$
$= | - z + H(a_1) + \lambda_1 z^* + \delta\mu' z'| + (1-\mu)\delta\mu'|\hat{z}|$
$= |-z' + \delta\mu' z'| + (1-\mu)\delta\mu'|\hat{z}| \leq (1-\delta\mu')|\hat{z}| + (1-\mu)\delta\mu'|\hat{z}| = (1-\mu\delta\mu')|\hat{z}|.$
Similarly,
$-\omega + \eta(a') + \lambda'\omega^* = -\omega + \eta(a_1) + \eta(a') - \eta(a_1) + \lambda'\omega^* \leq$
$-\omega + \eta(a_1) + \delta(|\hat{z}|\omega^* + \mu'\omega') + \lambda'\omega^* + (1-\mu)\delta\mu'|\hat{z}| =$
$-\omega + \eta(a_1) + \lambda_1\omega^* + \delta\mu'\omega' + (1-\mu)\delta\mu'|\hat{z}| \leq$
$-\omega' + \delta\mu'\omega' + (1-\mu)\delta\mu'|\hat{z}| \leq (1-\mu\delta\mu')|\hat{z}|.$
Hence,
$\max\{| - z + H(a') + \lambda' z^*|, -\omega + \eta(a') + \lambda'\omega^*\} \leq (1-\mu\delta\mu')|\hat{z}|,$
so, using $\lambda' - \lambda_1 = \delta|\hat{z}|$,
$\max\{| - z + H(a') + \lambda' z^*|, -\omega + \eta(a') + \lambda'\omega^*\} +$
$(\mu\mu'/2K\kappa)\max\{\rho(a', a_1), |\lambda' - \lambda_1|K\kappa\}$
$\leq (1-\mu\delta\mu')|\hat{z}| + (\mu\mu'/2K\kappa)\max\{\delta K|(|\hat{z}|\hat{z}^* + \mu'\hat{z})|, \delta|\hat{z}|K\kappa\} \leq$
$(1-\mu\delta\mu')|\hat{z}| + \mu\delta\mu'|\hat{z}|/2 < |\hat{z}| =$
$|\hat{z}| + (\mu\mu'/2K\kappa)\max\{\rho(a_1, a_1), |\lambda_1 - \lambda_1|K\kappa\},$
a contradiction of the optimality of (a_1, λ_1).

Continued proof of the theorem: Let $\hat{z} := (z, \omega)$, $z \in clB(H(a^*), \mu\mu'\hat{\mu}/4K\kappa)$, $\omega \in [\eta(a^*) - \mu\mu'\hat{\mu}/4K\kappa, \eta(a^*)]$, let γ be as in the conclusion of the theorem, and let $\phi(a, \lambda) := \max\{|H(a) + \lambda z^* - z|, \eta(a) + \lambda\omega^* - \omega\}$. Note that $\phi(a^*, 0) := \max\{|H(a^*) - z|, \eta(a^*) - \omega\} \leq \gamma\mu\mu'\hat{\mu}/4K\kappa$. Let the distance between (a, λ) and (a'', λ'') be $(\mu\mu'/2K\kappa)\max\{\rho(a, a''), |\lambda - \lambda''|K\kappa\}$ in the complete space $clB(a^*, \hat{\mu})\times[-\hat{\mu}/K\kappa, 0]$. By Aubin and Ekeland (1984, Theorem 1, p. 255), (Ekeland's variational principle), there exists a $(a_1, \lambda_1) \in clB(a^*, \hat{\mu})\times[-\hat{\mu}/K\kappa, 0]$ such that
$\phi(a_1, \lambda_1) \leq \phi(a, \lambda) + (\mu\mu'/2K\kappa)\max\{\rho(a, a_1), |\lambda - \lambda_1|K\kappa\}$
for all $(a, \lambda) \in clB(a^*, \hat{\mu})\times[-\hat{\mu}/K\kappa, 0]$ and
$\phi(a_1, \lambda_1) + (\mu\mu'/2K\kappa)\max\{\rho(a_1, a^*), |\lambda_1 - 0|K\kappa\} \leq \phi(a^*, 0)$
$\leq \mu\mu'\hat{\mu}\gamma/4K\kappa$, which gives $\rho(a_1, a^*) \leq \hat{\mu}\gamma/2, |\lambda_1| \leq \hat{\mu}\gamma/2K\kappa$.
By Lemma 13, $| - z + H(a_1) + \lambda_1 z^*| = 0$ and $-\omega + \eta(a_1) + \lambda_1\omega^* \leq 0$, so $z + \alpha z^* = H(a_1)$ and $\omega + \alpha\omega^* \geq \eta(a_1)$, for $\alpha = -\lambda_1 \in [0, \hat{\mu}\gamma/2K\kappa]$ and the proof is finished.

Corollary 2. (Attainability. Theorem 3 with $\eta(.)$ deleted) Let the entities in (46),a) be given, (C is not used). Let positive numbers $K, \hat{\mu}, \mu', \mu, \mu \in (0,1)$, and an element $z^* \in Y$ be given. Assume that the following properties hold for all $a \in clB(a^*, \hat{\mu})$:

For all $v \in Y$, with $|v - z^*| = \mu'$, for all $r > 0$, a pair $(a', \delta) \in A \times (0, r]$ exists, such that $|H(a') - H(a) - \delta v| \leq (1 - \mu)\delta\mu'|v|/(|z^*| + \mu')$, and $\rho(a', a) \leq \delta K|v|$. (48)

Then for all $z \in clB(H(a^*), \mu\mu'\hat{\mu}/4K(|z^*| + \mu'))$, there exists a pair $(a, \alpha) \in clB(a^*, \hat{\mu}\gamma/2) \times [0, \hat{\mu}\gamma/2K(|z^*| + \mu')]$, such that $z + \alpha z^* = H(a)$, where $\gamma := 4K(|z^*| + \mu')|H(a^*) - z|/\mu\mu'\hat{\mu}$.

Proof: The proof of this result can be obtained by deleting all arguments pertaining to $\eta(.)$ in the proof of Theorem 3, (the function minimized in such a proof will then be $|H(a) + \lambda z^* - z| + [\mu\mu'/2K\kappa]\max\{\rho(a, a_1), |\lambda - \lambda_1|K\kappa\}, \kappa = |z^*| + \mu')$. The present result can be seen to follow from Theorem 3 in a formal manner, and let us indicate this, perhaps not very exciting fact. First, observe that (48) in fact holds for all $v \in clB(z^*, \mu')$. (If $v = 0$, put $a' = a$, if not, for some $\lambda \in [1, \infty)$, $|\lambda v - z^*| = \mu'$, since $|v - z^*| \leq \mu', |\infty v - z^*| = \infty$. Then see the argument in the beginning of the proof of Theorem 3.) The corollary then follows from Theorem 3, for $\omega^* = 0$, by replacing a^* by $(a^*, 0)$ and A by $A \times \mathbb{R}$, and defining $\eta(a, \beta)$ for (a, β) in this set by $\eta(a, \beta) := \beta/K$ and defining H on the product set by $H(a, \beta) := H(a)$. To show (47) for $a \in clB(a^*, \hat{\mu})$, $\beta \in [-\hat{\mu}, \hat{\mu}]$, and for any $\hat{v} := (v, \omega)$ as in (47), take the (a', δ) furnished for this v by (48) (just generalized), and let (a', β') play the role of a' in (47), where $\beta' := \beta - \delta K|\hat{v}|$. (Note that $\eta(a', \beta') - \eta(a, \beta) - \delta\omega = (-\delta K|\hat{v}|)/K - \delta\omega \leq 0$.)

Corollary 3. (Exact penalization). Let the entities in (46) be given, (C is not used). Assume that a^* is optimal in the problem: $\min_a \eta(a)$, subject to $a \in A, H(a) = 0$. Assume that for all $a \in A$, for some number $W_a, |\eta(a') - \eta(a)| \leq W_a\rho(a', a)$ for all $a' \in A$. Moreover, assume that there exists a quadruple $(K, \mu, \hat{\mu}, \mu'), K > 0, \mu \in (0, 1)$, $\hat{\mu} > 0$, $\mu' > 0$, such that for all $a \in clB(a^*, \hat{\mu})$, all $v \in Y$ with $|v| = \mu'$, all $r > 0$, a pair (a', δ) $\in A \times (0, r]$ exists, such that,

$$|H(a') - H(a) - \delta v| \leq (1 - \mu)\delta\mu',$$
$$\rho(a', a) \leq \delta K\mu'. \tag{49}$$

Then a^* minimizes $a \to \phi(a) := \eta(a) + 4\ W_a K|H(a)|/\mu$
for a in $clB(a^*, \hat{\mu}/2)$.

Proof. Assume by contradiction that for some $a_* \in clB(a^*, \hat{\mu}/2), \phi(a_*) < \phi(a^*) = \eta(a^*)$. By "$\eta$-optimality" of a^*, $H(a_*) \neq 0$. Now, $\phi(a_*) < \phi(a^*)$

implies (the third of the inequalities) $W_{a_*}\ \hat{\mu}/2 \geq W_{a_*}\rho(a_*, a^*) \geq$

$$\eta(a^*) - \eta(a_*) > 4W_{a_*}K|H(a_*)|/\mu\ , \tag{50}$$

so $(\hat{\mu}/2)\mu/4K \geq |H(a_*)|$ and $0 \in clB(H(a_*), (\hat{\mu}/2)\mu/4K)$. Corollary 2 is now applied for a_* and $\hat{\mu}/2$ playing the roles of a^* and $\hat{\mu}$; as $a_* \in clB(a^*, \hat{\mu}/2)$, note that, by (49), (48) holds for $z^* = 0$ for all $a \in clB(a_*, \hat{\mu}/2) \subset clB(a^*, \hat{\mu})$. Hence, 0 is attainable, i.e. for some $a \in clB(a_*, (\hat{\mu}/2)\gamma/2), H(a) = 0$, where $\gamma = 4K|H(a_*)|/(\hat{\mu}/2)\mu$, so $\rho(a_*, a) \leq (\hat{\mu}/2)\gamma/2 \leq 2K|H(a_*)|/\mu$. Then, by the last inequality and (50), $\eta(a) = \eta(a) - \eta(a_*) + \eta(a_*) - \eta(a^*) + \eta(a^*) < W_{a_*}2K|H(a_*)|/\mu - 4W_{a_*}K|H(a_*)|/\mu + \eta(a^*) < \eta(a^*)$, a contradiction of the "η-optimality" of a^*.

Corollary 4 (Attainability) Let the entities of (46) be given. Assume that there exists a collection $(K, c^*, z^*, \omega^*, \mu'', \hat{\mu}, \mu'), K > 0, c^* \in C, \hat{z}^* := (z^*, \omega^*) \in Y \times \mathbb{R}, \mu'' \in (0, 1), \hat{\mu} > 0, \mu' > 0$, with the property that for all $(a, c) \in clB(a^*, \hat{\mu}) \times (C \cap clB(c^*, \hat{\mu}))$, all $\hat{v} := (v, \omega) \in Y \times (-\infty, \omega^*]$ with $|\hat{v} - \hat{z}^*| = \mu'$, all $r \in (0, 1]$, a triple $(a', c', \delta) \in A \times C \times (0, r]$ exists, such that,

$$\begin{aligned} &|H(a') - H(a) - \delta v - (c' - c)| \leq (1 - \mu'')\delta\mu'|\hat{v}|/(|\hat{z}^*| + \mu'), \\ &\quad \rho(a', a) \leq \delta K|\hat{v}|, |c' - c| \leq \delta K|\hat{v}|, \text{ and} \\ &\quad \eta(a') - \eta(a) - \delta\omega \leq (1 - \mu'')\delta\mu'|\hat{v}|/(|\hat{z}^*| + \mu'). \end{aligned} \tag{51}$$

Then, for any $\theta \in [1 - \mu''\mu'/2K(|\hat{z}^*| + \mu'), 1]$, for all pairs (z, ω),
$z \in clB(H(a^*) - \theta c^*, \mu''\mu'\hat{\mu}/8K(|\hat{z}^*| + \mu'))$,
$\omega \in [-\mu''\mu'\hat{\mu}/8K(|\hat{z}^*| + \mu') + \eta(a^*), \eta(a^*)]$,
there exists a triple (a, c, α) in
$clB(a^*, \hat{\mu}\lambda/2) \times (C \cap clB(c^*, \hat{\mu}\lambda/2)) \times [0, \hat{\mu}\lambda/2K(|\hat{z}^*| + \mu')]$,
such that $z + \alpha z^* = H(a) - \theta c$ and $\omega + \alpha\omega^* \geq \eta(a)$, where
$\lambda = 8K(|\hat{z}^*| + \mu')\max\{|H(a^*) - \theta c^* - z|, \eta(a^*) - \omega\}/\mu''\mu'\hat{\mu}$.

Proof: Let $\mu = \mu''/2$. Note that, for any $\theta \in [1 - \mu''\mu'/2K(|\hat{z}^*| + \mu'), 1]$ and c', c as in (51), $|(c' - c) - \theta(c' - c)| \leq (1 - \theta)\delta K|\hat{v}| \leq [\mu''\mu'/2K(|\hat{z}^*| + \mu')]\delta K|\hat{v}| \leq \mu''\mu'\delta|\hat{v}|/2(|\hat{z}^*| + \mu')$. Hence, replacing $c' - c$ by $\theta(c' - c)$ in the first inequality in (51) yields:

$$\begin{aligned} &|H(a') - H(a) - \delta v - \theta(c' - c)| \leq (1 - \mu)\delta\mu'|\hat{v}|/(\ |\hat{z}^*| + \mu'), \\ &\rho(a', a) \leq \delta K|\hat{v}|, |c' - c| \leq \delta K|\hat{v}|\ , \text{ and} \\ &\eta(a') - \eta(a) - \delta\omega \leq (1 - \mu)\delta\mu'|\hat{v}|/(|z^*| + \mu'). \end{aligned} \tag{52}$$

Write $A^* := A \times C$ and $H^*(a,c) := H(a) - \theta c$. Let $A^*, H^*, (a^*, H(a^*)), (a,c)$ and (a', c') play the roles of A, H, a^*, a, and a' in the premisses of Theorem 3. The conclusion of Corollary 4 then follows from the conclusion of Theorem 3.

Corollary 5. (Necessary condition for optimality) Let the entities of (46) be given, with C convex. Assume that a^* is optimal in the problem: $\min_a \eta(a)$ subject to $a \in A, H(a) \in C$. Then there exists no quintuple $(K, c^{**}, \mu'', \hat{\mu}, \mu')$, $K > 0, c^{**} \in C - H(a^*), \mu'' \in (0,1), \hat{\mu} > 0, \mu' > 0$, with the property that for all $(a,c) \in clB(a^*, \hat{\mu}) \times (C \cap clB(H(a^*), \hat{\mu}))$, all $\hat{v} := (v, \omega) \in Y \times (-\infty, 0]$ with $|\hat{v} - (c^{**}, 0)| = \mu'$, all $r \in (0,1]$, a triple $(a', c'', \delta) \in A \times C \times (0, r]$ exists, such that

$$
\begin{aligned}
&|H(a') - H(a) - \delta(v + c'' - c)| \leq (1 - \mu'')\delta\mu'|\hat{v}|/(|c^{**}| + \mu'),\\
&\rho(a', a) \leq \delta K|\hat{v}|, |c'' - c| \leq K|\hat{v}| \text{ , and}\\
&\eta(a') - \eta(a) - \delta\omega \leq (1 - \mu'')\delta\mu'|\hat{v}|/(|c^{**}| + \mu'). \qquad (53)
\end{aligned}
$$

This corollary says that, for (a,c) near $(a^*, H(a^*))$, if $H(a) - c$ can be moved roughly in all directions by perturbation of a and c chosen from A and C, then the corresponding changes in $\eta(.)$ cannot be "too negative".

Proof. Condition (53) implies (51), when $\omega^* = 0, c^* := H(a^*), z^* = c^{**}, c' = \delta c'' + (1 - \delta)c$. Then, assume by contradiction that (53) holds. Let $\kappa := |c^{**}| + \mu', \theta := 1 - \mu''\mu'/2K\kappa$. By shrinking μ'' if necessary, it can be assumed that $\theta \geq 0$, with (53) still holding. Let $\varepsilon \in (0, \mu''\mu'/4K\kappa]$ be so small that $\varepsilon c^{**} \in clB(0, \mu''\mu'\hat{\mu}/8K\kappa)$, and that $\lambda := (8K\kappa/\mu''\mu'\hat{\mu})\varepsilon|c^{**}|$ satisfies $\lambda\hat{\mu} \leq \mu''\mu'/2$. Then, by Corollary 4, for $z = c^* - \theta c^* + \varepsilon c^{**}, \omega = \eta(a^*) - \varepsilon|c^{**}|$, for some $\alpha \in [0, \lambda\hat{\mu}/2K\kappa] \subset [0, \mu''\mu'/4K\kappa]$, and for some $a \in A, c \in C$, we have that $c^* - \theta c^* + \varepsilon c^{**} + \alpha c^{**} = z + \alpha c^{**} = H(a) - \theta c$ and $\eta(a^*) - \varepsilon|c^{**}| = \omega + \alpha \cdot 0 \geq \eta(a)$. Write $c^{**} = \tilde{c}^{**} - c^*, \tilde{c}^{**} \in C$, and note that $\theta + \varepsilon + \alpha \leq 1 - \mu''\mu'/2K\kappa + \mu''\mu'/4K\kappa + \mu''\mu'/4K\kappa \leq 1$. Then, $H(a) = c^* - \theta c^* + (\varepsilon + \alpha)c^{**} + \theta c = (1 - \theta - \varepsilon - \alpha)c^* + \theta c + (\varepsilon + \alpha)\tilde{c}^{**} \in C$, and $\eta(a^*) > \eta(a)$, contradicting the optimality of a^*.

Corollary 6. (Exact penalization) Let the entities in (46) be given. Assume that a^* is optimal in the problem: $\min_a \eta(a)$, subject to $a \in A, H(a) \in C$. Assume that for all $a \in A$, for some number $W_a, |\eta(a') - \eta(a)| \leq W_a\rho(a', a)$ for all $a' \in A$. Moreover, assume that there exists a quadruple $(K, \mu, \hat{\mu}, \mu'), K > 0, \mu \in (0,1), \hat{\mu} > 0, \mu' > 0$, such that for all $(a,c) \in clB(a^*, \hat{\mu}) \times (C \cap clB(H(a^*), \hat{\mu}))$, all $v \in Y$ with $|v| = \mu'$, all $r \in (0,1]$, a triple $(a', c', \delta) \in A \times C \times (0, r]$ exists, such that,

$$
\begin{aligned}
&\text{(i)}|H(a') - H(a) - \delta v - (c' - c)| \le (1-\mu)\delta\mu', \\
&\text{(ii) } \rho(a', a) \le \delta K\mu', \text{ and } |c' - c)| \le \delta K\mu'.
\end{aligned} \tag{54}
$$

Then $(a^*, H(a^*))$ minimizes $(a,c) \to \phi(a,c)$ in $clB(a^*, \hat{\mu}/2)\times$ $(C \cap clB(H(a^*), \hat{\mu}/2))$, where

$$\phi(a,c) := \eta(a) + 4W_a K|H(a) - c|/\mu.$$

Proof. Let $A^* := A \times C$, and $H^*(a,c) = H(a) - c$. Let $A^*, H^*, (a^*, H(a^*))$, (a,c) and (a',c') play the roles of A, H, a^*, a, and a' in Corollary 3. The conclusion of Corollary 6 then follows from the conclusion of Corollary 3.

The next corollary shows that a modification of (54) works, in case C is convex.

Corollary 7. (Exact penalization) Let the entities in (46) be given with C convex. Assume that a^* is optimal in the problem: $\min_a \eta(a)$, subject to $a \in A, H(a) \in C$. Assume that for all $a \in A$, for some number $W_a, |\eta(a') - \eta(a)| \le W_a \rho(a', a)$ for all $a' \in A$. Assume also that there exists a quadruple $(K', \mu'', \mu^*, \mu'), K' > 0, \mu'' \in (0,1)$, $\mu^* > 0, \mu' > 0$, such that for all $a \in clB(a^*, \mu^*)$, all $v \in Y$ with $|v| = \mu'$, all $r \in (0,1]$, a triple (a', c'', δ) $\in A\times$ $C \times (0, r]$ exists, such that,

$$
\begin{aligned}
&\text{(i) } |H(a') - H(a) - \delta v - \delta(c'' - H(a^*))| \le (1-\mu'')\delta\mu', \\
&\text{(ii) } \rho(a', a) \le \delta K'\mu', \text{ and } |c'' - H(a^*)| \le K'\mu'.
\end{aligned} \tag{54'}
$$

Then, $(a^*, H(a^*))$ minimizes $(a,c) \to \check{\phi}(a,c)$ in $clB(a^*, \hat{\mu}/2)\times$ $(C \cap clB(H(a^*), \hat{\mu}/2))$, where

$$
\begin{aligned}
&\check{\phi}(a,c) := \eta(a) + 8W_a K|H(a) - c|/\mu'', \\
&\hat{\mu} := \min\{\mu^*, \mu''\mu'/2\}, \text{ and } K := \max\{K', 1 + K'\mu'/\hat{\mu}\}.
\end{aligned}
$$

Proof. Let $c^* := H(a^*)$. It suffices to show that (54') implies (54) for K and $\hat{\mu}$ as just defined, with $\mu{:=}\mu''/2$ and $c' = \delta c'' + (1-\delta)c$, where c is arbitrary in $C \cap clB(c^*, \hat{\mu})$. Note that when c^* is replaced by c in (54')(i), then $\delta\hat{\mu}$ has to be added on the right hand side, which yields (54)(i), because $\delta\hat{\mu} \le \delta\mu''\mu'/2$. Moreover, the inequalities $|c - c^*| \le \hat{\mu}, |c'' - c^*| \le K'\mu'$, (see (54')(ii)) imply $|c'' - c| \le K'\mu' + \hat{\mu} \le K\hat{\mu} \le K\mu''\mu'/2 \le K\mu'$, and (54)(ii) follows.

11. Proofs of Remarks 2 and 3 and Corollary 1

Proof of Remark 2 a. A proof is only given for systems of the type

(37). Given any quintuple $(f, v, \tilde{f}, c'', \tilde{\delta})$ such that the version of (15) stated in Remark 2 a. holds. Let $q(.) \in Q(\tilde{f}, f)$ and observe that $|q(.)| \leq \tilde{\delta}K$, where $K = 2M \exp(TM^{f^*})$. Choose $\hat{\mu} \in (0, \hat{\mu}']$ such that $\breve{\mu}\mu'/2 \geq 2\hat{\mu}KM^{f^*} \exp(2TM^{f^*})$. Let $f \in clB(f^*, \hat{\mu})$. Note that $Df^*(s, x^f(.))$ differs from $Df(s, x^f(.))$ only on C_f, so $dq(t)/dt \in Df^*(t, x^f(.))(q(.))+$
$clB(0, 1_{C_f} M^{f^*}\tilde{\delta}K) + \tilde{f}(t, x^f(.)) - f(t, x^f(.)), (M^f = M^{f^*})$. By Lemma 10, there exists an antidifferentiable $q^*(.)$, with
$dq^*(t)/dt \in Df^*(t, x^f(.))(q^*(.)) + \tilde{f}(t, x^f(.)) - f(t, x^f(.))$ a.e. such that $|q(T) - q^*(T)| \leq 2(\int_0^T \tilde{\delta}K 1_{C_f} M^{f^*} dt) \exp(2TM^{f^*}) \leq \tilde{\delta}\breve{\mu}\mu'/2$. Hence, (15) follows for $\mu = \breve{\mu}/2$.

Proof of Corollary 1.

It is only needed to consider systems of the form (37), (ordinary retarded equations). In this case, $q^{\tilde{f},f}(s)(t)$ is actually independent of t, we write $q_{\tilde{f},f}(s) = q^{\tilde{f},f}(s)(s)$. Define

$$K^f := cl\{G_x(x^f(T))q_{\tilde{f},f}(T) - c + G(x^*(T)) :\\ c \in C \cap clB(G(x^*(T)), 1), \tilde{f} \in coF\}. \tag{55}$$

By (19), K^{f^*} is a closed convex body. If the origin $\mathbf{0}$ in Y is not an interior point of K^{f^*}, then, for some nonzero continuous linear functional $\lambda^*, \langle K^{f^*}, \lambda^* \rangle \leq \langle \mathbf{0}, \lambda^* \rangle$. With $\lambda_0 = 0$, then (20) holds. So let us consider the nontrivial case where $\mathbf{0}$ is an interior point of K^{f^*}. Then for some $\kappa > 0, -\kappa z^* \in K^{f^*}$. Consider the set $B_z := co\{z, clB(z^*, \varepsilon)\}$. Evidently, $\mathbf{0}$ is an interior point in B_z if $z = -\kappa z^*$, and even if the equality $z = -\kappa y^*$ is only an approximate one. In fact, there exist numbers $\rho > 0$ and $\xi > 0$, such that $clB(\mathbf{0}, \xi) \subset B_z$ for all $z \in clB(-\kappa z^*, \rho)$. Because $-\kappa z^* \in K^{f^*}$, there exist a $\tilde{f} \in coF$, and a $c \in (C \cap clB(G(x^*(T), 1)) - G(x^*(T))$, such that $|-\kappa z^* - (G_x(x^*(T))q_{\tilde{f},f^*}(T) - c)| < \rho/4$. For some $\beta \in (0, \hat{\mu}]$, small enough, for $f \in clB(f^*, \beta), |G_x(x^*(T))q_{\tilde{f},f^*}(T) - G_x(x^f(T))q_{\tilde{f},f^*}(T)| < \rho/4$, by the continuity assumption on G_x in Corollary 1. The two inequalities involving $\rho/4$ yield $|-\kappa z^* - (G_x(x^f(T))q_{\tilde{f},f^*}(T) - c)| < \rho/2$, for all $f \in clB(f^*, \beta)$. By shrinking β, if necessary, we shall show below that, when $f \in clB(f^*, \beta)$,

$$|G_x(x^f(T))q_{\tilde{f},f}(T) - G_x(x^f(T))q_{\tilde{f},f^*}(T)| \leq \rho/2. \tag{56}$$

Assume for the moment that (56) holds. From the two inequalities involving $\rho/2$, for $z := G_x(x^f(T))q_{\tilde{f},f}(T) - c$, it follows that $z \in B(-\kappa z^*, \rho)$ for all $f \in clB(f^*, \beta)$, hence, by (19), $B_z := co\{z, clB(z^*, \varepsilon)\} \subset co\{z, K^f\} \subset K^f$, (since $z \in K^f$ and K^f is convex). Thus, $B(\mathbf{0}, \xi) \subset K^f, f \in clB(f^*, \beta)$. Then, (15) is satisfied for $\mu' = \xi/2$ and $\mu = 1/2$ (say), and $\hat{\mu}$ replaced by β, and the conclusion of Theorem 1 holds. The inequality $0 \geq \inf \Omega_{c,\tilde{f}} = \phi_x(x^*(T))q_{\tilde{f},f^*}(T) - \Lambda|G_x(x^*(T))q_{\tilde{f},f^*}(T) - c + G(x^*(T))|$ (for all $\tilde{f} \in coF, c \in C$), implies that the open convex body $L := \{(w, v) \in \mathbb{R} \times X, w - \Lambda|v| > 0\}$ is disjoint from the convex set $D := \{(\phi_x(x^*(T))q_{\tilde{f},f^*}(T), G_x(x^*(T))q_{\tilde{f},f^*}(T) - c : \tilde{f} \in coF, c \in C - G(x^*(T))\}$. Thus, L and D can be separated by a nonzero continuous linear functional, in fact for some $(\lambda_0, \lambda^*) \in \mathbb{R} \times Y^*, (\lambda_0, \lambda^*) \neq 0$, the inequality $\sup\langle D, (\lambda_0, \lambda^*)\rangle \leq \inf\langle L, (\lambda_0, \lambda^*)\rangle$ holds, which yields $\lambda_0 \geq 0$ and, generally, the conclusion in Corollary 1.

Let us now prove (56): Define $M' := |q_{\tilde{f},f^*}(.)|$. Choose γ' such that $\gamma' \exp(M^*T) = \rho/8M_G$, choose $\gamma \in (0, \zeta)$ so small that
$2M\gamma \exp(M^*T) \leq \rho/8M_G, \gamma'' := 2M^*M'\gamma$ satisfies $\gamma'' \exp(M^*T) \leq \rho/8M_G$, $\gamma''' := 2MM^*\gamma \exp(M^*T)$ satisfies $\gamma''' \exp(M^*T) \leq \rho/8M_G$, and such that

$$\int_0^T |f_x^*(\tau, x^f(\to \tau))q_{\tilde{f},f^*}(\to \tau) - f_x^*(\tau, x^*(\to \tau))q_{\tilde{f},f^*}(\to \tau)|d\tau < \gamma' \quad (56')$$

when $f \in clB(f^*, \gamma)$, (for each τ, the integrand converges to zero, when $f \to f^*$, by (24)(i), moreover, dominated convergence and the bound M^{f^*} on f_x^* is also used). Let f be any element in F and let C_f be a set of measure $\sigma(f, f^*)$ such that $\{s : f(s, y(.)) \neq f^*(s, y(.))\} \subset C_f$ for all $y(.)$ in $C(J, X)$. Note that if $\sigma(f, f^*) \leq \gamma$, then
$\int_0^t |[f_x(t, x^f(\to \tau)) - f_x^*(\tau, x^f(\to \tau))]q_{\tilde{f},f^*}(\to \tau)|d\tau \leq 2M^*M'\text{meas}(C_f) = 2M^*M'\sigma(f, f^*) \leq \gamma''$ and $|x^f(.) - x^{f^*}(.)| \leq 2M\gamma \exp(TM^*)$, see (18).

Then, by (18), and (24)i), for $f \in clB(f^*, \gamma)$, $|q_{\tilde{f},f}(t) - q_{\tilde{f},f^*}(t)| \leq$

$$\begin{aligned}
&|\tilde{f}(\tau, x^f(.)) - \tilde{f}(\tau, x^{f^*}(.)) - f(\tau, x^f(.)) + f^*(\tau, x^{f^*}(.))| + \\
&\int_0^t |f_x^*(\tau, x^f(\to \tau))q_{\tilde{f},f}(\to \tau) - f_x^*(\tau, x^{f^*}(\to \tau))q_{\tilde{f},f^*}(\to \tau)|d\tau \leq \\
&M^*|x^f(.) - x^{f^*}(.)| + 2M\sigma(f, f^*) + \\
&\int_0^t |f_x^*(\tau, x^f(\to \tau))q_{\tilde{f},f}(\to \tau) - f_x^*(\tau, x^f(\to \tau))q_{\tilde{f},f^*}(\to \tau)|d\tau + \\
&\int_0^t |f_x^*(\tau, x^f(\to \tau))q_{\tilde{f},f^*}(\to \tau) - f_x^*(\tau, x^{f^*}(\to \tau))q_{\tilde{f},f^*}(\to \tau)|d\tau \leq \\
&\gamma''' + 2M\gamma + \int_0^t M^*|q_{\tilde{f},f}(.) - q_{\tilde{f},f^*}(.)|_\tau d\tau + \gamma'' + \gamma'.
\end{aligned}$$

Even $|q_{\tilde{f},f}(.) - q_{\tilde{f},f^*}(.)|_t \leq \int_0^t M^*|q_{\tilde{f},f}(.) - q_{\tilde{f},f^*}(.)|_\tau d\tau + \gamma''' + 2M\gamma + \gamma'' + \gamma'$. Hence, by Gronwall's inequality, $|q_{\tilde{f},f}(.) - q_{\tilde{f},f^*}(.)|_t \leq (\gamma''' + 2M\gamma + \gamma'' + \gamma') \exp(M^*t)$, and

$$|q_{\tilde{f},f}(T) - q_{\tilde{f},f^*}(T)| \leq \rho/2M_G, \text{ for any } f \text{ such that } \sigma(f, f^*) \leq \gamma. \quad (57)$$

Then, (56) follows, for $\beta \leq \gamma$.

A proof of Remark 3 is obtained by observing that in this case (56) can be obtained for $\rho = \varepsilon$, for any $\tilde{f} \in coF$, ((56') now holds for any $\tilde{f} \in coF$, for $\beta \in (0, \zeta]$ suitably chosen). Assuming that (19) holds for $f = f^*$, then this version of (56) implies that (19) holds for $\hat{\mu} = \beta$, (i.e. all $f \in clB(f^*, \beta)$), for ε replaced by $\varepsilon/2$.

To obtain a proof in the case of Remark 2. b., K^f must be redefined to equal $clco\{\hat{\delta}^{-1}G_x(x^f(T))q_{\hat{f},f}(T) - c + G(x^*(T)) : \hat{\delta} > 0, \hat{f} \in B(f^*, \hat{\delta}), c \in C \cap clB(G(x^*(T)), 1)\}$. In the nontrivial case where $0 \in \mathrm{int}K^{f^*}$, for some finite collection $(\tilde{f}_i,\tilde{\delta}_i, c_i, \lambda_i)$, where $\lambda_i > 0, \sum \lambda_i = 1, \tilde{\delta}_i > 0, \tilde{f}_i \in B(f^*, \tilde{\delta}_i)$, $c_i \in C \cap clB(G(x^*(T)), 1) - G(x^*(T))$, the inequality $|-\kappa z^* - \sum_i \lambda_i[\tilde{\delta}_i^{-1}(G_x(x^*(T))q_{\tilde{f}_i,f^*}(T) - c_i]| < \rho/4$ holds. Again, by the continuity assumption on G_x, for β small enough, $|-\kappa z^* - \sum_i \lambda_i[\tilde{\delta}_i^{-1}(G_x(x^f(T))q_{\tilde{f}_i,f^*}(T) - c_i]| < \rho/2$ for all $f \in clB(f^*, \beta)$, and $\sigma(\tilde{f}_i, f^*) + \beta < \tilde{\delta}_i$ for all $i, (\Rightarrow \sigma(\tilde{f}_i, f) < \tilde{\delta}_i)$, for $f \in clB(f^*, \beta))$. Evidently, for $\rho/2$ replaced by $\tilde{\delta}_i\rho/2$ and $\tilde{f}$ by $\tilde{f}_i$, (56) holds for small β, for each i. Thus, by shrinking β, for $f \in clB(f^*, \beta))$, the inequality $|-\kappa z^* - \sum_i \lambda_i[\tilde{\delta}_i^{-1}(G_x(x^f(T))q_{\tilde{f}_i,f}(T) - c_i]| < \rho$, can be obtained. Hence, as before, for $f \in clB(f^*, \beta)$, $B(\mathbf{0},\xi) \subset K^f$. This inclusion yields that for any $v \in Y$ with $|v| = \mu' := \xi/2$, for any $f \in clB(f^*, \beta)$, for some collection $(\tilde{f}_i,\tilde{\delta}_i, c_i, \lambda_i), \lambda_i > 0, \sum \lambda_i = 1, \tilde{\delta}_i > 0, \tilde{f}_i \in B(f, \tilde{\delta}_i)$, we have $|v - \sum_i \lambda_i[\tilde{\delta}_i^{-1}(G_x(x^f(T))q_{\tilde{f}_i,f}(T) - c_i]| \leq \mu'/2$. Define $\varkappa = \sum_i \lambda_i\tilde{\delta}_i^{-1}$ and $\tilde{C}^i$ to be measurable sets such that meas $(\tilde{C}^i) = \sigma(\tilde{f}_i, f)$ and $\tilde{C}^i \supset \{s : \tilde{f}_i(s, x(.)) \neq f(s, x(.))\}$ for all continuous $x(.)$. There exist disjoint measurable sets C_i, (stemming from a rapid switching between the $\tilde{f}_i$'s with weights $\lambda_i\tilde{\delta}_i^{-1}/\varkappa$), such that $|\sum_i \lambda_i[\tilde{\delta}_i^{-1}(G_x(x^f(T))q_{\tilde{f}_i,f}(T)-c_i]-\tilde{\delta}^{-1}G_x(x^f(T))q_{\tilde{f},f}(T)+\sum_i \lambda_i c_i| \leq \mu'/4$, where $\tilde{\delta} := \varkappa^{-1}$, $\tilde{f} := \sum_i 1_{C_i}\tilde{f}_i$, and $\int_J 1_{C_i}1_{\tilde{C}^i} < (\lambda_i\tilde{\delta}_i^{-1}/\varkappa)\tilde{\delta}_i = \lambda_i/\varkappa$. As $C_{\tilde{f}} \subset \cup_i C_i \cap \tilde{C}^i$, $\sigma(\tilde{f}, f) < \sum_i \lambda_i/\varkappa = 1/\varkappa = \tilde{\delta}$. Moreover, by the inequalities involving $\mu'/2$ and $\mu'/4$, $|v - \tilde{\delta}^{-1}G_x(x^f(T))q_{\tilde{f},f}(T) + \sum_i \lambda_i c_i| \leq 3\mu'/4$, so (15) holds for $\mu = 1/4, \mu' = \xi/2$.

Remark 5. In the context of Theorem 1, it seems possible to do without the uniformity requirements in (10) and (11), but only at the expence of a more complicated controllability requirement. Assume in this remark that the uniformity requirements on the d.d. containers D_3f, DG, and $D\phi$ of (10) and (11) are removed, (with this change, (6) - (11) are postulated). Then, (15) has to be replaced by the following condition: For any $f \in clB(f^*, \hat{\mu})$, for any $(v, y(.)) \in Y \times L_1(J, C(J, X))$,with $|(v, y(.))| = \mu'$, a quadruple $(\tilde{f}, c'', \tilde{\delta}, z(.))$ exists, $\tilde{f} \in coB(f, \tilde{\delta})$, $c'' \in C \cap clB(G(x^*(T)), 1), \tilde{\delta} > 0, z(.) \in clB(0,1) \subset L_1(J, C(J, X))$, such that

$$\max\{\sup_{\check{y} \in \nabla^{G,f}(z(.))} |\tilde{\delta}v - \check{y} + \tilde{\delta}[c'' - c^*]|, \sup_{\hat{z}(.) \in \nabla^f(z(.), \tilde{f})} |\tilde{\delta}y(.) - \hat{z}(.)|\}$$
$$\leq (1-\mu)\tilde{\delta}\mu',$$

where $c^* := G(x^*(T))$, $\nabla^{\phi,f}(z(.)) := D\phi(x^f(T))(\int_J z(.)), \nabla^{G,f}(z(.)) := DG(x^f(T))(\int_J z(.)), \nabla^f(z(.), \tilde{f}) := \{\hat{z}(.) \in L_1(J, C(J, X)) : \hat{z}(s) \in z(s) - D_3 f(., s, x^f(.))(\int_0^s z(\tau)d\tau) - (\tilde{f}(., s, x^f(.)) - f(., s, x^f(.)))$ a.e.$\}$.
In this case, the following modification of the necessary condition in Theorem 1 presumably holds: For some positive constants Λ' and Λ'', for all $c \in C, \tilde{f} \in coF, z(.) \in L_1(J, C(J, X))$, we have $0 \geq \inf \tilde{\Omega}_{\tilde{f},c,z(.)}$, where $\tilde{\Omega}_{\tilde{f},c,z(.)} :=$

$$\{w - \Lambda'|v - c + c^*| - \Lambda''|\hat{z}(.)| : w \in \nabla^{\phi,f^*}(z(.)), v \in \nabla^{G,f^*}(z(.)),$$
$$\hat{z}(.) \in \nabla^{f^*}(z(.), \tilde{f})\}.$$

A proof would consist in applying the abstract theory in Section 10 to the problem

$\max_{f \in F, \check{x}(.) \in L_1(J,C(J,X))} \phi(x_0(T) + \int_J \check{x}(\tau)d\tau)$, subject to
$G(x_0(T) + \int_J \check{x}(\tau)d\tau) \in C$, $\check{x}(.) - f(., ., x_0(.) + \int_0^. \check{x}(\tau)d\tau) = 0$.

Appendix.

A. Proof of Lemma 10

(Adapted from Clarke (1983).) Assume first that $H = C(J, X)$. Let $\check{q}_0(t) := dq_0(t)/dt$ and choose a measurable function $\check{q}_1(t)$ such that $\check{q}_1(t) \in A(t, q_0(.)) \cap clB(\check{q}_0(t), 2\lambda_0(t))$ a.e., (the right hand side is a nonempty measurable set function with closed image sets, all contained in a separable set independent of t for a.e. t, so Kuratowskii's selection theorem applies, see e.g. Aubin and Frankowska (1990)). Note that $|\check{q}_1(t) - \check{q}_0(t)| \leq 2\lambda_0(t)$ a.e., (implying integrability of $\check{q}_1(t)$). Define $q_1(t) = q_0(0) + \int_0^t \check{q}_1(s)ds$, and note that by Lipschitz continuity, for some $a_t \in A(t, q_1(.)), |\check{q}_1(t) - a_t| \leq \kappa(t)|q_1(.) - q_0(.)|_t$, where $|q_1(.) - q_0(.)|_t = \sup_{s \leq t} |q_1(s) - q_0(s)|$, so $\lambda_1(t) := dist(\check{q}_1(t), A(t, q_1(.))) \leq \kappa(t)|q_1(.) - q_0(.)|_t$, (implying integrability of $\lambda_1(.)$). By induction, assume that an integrable function $\check{q}_n(.)$ is defined such that $dist(\check{q}_n(t), A(t, q_n(.))) =: \lambda_n(t)$ is integrable, where $q_n(t) = q_0(0) + \int_0^t \check{q}_n(s)ds$. Then, choose a measurable function $\check{q}_{n+1}(t)$ such that $\check{q}_{n+1}(t) \in A(t, q_n(.)) \cap clB(\check{q}_n(t), 2\lambda_n(t))$ a.e., and note that $|\check{q}_{n+1}(t) - \check{q}_n(t)| \leq 2\lambda_n(t)$ a.e.,(implying integrability of $\check{q}_{n+1}(t)$). Define $q_{n+1}(t) = q_0(0) + \int_0^t \check{q}_{n+1}(s)ds$, and note that by Lipschitz continuity, for some $a_t \in A(t, q_{n+1}(.))$, $|\check{q}_{n+1}(t) - a_t| \leq \kappa(t)|q_{n+1}(.) - q_n(.)|_t$, so $\lambda_{n+1}(t) := dist(\check{q}_{n+1}(t), A(t, q_{n+1}(.))) \leq \kappa(t)|q_{n+1}(.) - q_n(.)|_t$, (implying integrability of $\lambda_{n+1}(.)$). Hence, $\check{q}_n(.)$ is defined for all n.

Now, $|\check{q}_{n+1}(t) - \check{q}_n(t)| \leq 2\lambda_n(t) \leq 2\kappa(t)|q_n(.) - q_{n-1}(.)|_t$, a.e. Consider for a moment the iterative solution $\alpha_{n+1}(t) = \int_0^t 2\kappa(s)\alpha_n(s)ds$ of the equation $\alpha(t) = \int_0^t 2\kappa(s)\alpha(s)ds, \alpha(0) = 0$, with $\alpha_1(t) = |q_1(.) - q_0(.)|_t$. By induction, it is easily seen that $\alpha_{n+1}(t) \leq \alpha_1(T)(2\kappa^*(t))^n/n!$, where $\kappa^*(t) = \int_0^t \kappa(\tau)d\tau$. By induction, evidently $|q_{n+1}(.) - q_n(.)|_t \leq \alpha_{n+1}(t)$, so $|\check{q}_{n+1}(t) - \check{q}_n(t)| \leq 2\kappa(t)\alpha_n(t)$ a.e., which gives that $\check{q}_n(t)$ is an a.e. pointwise - , and L_1 - Cauchy sequence. Evidently, $\lim_n \check{q}_n(t) = \check{q}(t)$ a.e. exists and yields an integrable function, moreover $\check{q}(t) \in A(t, q(.))$ a.e., (by Lipschitz continuity), where $q(t) := q_0(0) + \int_0^t \check{q}(s)ds = \lim_n q_n(t)$ (uniform limit). Note that $|q(t) - q_0(t)| = \lim_n |q_{n+1}(t) - q_0(t)| = \lim_n |\sum_{0\leq j\leq n} q_{j+1}(t) - q_j(t)| \leq \beta(t)$, where $\beta(t) := \sum_{n\geq 1} \alpha_n(t)$, with $\beta(t)$ satisfying $\beta(t) = \alpha_1(t) + \int_0^t 2\kappa(s)\beta(s)ds$. By Gronwall's inequality, (as $\alpha_1(.)$ is nondecreasing), $|q(.) - q_0(.)|_t \leq \beta(t) \leq |q_1(.) - q_0(.)|_t \exp(\int_0^t 2\kappa(s)ds) \leq (\int_0^t 2\lambda_0(s)ds)\exp(\int_0^t 2\kappa(s)ds)$.

If H is a subset of $C(J, X)$, with $B(q_0(.), \varepsilon) \subset H$, then the applications of Lipschitz continuity above require that we know that all $q_n(.)$ belong to $B(q_0(.), \varepsilon)$ (or at least to H). Now, all $q_n(.)$ satisfy $|q_n(t) - q_0(t)| \leq \beta(t) < \varepsilon$. This ends the proof.

B. Further comments on d.d. containers

For the definition of d.d. containers to apply, there is no need for the functions to be locally Lipschitz continuous functions (not even continuous), though local Lipschitz continuity is generally assumed in this paper, (all sections except Section 8 and the present section B) . In this section, neither it is assumed that d.d. containers are bounded, and linearly homogeneous. We shall nevertheless see that it is natural to assume such properties, at least for locally Lipschitz continuous functions.

Loosely speaking, the smaller the d.d. containers are, the better tools they are. However, in general, no loss is incurred by working with $clDE(x_0)(v)$, rather than $DE(x_0)(v)$. (In this remark, as in Section 8, $E{:}X \to Y$ and X and Y are normed spaces.) By convention $DE(x_0)(0) = \{0\}$. Any enlargement of a d.d. container is a container, and the enlargement is uniform in $v \in V$, if the original d.d. container is. For each $r > 0, \Delta^* E(x_0)(v,r)$ is a d.d. container. Note that, for any d.d. container $DE(x_0)(v)$ that E might have at $x_0, D^* E(x_0)(v) := \cap_{r>0} cl \Delta^* E(x_0)(v,r)$ is a subset of $clDE(x_0)(v)$. If $D^* E(x_0)(v)$ is a d.d. container, then E is said to have a set-valued directional derivative, and in that case we write $D^s E(x_0)(v)$ instead of $D^* E(x_0)(v)$. If $cl\Delta^* E(x_0)(v,r)$ is norm-compact for some $r > 0$, then $D^* E(x_0)(v)$ is automatically a d.d. container, and so a set-valued directional derivative, which in this "compact" case is called a directional multiderivative, as in Seierstad (1997). (It is automatically nonempty.)

The intersection of two d.d. containers of E at x_0 in direction v, is not necessarily a d.d. container, (the intersection might even be empty.).

a. If E has a set-valued directional derivative at x_0 in direction v, then $D^s E(x_0)(v)$ is the smallest of all closed d.d. containers.

b. For $\mu > 0, \Delta E(x_0)(\mu v, \lambda)/\mu = \Delta E(x_0)(v, \mu\lambda)$ and $\Delta^* E(x_0)(\mu v, r)/\mu = \Delta^* E(x_0)(v, \mu r)$. Thus, if $DE(x_0)(v)$ is a d.d. container of E in direction v, then $\mu DE(x_0)(v)$ is a d.d. container in direction μv: If $DE(x_0)(v) + B(0, \varepsilon/\mu)$ contains $\Delta^* E(x_0)(v, \mu r)$, then $\mu DE(x_0)(v) + B(0, \varepsilon)$ contains $\mu \Delta^* E(x_0)(v, \mu r) = \Delta^* E(x_0)(\mu v, r)$.

c. If E has a d.d. container $DE(x_0)(v)$ at x_0 in a given direction v, then, by b., automatically, d.d. containers $DE(x_0)(\mu v)$ in directions μv are obtained for all $\mu \in (0, \infty)$, by defining $DE(x_0)(\mu v) = \mu DE(x_0)(v)$, and for any $K > 0$, they are uniform with respect to $\mu \in (0, K]$. For the moment it is assumed that $DE(x_0)(v)$ originally is defined for v belonging to a given sphere, here chosen to be the unit sphere, (i.e. $|v| = 1$). Extend the definition of $DE(x_0)(v)$ to v's of norms different from 1, by $DE(x_0)(v) := |v| DE(x_0)(v/|v|)$. Then $DE(x_0)(v)$ is (positively) linearly homogeneous in v. If $DE(x_0)(v)$ is uniform in $v \in V \subset \{v \in X : |v| = 1\}$, then, for any K, $DE(x_0)(\mu v)$ is uniform in $v \in V, \mu \in (0, K]$. If, at the outset, d.d. containers are specified for all $v \in V', V'$ a cone, then a natural "consistency assumption" on the set function $DE(x_0)(v), v \in V'$, would be that $DE(x_0)(v)$ is linearly homogeneous.

d. If $E(x)$ is Lipschitz continuous in $B(x_0, \gamma)$ of rank M_E, then for $\lambda v \in B(0, \gamma)$, (or $\lambda \le \gamma/|v|$), $|\Delta E(x_0)(v, \lambda)| \le M_E |v|, |\Delta^* E(x_0)(v, r)| \le M_E |v|$ and $|cl\Delta^* E(x_0)(v, r)| \le M_E |v|$ for $r < \gamma/|v|$, and, for any $\varepsilon > 0$, these three set functions are Lipschitz continuous of rank $M_E + \varepsilon$ in $v \in V$, in any given bounded set V, for $\lambda, r \le \gamma/|V|$(the two first ones have actually rank M_E).

e. Let $E(x)$ be Lipschitz continuous in $B(x_0, \gamma)$ of rank M_E. Let $DE(x_0)(v)$ be a d.d. container of E, uniform in $v \in V$, V a bounded set. Then, for any $\varepsilon > 0$, for some $r \in (0, \gamma/|V|], \Delta^* E(x_0)(v, r) \subset DE(x_0)(v) + B(0, \varepsilon)$, so for each $y \in \Delta^* E(x_0)(v, r)$, there exists a $y' \in DE(x_0)(v)$, with $|y - y'| < \varepsilon$. But then $y' \leq |y| + \varepsilon \leq M_E|v| + \varepsilon \leq M_E|V| + \varepsilon$, so $DE(x_0)(v) \cap clB(0, M_E|V| + \varepsilon)$ is also a d.d. container for E at x_0, uniform in $v \in V$, if $DE(x_0)(v)$ is. If in particular $DE(x_0)(v)$ is linearly homogeneous on a cone $\tilde{V}$, uniform in a set $V \subset V' := \tilde{V} \cap \{x \in X : |x| = 1\}$, then $DE(x_0)(v) \cap clB(0, (M_E + \varepsilon)|v|)$ is a linear homogeneous d.d. container at x_0 for all v in $\tilde{V}$, uniform in $v \in V'' = \{\lambda q : \lambda \in (0, 1], q \in V\}$.

Hence, a bound of the form $M|v|$ on d.d. containers of locally Lipschitz continuous functions is not seldom a natural assumption.

Modification of d.d. containers to obtain Lipschitz continuity is not so simple, but at least recall that the sets they shall "approximately" contain, namely the sets of difference quotients are Lipschitz continuous in the direction if the function itself is locally Lipschitz continuous. Note also the following property:

f. Let E have Lipschitz rank M_E in $B(x_0, \gamma)$ and let E have a set-valued directional derivative $D^s E(x_0)(v)$ at x_0 for all $v \in V$ in a given set V. Then, for any $\varepsilon > 0, D^s E(x_0)(v)$ is Lipschitz continuous in $v \in V$ with rank $M_E + \varepsilon$.

Proof. Given v and v' in $V, v \neq v'$, and $\varepsilon > 0$. Let $x \in D^s E(x_0)(v)$. There exists a $r \in (0, \gamma / \max\{|v|, |v'|\}]$ such that $\Delta^* E(x_0)(v', r) \subset D^s E(x_0)(v') + B(0, \varepsilon|v - v'|)$. We also have $D^s E(x_0)(v) \subset cl\Delta^* E(x_0)(v, r)$. So, for some $\lambda \in (0, r], x'' := \Delta E(x_0)(v, \lambda)$ satisfies $|x'' - x| < \varepsilon|v - v'|$. Let $x' := \Delta E(x_0)(v', \lambda)$, and note that $|x'' - x'| \leq M_E|v - v'|$. By the first inclusion, there exists a x^* in $D^s E(x_0)(v')$ such that $|x^* - x'| < \varepsilon|v - v'|$. The three last inequalities yield $|x - x^*| \leq (M_E + 2\varepsilon)|v - v'|$.

Local Lipschitz continuity of E, and nonemptyness of the contingent derivative $D^* E(x_0)(v)$ does not necessarily imply Lipschitz continuity of $D^* E(x_0)(v)$ in v.

g. If E has a d.d. container $DE(x_0)(v)$ which is a one point set, then $DE(x_0)(v) = \cap_{r>0} cl\Delta^* E(x_0,)(v, r)$, (i.e. E has a set-valued directional derivative at x_0 in direction v), and in fact E has a directional derivative at

x_0 in direction v.

Proof: Let $DE(x_0)(v) = \{y\}$ and let r be > 0. Then, for all $\varepsilon = 1/n$, there exists an r_n in $(0, r]$ such that $\Delta^* E(x_0)(v, r_n) \subset \{y\} + B(0, 1/n)$, which implies $y \in \Delta^* E(x_0)(v, r_n) + B(0, 1/n) \subset \Delta^* E(x_0)(v, r) + B(0, 1/n)$, for all n. Hence, $y \in cl\Delta^* E(x_0)(v, r)$ for all r. On the other hand, $\{y\} = clDE(x_0)(v) \supset \cap_{r>0} cl\Delta^* E(x_0)(v, r)$.

h. If E is continuous in a ball $B(x_0, \lambda)$, then for each $v, \Delta^* E(x_0)(v, r)$ is a separable set for $r < \lambda/|v|$. In fact, let S be a countable dense set in $(0, \infty)$. Then, $\Delta' E(x_0)(v, r) := \{\Delta E(x_0)(v, \lambda) : \lambda \in S, \lambda \leq r\}$ is dense in $\Delta^* E(x_0)(v, r)$.

i. Let E be continuous in $B(x_0, \gamma)$. Let $DE(x_0)(v)$ be a closed d.d. container of E at x_0 (existing and) uniform for $v \in V$, V a given bounded set. Then for each n, there exists $r_n \in (0, \gamma/|V|)$ such that $\Delta' E(x_0)(v, r_n) \subset DE(x_0)(v) + B(0, 1/2n)$. To each point y in $\Delta' E(x_0)(v, r_n)$, there corresponds a point $z_{y,n} \in DE(x_0)(v)$ such that $|y - z_{y,n}| < 1/2n$. Thus, for $D_v := \cup_n \{z_{y,n} : y \in \Delta' E(x_0)(v, r)\} \subset DE(x_0)(v)$, we have that $\Delta' E(x_0)(v, r_n) \subset D_v + B(0, 1/2n)$. Then, by h., $\Delta^* E(x_0)(v, r_n) \subset D_v + B(0, 1/n)$, so D_v (and then also clD_v), is a separable d.d. container smaller that $DE(x_0)(v)$, uniform in $v \in V$.

Hence, it is often natural to assume at the outset that d.d. containers are separable sets.

j. Let $g(s, x)$ be strongly measurable in s and Lipschitz continuous in x in $clB(x_0, \lambda)$. For each s, assume that at x_0 , $x \to g(s, x)$ has a closed d.d. container $D_2 g(s, x_0)(v), v$ a given vector. Choose functions $r_n(s) \in (0, \lambda/|v|]$ such that $\Delta_2^* g(s, x_0)(v, r_n(s)) \subset D_2 g(s, x_0)(v) + B(0, 1/2n)$. Let S be the set of rational numbers in $(0, \infty)$. To each function $y(.)$ of the form $\Delta_2 g(s, x)(v, \lambda), \lambda \in S$, if $\lambda \leq r_n(s)$, there corresponds a point $z_{\lambda,n}(s) \in D_2 g(s, x_0)(v)$ such that $|y(s) - z_{\lambda,n}(s)| < 1/2n$, while if $\lambda > r_n$, put $z_{\lambda,n}(s) = 0$. Thus, $D_v(s) := \{z_{\lambda,n}(s) : y(.) = \Delta_2 g(., x_0)(v, \lambda), \lambda \in S, n = 1, 2, ...\}$ is a countable set, and $D_v(s)$ would have been measurable if each $s \to z_{\lambda,n}(s)$ had been measurable. Assume this measurability for the moment. Then also, by measurability, one might have assumed these functions to be separably valued, which would entail the existence of a separable set X_v such that $D_v(s) \subset X_v$ for a.e. s. By i., $clD_v(s)$ is a d.d. container of $x \to g(s, x)$ at x_0 in direction v, by continuity of $\lambda \to \Delta_2 g(s, x_0)(v, \lambda)$ and density of S : For each s, each n, $\Delta_2^* g(s, x_0)(v, r_n(s)) \subset clD_v(s) + B(0, 1/n))$. Unfortunately, no selection result exists that can give that the $z_{\lambda,n}$-functions can be chosen to be measurable, for the simple reason that the mere postulation of d.d. containers for each s, does not imply any relation between these d.d. containers for various s's.(Perhaps even the $r_n(.)$'s are nonmeasurable.) Thus, we have to *assume* that $s \to D_2g(s,x_0)(v)$ is measurable and essentially separably valued, if we want to "integrate" $D_2g(s,x_0)(v)$, i.e. integrate functions taking values in this set. (We must know that there are measurable functions taking values in this set.) Note that the set $D^{''}(s,v,1/m) = \{y(s) : y(.) = \Delta_2 g(s,x_0)(v,\lambda)$ for some $\lambda \in S \cap (0,1/m]\}, m \in \{1,2,...\}$, is a measurable set function of s, so $clD^{''}(s,v,1/m) = cl\Delta_2^* g(s,x_0)(v,1/m)$ and $\cap_m cl\Delta_2^* g(s,x_0)(v,1/m)$ are measurable in s, as well as essentially separably valued. If $D_2g(s,x_0)(v)$ is not only a d.d. container but even a set-valued directional derivative, then, by definition, $D_2g(s,x_0)(v)$ equals $\cap_m cl\Delta_2^* g(s,x_0)(v,1/m)$. In this case, $D_2g(s,x_0)(v)$ is automatically measurable and essentially separably valued.

The assumption that d.d. containers of the type $D_2g(s,x_0)(v)$), (compare (10)), are measurable and essentially separably valued should be illuminated by the above discussion. The two properties just mentioned are needed for the measurable selection occuring in the following proof.

Proof of F in Section 8. Let $\varepsilon > 0$. Note that $\Delta^*\tilde{f}(x_0)(v,r) = \{\int_I \lambda^{-1}[g^*(s,x_0+\lambda v) - g^*(s,x_0)]ds : \lambda \in (0,r]\}$. For each s, each $v \in V$,

$\eta(v,s,r) := \sup_{\lambda\in(0,r]} \text{dist}(\lambda^{-1}[g^*(s,x_0+\lambda v) - g^*(s,x_0)], D_2g^*(s,x_0)(v))$ converges to zero with r, moreover, η is bounded by $\kappa^*(s) := \kappa(s)|V| + \kappa'(s)$, for $r < \gamma|V|$. For fixed v, by dominated convergence, $\int_I \eta(v,s,r)ds < \varepsilon/2$, for r small enough. By measurable selection, for any function $w(v,s,\lambda) := [g^*(s,x_0+\lambda v) - g^*(s,x_0)]/\lambda, \lambda \in (0,r]$, there exists an integrable function $s \to w^*(v,s) \in D_2g^*(s,x_0)(v)$ a.e., such that $|w(v,s,\lambda) - w^*(v,s)| \le 2\eta(v,s,r)$, hence $|\int_I w(v,s,\lambda)ds - \int_I w^*(v,s)ds| < \varepsilon$. This shows that $D\tilde{f}(x_0)(v)$ is a d.d. container of $\tilde{f}(x)$ at x_0 in direction v. Next, by contradiction, assume that it is not uniform in $v \in V$, i.e. for some $\varepsilon > 0$, for all n, some pair $(v_n,\lambda_n) \in V\times\ (0,1/n]$ exists, such that $w(v_n,s,\lambda_n)$ does not satisfy $\int_I w(v_n,s,\lambda_n)ds \in D\tilde{f}(x_0)(v_n) + B(0,\varepsilon)$. Define $\eta(s,r) = \sup_n \eta(v_n,s,r) \le \kappa^*(s)$. By uniformity of $D_2g^*(s,x_0)(v)$ in V and dominated convergence, $\int_I \eta(s,r)ds < \varepsilon/2$, if r is small enough, say $r < r'$. Choose an n such that $1/n < r'$. Evidently, $|\int_I w(v,s,\lambda)ds - \int_I w^*(v,s)ds| < \varepsilon$ can then be obtained for $v = v_n$, which gives a contradiction.

C. Mathematical programming results

To illuminate the control theory results, results in mathematical programming following from the abstract results of Section 10 are presented. In the control theory above, we use rapid switching in order to obtain approximate convexity. Here we shall see what comes out of the "abstract" setting when (true) convexity is assumed. In addition to d.d. containers, also contingent derivatives are used. Throughout this Section C, A is a complete convex set in a normed space $\hat{A}, Y$ is a normed space, C is a complete convex set in Y, $H : A \to Y$ and $\eta : A \to \mathbb{R}$ are continuous, and a^* is a given point in A. By redefinitions of $\hat{A}$, A, and H, problems of the type $\min_a \eta(a)$, subject to $a \in A, H(a) \in C$, can be reduced to the case where $C = \{0\}$. Though confining the discussion to the latter case would save space, for easy comparison with results above, below we stick to the general problem.

Below, the contingent derivative of $a \to (H(a), \eta(a))$ is written $D^*(H,\eta)(a)(v), (a \in A, v \in \hat{A})$. Consider the following condition: For some $\check{\mu} \in (0,1), \hat{\mu} > 0, \mu' > 0, c^* \in C, \hat{z}^* := (z^*, \omega^*) \in Y \times \mathbb{R}, K > 0$,

for all $(a,c) \in (A \cap clB(a^*, \hat{\mu})) \times (C \cap clB(c^*, \hat{\mu}))$, for all $\hat{v} := (v, \omega) \in Y \times (-\infty, \omega^*]$ with $|\hat{v} - \hat{z}^*| = \mu'$, there exist elements $a'' \in A \cap clB(a, K|\hat{v}|)$, $c'' \in C \cap clB(c, K|\hat{v}|)$, and (v'', ω'') such that

$$(*) \qquad \begin{aligned} &(v'', \omega'') \in D^*(H,\eta)(a)(a'' - a), \\ &|v'' - v - (c'' - c)| \le (1 - \check{\mu})\mu'|\hat{v}|/(|\hat{z}^*| + \mu') \text{ and} \\ &\omega'' - \omega \le (1 - \check{\mu})\mu'|\hat{v}|/(|\hat{z}^*| + \mu'). \end{aligned} \tag{58}$$

This condition implies (51) for $\mu'' = \check{\mu}/2$. To show this, note that, since $(v'', \omega'') \in cl\Delta^*(H,\eta)(a)(a'' - a, r)$ for all r, if $\hat{v} \ne 0$, there exists a $\delta > 0$, arbitrary small, such that $|[(H(a'), \eta(a')) - (H(a), \eta(a))]/\delta - (v'', \omega'')| \le (\check{\mu}/2)\mu'|\hat{v}|/(|\hat{z}^*| + \mu')$, where $a' = \delta a'' + (1 - \delta)a$. Note that this inequality trivially holds if $\hat{v} = 0$ ($\Rightarrow a'' = a, c'' = c$). Hence,

$$\begin{aligned} &(58) \text{ implies the existence of arbitrary small } \delta > 0, \text{ such that} \\ &|(H(a') - H(a))/\delta - v - (c'' - c)| \le (1 - \check{\mu}/2)\mu'|\hat{v}|/(|\hat{z}^*| + \mu') \text{ and} \\ &(\eta(a') - \eta(a))/\delta - \omega \le (1 - \check{\mu}/2)\mu'|\hat{v}|/(|\hat{z}^*| + \mu'). \end{aligned} \tag{59}$$

(For $\hat{v} = 0$,(59) trivially holds, put $a'' = a, c'' = c$.)

Below, we need the following fact. If (58) holds, then by slightly decreasing $\check{\mu}$ and μ' if necessary, we can always obtain that (58) holds with the following assumption added:

$$|\hat{z}^*| \neq \mu'. \tag{59'}$$

Assume that, initially, (58) holds with $\mu' = |\hat{z}^*|$. Then choose $\hat{\mu}' > 0$ and $\mu'' \in (0,1)$ slightly smaller than μ' and $\check{\mu}$, respectively, such that $(1-\mu'')\hat{\mu}' \geq (1-\check{\mu})\mu'$. Note that if the element $\hat{v} := (v,\omega) \neq 0, \omega \leq \omega^*$, satisfies $|\hat{v}-\hat{z}^*| = \hat{\mu}'$, then $|\infty\hat{v}-\hat{z}^*| = \infty$, so for some $\lambda > 1, |\lambda\hat{v}-\hat{z}^*| = \mu'$). Using (58) for $\hat{v}$ replaced by $\lambda\hat{v}$ yields elements a'', c'' , v'', and ω'' corresponding to $\lambda\hat{v}$, which we write as a_λ, c_λ , v_λ, and ω_λ instead, such that (58)(*) is satisfied. Dividing by λ in this version of (58)(*) gives that (*) holds for $\hat{v}$, together with $a'' = (a_\lambda - a)/\lambda + a, c'' = (c_\lambda - c)/\lambda + c, v'' = v_\lambda/\lambda, \omega'' = \omega_\lambda/\lambda$. Dividing by λ in the inequalities in (58) gives that they are satisfied by $\hat{v}$ and these a'', c'', v'', ω''. On the right hand side of the inequalities of course $(1-\check{\mu})\mu'$ can be replaced by $(1-\mu'')\hat{\mu}'$, so (59') follows.

When (59') holds, then in (58), if we replace
$clB(a, K|\hat{v}|)$ and $clB(c, K|\hat{v}|)$ by $clB(a, \check{K})$ and $clB(c, \check{K})$ (59")

respectively, where $\check{K}$ is some positive constant, then, we get a condition (denoted (58")) implying (58), for $K = \check{K}/\xi, \xi := |\mu' - |\hat{z}^*||$. This follows simply from the fact that if $|\hat{v} - \hat{z}^*| = \mu'$, then $|\hat{v}| \geq \xi$, so $clB(a, \check{K}) \subset clB(a, K|\hat{v}|)$ and $clB(c, \check{K}) \subset clB(c, K|\hat{v}|)$.

Let the contingent derivative $D^*(H,\eta)(a)(v)$ be Lipschitz continuous in v, with a rank independent of a, for $a \in A$. Then (58") is implied by the following condition, for $\check{K} := \hat{K} + \hat{\mu}'$. For some $\mu^+ \in (0,1), \hat{\mu}' > 0, \mu' > 0, c^* \in C, \hat{z}^* := (z^*, \omega^*) \in Y \times \mathbb{R}, \hat{K} > 0$, for all $a \in clB(a^*, \hat{\mu}')$,

for all $\hat{v} := (v,\omega) \in Y \times (-\infty, \omega^*]$ with $|\hat{v} - \hat{z}^*| = \mu'$, there exist elements $a'' \in A \cap clB(a^*, \hat{K})$, $c'' \in C \cap clB(c^*, \hat{K})$, and $(\hat{v}'', \hat{\omega}'') \in D^*(H,\eta)(a)(a'' - a^*)$, such that $|\hat{v}'' - v - (c'' - c^*)| \leq (1-\mu^+)\mu'|\hat{v}|/(|\hat{z}^*| + \mu')$ and $\hat{\omega}'' - \omega \leq (1-\mu^+)\mu'|\hat{v}|/(|\hat{z}^*| + \mu')$.(60)

The number μ' can be taken to be the same as the one occurring in (58"), except when $\mu^+ = |\hat{z}^*|$. In the latter case, μ^+ can be decreased slightly so that inequality obtains, (see arguments connected with (59')), we assume in what follows that this has already been carried out. Thus, with $\mu' \neq |\hat{z}^*|$, consider a quadruple $(a'', c'', (\hat{v}'', \hat{\omega}''), (v, \omega))$ for which (60) is satisfied, Evidently, $\xi = |\mu' - |\hat{z}^*|| = \inf\{|\hat{v}| : |\hat{v} - \hat{z}^*| = \mu'\}$. Then for $\hat{\mu} \in (0, \hat{\mu}']$, small enough, for $a \in A \cap clB(a^*, \hat{\mu}), c \in C \cap clB(c^*, \hat{\mu})$, there exists an element $\hat{v}' := (v', \omega')$ in $D^*(H,\eta)(a)(a'' - a)$ such that the inequalities $|\hat{v}''-\hat{v}'| \le \mu^+\mu'\xi/4(|\hat{z}^*|+\mu')$, and $|c''-c^*-(c''-c)| \le \mu^+\mu'\xi/4(|\hat{z}^*|+\mu')$ hold. If, in the inequalities in (60), $\hat{v}'', \hat{\omega}'', a^*$, and c^* are replaced by v', ω', a, and c, respectively, then the validity of the inequalities are rescued by adding $\mu^+\mu'\xi/2(\ |\hat{z}^*|+\mu')$ on the right hand sides. Moreover, $|a''-a^*| \le \hat{K}$, $|a-a^*| \le \hat{\mu}$ implies $|a''-a| \le \hat{K}+\hat{\mu} \le \check{K}$. Similarly, $|c''-c| \le \check{K}$. Thus (58") holds, and, hence, (51) is implied by (60).

If the uniform Lipschitz continuity of $D^*(H,\eta)(a)(v)$ does not hold, it is only possible to prove, (using part of the arguments above), that (58") is implied by the following condition: For some $\mu^+ \in (0,1), \hat{\mu}' > 0, \mu' > 0, c^* \in C, \hat{z}^* := (z^*, \omega^*) \in Y \times \mathbb{R}, \hat{K} > 0$, for all $a \in clB(a^*, \hat{\mu}')$,

for all $\hat{v} := (v, \omega) \in Y \times (-\infty, \omega^*\ ]$ with $|\hat{v} - \hat{z}^*| = \mu'$, there exist elements $a'' \in A \cap clB(a, \hat{K})$, $c'' \in C \cap clB(c^*, \hat{K})$, and $(\hat{v}'', \hat{\omega}'') \in D^*(H,\eta)(a)(a''-a)$, such that $|\hat{v}''-v-(c''-c^*)| \le (1-\mu^+)\mu'|\hat{v}|/(|\hat{z}^*|+\mu')$ and $\hat{\omega}''\ -\omega \le (1-\mu^+)\mu'|\hat{v}|/(|\hat{z}^*|+\mu')$. (60')

Next, consider the case where $\omega^* = 0$, $z^* \in C - c^*$ in (60'). In this case, (60') also holds for $\hat{\mu}'$, μ', z^*, and $\hat{K}$ replaced by $\hat{\mu}'/2$, $\mu'/2$, 0, and K'' defined below. To see this, let $\hat{v}, a'', c'', \hat{v}'', \hat{\omega}''$ satisfy (60'), and write $v' = v - z^*$, $z^* := \tilde{c} - c^*$, $\tilde{c} \in C$, $\tilde{c}'' = (\frac{1}{2})\tilde{c} + (\frac{1}{2})c'' \in C$. Note that $|\tilde{c}''-c^*| = |\tilde{c}/2+c''/2-c^*/2-c^*/2| = |(c''-c^*)/2\ -c^*/2+\tilde{c}/2| \le \hat{K}/2+|\tilde{c}/2-c^*/2| =: K''$. Furthermore, note that $|\hat{v}''/2-v/2-(c''-c^*)/2| = |\hat{v}''/2-v'/2-z^*/2-(c''-c^*)/2| =$
$|\hat{v}''/2-v'/2-\tilde{c}/2+c^*/2-(c''-c^*)/2)| =$
$|\hat{v}''/2-v'/2-(\tilde{c}''-c^*)| \le (1-\mu^+)(\mu'/2)|\hat{v}|/(|\hat{z}^*|+\mu') \le (1-\mu^+)(\mu'/2)$, and $\hat{\omega}''/2\ -\omega/2 \le (1-\mu^+)(\mu'/2)|\hat{v}|/(|\hat{z}^*|+\mu') \le (1-\mu^+)(\mu'/2)$. Note that $(\hat{v}''/2, \hat{\omega}''/2) \in D^*(H,\eta)(a)[\tilde{a}''-a)]$, where $\tilde{a}'' := a''/2+a/2$ and that $|\tilde{a}''-a| \le \hat{K}/2 \le K''$.

Now, recall that when $z^* \in C - H(a^*), \omega^* = 0$, (60') implies (53) in Corollary 5. The calculations just carried out show the equivalence of the following two necessary condition (i) and (ii) in (61) for optimality of a^*.

(i) There exists no quintuple $(\mu^+, \hat{\mu}', \mu', z^*, \hat{K}), z^* \in C - H(a^*)$, such that (60') holds for $\omega^* = 0$, $c^* = H(a^*)$.
(ii) There exists no quadruple $(\mu^+, \hat{\mu}', \mu', \hat{K})$, such that (60') holds for $\omega^* = 0, c^* = H(a^*), z^* = 0$. (61)

Assume next that $a \to (H(a), \eta(a))$ has a d.d. container denoted

$D(H,\eta)(a)(v), a \in A, v \in \hat{A}$, and consider the following condition: For some $\breve{\mu} \in (0,1), \hat{\mu} > 0, \mu' > 0, c^* \in C, \hat{z}^* := (z^*, \omega^*) \in Y \times \mathbb{R}, K > 0$,

$$\text{for all } (a,c) \in (A \cap clB(a^*,\hat{\mu})) \times (C \cap clB(c^*,\hat{\mu})), \text{ for all } \hat{v} := (v,\omega) \in Y \times (-\infty,\omega^*)] \text{ with } |\hat{v} - \hat{z}^*| = \mu', \text{ there exist an } a'' \in A \cap clB(a, K|\hat{v}|) \text{ and a } c'' \in C \cap clB(c, K|\hat{v}|), \text{ such that for all } (v'',\omega'') \in D(H,\eta)(a)(a''-a),$$
$$|v'' - v - (c''-c)| \le (1-\breve{\mu})\mu'|\hat{v}|/(|\hat{z}^*| + \mu') \text{ and}$$
$$\omega'' - \omega \le (1-\breve{\mu})\mu'|\hat{v}|/(|\hat{z}^*| + \mu'). \tag{62}$$

This condition implies (51) for $\mu'' = \breve{\mu}/2$. To show this, we need only consider the case $\hat{v} \neq 0$. Note that there exists a $r > 0$ such that for all $\delta \in (0,r]$, for some $(v_\delta, \omega_\delta) \in D(H,\eta)(a)(a''-a)$, $|[(H(a_\delta),\eta(a_\delta)) - (H(a),\eta(a))]/\delta - (v_\delta,\omega_\delta)| \le (\breve{\mu}/2)\mu'|\hat{v}|/(|\hat{z}^*| + \mu')$, where $a_\delta = \delta a'' + (1-\delta)a$. Combining this with (62), we evidently get: For all $\delta \in (0,r]$,

$$|[H(a_\delta) - H(a)]/\delta - v - (c''-c)| \le (1-\breve{\mu}/2)\mu'|\hat{v}|/(|\hat{z}^*| + \mu'), \text{ and}$$
$$[\eta(a_\delta) - \eta(a)]/\delta - \omega \le (1-\breve{\mu}/2)\mu'|\hat{v}|/(|\hat{z}^*| + \mu'). \tag{63}$$

This is more than we need, ((63) says that (51) holds for triples (a', c'', δ) for *all* $\delta \in (0,r]$). When we do not have recourse to contingent derivatives, but only to d.d. containers, we cannot simply assume in (62) that the two inequalities hold for some $(v'',\omega'') \in D(H,\eta)(a)(a''-a)$. The set $D(H,\eta)(a)(a''-a)$ may contain elements unrelated to any term of the form $[(H(a_\delta),\eta(a_\delta)) - (H(a),\eta(a))]/\delta$, but we need an inequality that holds for some element of this form.

We can work with a slight modification of (62): For a function $E : X \to Y$, define a directional derivative trap at x_0 in direction v to be a set $D^tE(x_0)(v)$ with the property that, for all $\varepsilon > 0$, for all $r > 0$, $\Delta^*E(x_0)(v,r) \cap (D^tE(x_0)(v) + B(0,\varepsilon)) \neq \emptyset$. This property is equivalent to $(\text{cl}\Delta^*E(x_0)(v,r)) \cap (D^tE(x_0)(v) + B(0,\varepsilon)) \neq \emptyset$, so if E has a nonempty contingent derivative, then it is a d.d. trap. And, of course, a d.d. container is a d.d. trap. To know that (63) holds, for some $\delta \in (0,r]$, (and this is all we need), it suffices to assume in (62) that $D(H,\eta)(a)(v)$ is a d.d. trap. We have not worked with d.d. traps, because their calculus is proor. Let us nevertheless briefly comment on some properties that do hold:

If in B. in Section 8, $V = \{v\}$ for some v, and $DF(x_0)(v)$ is only a d.d. trap, then $DG(F(x_0))(DF(x_0)(v))$ is a d.d. trap of $H(x) = G(F(x))$. Much of the arguments in the proof can be kept. For any r, for some $\lambda_r \in (0,r]$, for $w' = \Delta F(x_0)(v,\lambda_r)$, we have $w' \in DF(x_0)(v) + B(0,\varepsilon/2M_G)$, so for some

$w \in DF(x_0)(v)$, $|w-w'| < \varepsilon/2M_G$. For r small enough, $\Delta^*G(F(x_0))(\tilde{w}, r) \subset DG(F(x_0))(\tilde{w}) + B(0, \varepsilon/2)$, for any $\tilde{w} \in DF(x_0)(v)$. By Lipschitz continuity, $\Delta^*G(F(x_0))(w', r) \subset DG(F(x_0))(w) + B(0, \varepsilon)$. In particular, $\Delta G(F(x_0))(w', \lambda_r) = \Delta H(x_0)(v, \lambda_r) \in DG(F(x_0))(w) + B(0, \varepsilon)$.).

It is also easily shown that $D^tE(x_0)(v) + DF(x_0)(v)$ is a d.d. trap of $E+F$ at x_0, (evidently we have here assumed that the trap $D^tE(x_0)(v)$ and the container $DF(x_0)(v)$ exists), similarly $D^tF(x_0)(v) \times D\tilde{F}(x_0)(v)$ is a d.d. trap of $F \times \tilde{F}$.

Note also in connection with Lemma 11, that if $h(v) := v \to x(., v)$, then any element of the contingent derivative $D^*h(v_0)(v)$ belongs to cl$Q(v)$, even when the uniformity condition on $D_2g(s, x(., v_0))(q^*(.))$ is dropped. To prove this let, $q^*(.) \in D^*h(v_0)(v)$, let $\Delta h(v_0)(v, \lambda_n)(.) \to q^*(.)$, when $\lambda_n \searrow 0$, and for any $\varepsilon > 0$, choose λ_n so small that $\int_J |\Delta h(v_0)(v, \lambda_n)(s) - q^*(s)|\kappa(s)ds < \varepsilon/4$ and $\int_J \gamma(t)dt < \varepsilon/4$, where $\gamma(t) := \text{dist}([g(t, x(., v_0) + \lambda_n q^*(.)) - g(t, x(., v_0))]/\lambda_n, D_2g(t, x(., v_0))(q^*(.)))$. Then, by Lipschitz continuity, the distance between $[g(t, x(., v_0)+\lambda_n \Delta h(v_0)(v, \lambda_n)(.)) - g(t, x(., v_0))]/\lambda_n$ and $D_2g(t, x(., v_0))(\Delta h(v_0)(v, \lambda_n))$ is $\leq \gamma(t) + 2\kappa(t)|\Delta h(v_0)(v, \lambda_n) - q^*(t)| =: \gamma'(t)$. By Lemma 10, there exists a $q(.) \in Q(v)$, such that $|\Delta h(v_0)(v, \lambda_n))(.) - q(.)| \leq (\int_J 2\gamma'(s)ds) \exp \int_J 2\kappa(s)ds \leq \varepsilon \exp \int_J 2\kappa(s)ds$.

In the case where C is nonconvex, if C has a nonempty adjacent cone at c, then in (62), we can require that $c'' - c$ belongs to such a cone, (then dropping $c'' \in C$), with (63) (for *some* $\delta \in (0, r]$ and c'' replaced by some $\tilde{c}'' \in C$ "adjacent" to c''), still following even in the trap case. This does not work if the cone is a contingent one. However, we can work with the following slight modification of (58) in the case where H and η are locally Lipschiz continuous. (It would work also without such a continuity assumption, but then the reader would perhaps object to the word "contingent"). Assume in (58) that we drop $(*)$ and $c'' \in C$, and instead assume that the triple (v'', ω'', c'') is a sort of "contingent" triple defined by the property that $\liminf_{\delta \searrow 0} \alpha(\delta) = 0$, where $\alpha(\delta) := \delta^{-1}\text{dist}(C, c + \delta(c'' - c)) + |v'' - [H(a + \delta(a'' - a)) - H(a)]/\delta| + |\omega'' - [\eta(a + \delta(a'' - a)) - \eta(a)]/\delta|$. (Then $c'' - c$ belongs to the contingent cone of C at c. We need however, the "simultaneous contingency" implied by the next to last equality). Then again (51) follows.

For the remaining part of this section we turn back to the case where C is convex. In (62) , (59') can be assumed, (the same arguments apply). Moreover, when the replacements given in (59") are carried out in (62), it yields a condition (denoted (62')) stronger than (62) with $K = \check{K}/\xi$, (the same arguments apply).

Let the d.d. container $D(H,\eta)(a)(v)$ be Lipschitz continuous in v, with a rank independent of $a \in A$. Consider the condition: For some $\mu^+ \in (0,1), \hat{\mu}' > 0, \mu' > 0, c^* \in C, \hat{z}^* := (z^*, \omega^*) \in Y \times \mathbb{R}, \hat{K} > 0$,

$$\begin{aligned}
&\text{for all } a \in clB(a^*, \hat{\mu}') \text{ for all } \hat{v} = (v,\omega) \in Y \times (-\infty, \omega^*] \text{ with}\\
&|\hat{v} - \hat{z}^*| = \mu', \text{ there exist an } a'' \in A \cap clB(a^*, \hat{K}) \text{ and a } c'' \in\\
&C \cap clB(c^*, \hat{K}), \text{ such that for all } (\hat{v}'', \hat{\omega}'') \in D(H,\eta)(a)(a'' - a^*),\\
&|\hat{v}'' - v - (c'' - c^*)| \le (1-\mu^+)\mu'|\hat{v}|/(|\hat{z}^*| + \mu')\\
&\text{and } \hat{\omega}'' - \omega \le (1-\mu^+)\mu'|\hat{v}|/(|\hat{z}^*| + \mu') \qquad (64)
\end{aligned}$$

The condition (64) implies that (62') holds for some $\hat{\mu} \in (0, \hat{\mu}']$, for some $\check{K} > 0, \check{\mu} = \mu^+/4$, and μ' perhaps slightly decreased. (If $|\hat{z}^*| = \mu'$, a slight decrease of μ^+ and μ' is necessary, with (64) still holding. Assume it has been carried out, so $\xi := ||\hat{z}^*| - \mu'| > 0$.) To prove the assertion, choose $\hat{\mu} \in (0, \min\{\hat{\mu}', \mu^+\mu'\xi/4(|\hat{z}^*| + \mu')\}]$ so small that, for any $a \in clB(a^*, \hat{\mu})$, $a'' \in clB(a^*, \hat{K})$, for any $(v'', \omega'') \in D(H,\eta)(a)(a'' - a)$, there exists a pair $(\hat{v}'', \hat{\omega}'') \in D(H,\eta)(a)(a'' - a^*)$ such that $|(\hat{v}'', \hat{\omega}'') - (v'', \omega'')| \le (\mu^+/4)\mu'\xi/(|\hat{z}^*| + \mu')$. As $c \in clB(c^*, \hat{\mu})$ implies $|c - c^*| \le \mu^+\mu'\xi/4(|\hat{z}^*| + \mu')$, then, for all pairs $(v'', \omega'') \in D(H,\eta)(a)(a'' - a)$, the two inequalities in (62) hold if the corresponding inequalities hold in (64). Finally, define $\check{K} := \hat{K} + \hat{\mu}'$, and note that $|c'' - c| \le |c'' - c^*| + |c - c^*| \le \hat{K} + \hat{\mu}' =: \check{K}$. A similar calculation yields $|a'' - a| \le \check{K}$. Thus (62') holds.

If the uniform Lipschitz continuity of $D(H,\eta)(a)(v)$ fails to hold, then $a'' - a^*$ has to replaced by $a'' - a$ in (64), giving rise to a condition we call (64'), see the parallell discussion leading to (60'). Moreover, two (again equivalent) necessary conditions are obtained by replacing (60') by (64') in (61).

Assume now that $D(H,\eta)(a)(v) = (DH(a)(v), D\eta(a)(v))$ is a directional derivative that exists for all $a \in A, v \in \hat{A}$. To signalize the existence of the directional derivative, we write D^d instead of D. Then (62) simplifies to

$$\begin{aligned}
&\text{For all } a \in A \cap clB(a^*, \hat{\mu}), c \in C \cap clB(c^*, \hat{\mu}), \text{ for all } \hat{v}\\
&:= (v,\omega) \in Y \times (-\infty, \omega^*] \text{ with } |\hat{v} - \hat{z}^*| = \mu',\\
&\text{there exist an } a'' \in A \cap clB(a, K|\hat{v}|) \text{ and a } c'' \in C \cap clB(c, K|\hat{v}|),\\
&\text{such that } |D^dH(a)(a'' - a) - v - (c'' - c)| \le (1-\check{\mu})\mu'|\hat{v}|/(|\hat{z}^*| + \mu')\\
&\text{and } D^d\eta(a)(a'' - a) - \omega \le (1-\check{\mu})\mu'|\hat{v}|/(|\hat{z}^*| + \mu'). \qquad (65)
\end{aligned}$$

Next, (64') simplifies to:

For all $a \in A \cap clB(a^*, \hat{\mu}')$, for all $\hat{v} := (v,\omega) \in$

$Y \times (-\infty, \omega^*]$ with $|\hat{v} - \hat{z}^*| = \mu'$ there exist an $a'' \in$
$A \cap clB(a, \hat{K})$ and a $c'' \in C \cap clB(c^*, \hat{K})$, such that
$|D^d H(a)(a'' - a) - v - (c'' - c^*)| \leq (1 - \mu^+)\mu'|\hat{v}|/(\ |\hat{z}^*| + \mu')$
and $D^d \eta(a)(a'' - a) - \omega \leq (1 - \mu^+)\mu'|\hat{v}|/(|\hat{z}^*| + \mu')$. (66)

In the problem: $\min_{a \in A} \eta(a)$, subject to $H(a) \in C$, the following necessary condition holds. There exists no quadruple $(\hat{K}, \mu, \mu^*, \mu')$ with $\omega^* = 0, z^* = 0$, such that (66) holds.

If $D^d H(a)(v)$ and $D^d \eta(a)(v)$ are uniformly continuous in v, uniformly in a in A, (in this case uniform *Lipschitz* continuity is not required), then in (66), $a'' - a$ can be replaced by $a'' - a^*$.

In the above uniform continuity case, for some $\gamma > 0, |D^d H(a)(v)| \leq 1$ and $|D^d \eta(a)(v)| \leq 1$ for all $v \in B(0, \gamma)$, all $a \in A$, hence $H(a)$ and $\eta(a)$ are Lipschitz continuous in A, and this is more than we need for another and more effective necessary condition do hold. We formulate it for less demanding differentiability requirements, we here use d.d. traps:

Corollary 8. Assume that $a \to H(a)$ has a d.d. trap $D^t H(a)(v)$, for all $a \in A, v \in \hat{A}$. Assume also that for all $a \in A$, there exist constants W_a, and W^a such that $|\eta(a') - \eta(a)| \leq W_a|a' - a|$, and $|H(a') - H(a)| \leq W^a|a' - a|$ for all $a' \in A$. Furthermore, assume that for each a, $W_{\lambda a + (1-\lambda)a^*} \to W_{a^*}$ when $\lambda \searrow 0$. Moreover, assume, for some $\check{\mu} \in (0, 1], \mu^* > 0, \mu' > 0$, that,

for all $a \in A \cap clB(a^*, \mu^*)$, for all $v \in Y$ with $|v| = \mu'$,
there exist an $a'' \in A \cap clB(a, 1)$ and a $c'' \in C \cap clB(H(a^*), 1)$,
such that for all $v'' \in D^t H(a)(a'' - a)$,
$|v'' - v - (c'' - H(a^*))| \leq (1 - \check{\mu})\mu'$. (67)

Finally, assume that a^* is optimal in the problem: $\min_a \eta(a)$, subject to $a \in A$, $H(a) \in C$. Then, for each $(a, c) \in A \times C$, for $c^* = H(a^*)$,

$$\limsup_{\lambda \searrow 0} \lambda^{-1}[\eta(a^* + \lambda(a - a^*)) - \eta(a^*) + 16KW_{a^*}|H(a^* + \lambda(a - a^*)) - c^* - \lambda(c - c^*)|/\check{\mu}] \leq 0, \quad (68)$$

where $K = \max\{1/\mu', 1 + 1/\hat{\mu}\}, \hat{\mu} = \min\{\mu^*, \check{\mu}\mu'/4\}$.

Proof. From (67), evidently, for some δ arbitrarily small, for $a_\delta = \delta a'' + (1 - \delta)a, |[H(a_\delta) - H(a)]/\delta - v - (c'' - c^*)| \leq (1 - \check{\mu}/2)|\hat{v}|$ follows, so (54') holds for $\mu'' = \check{\mu}/2, K' = 1/\mu'$, and Corollary 7 applies.

Note that by Corollary 7,
$\lambda^{-1}[\eta(a^*+\lambda(a-a^*))-\eta(a^*)+8KW_{\lambda a+(1-\lambda a^*)}|H(a^*+\lambda(a-a^*))-c^*-\lambda(c-c^*)|/(\check{\mu}/2)]\leq 0$, for λ so small that $\lambda(a-a^*)\in clB(0,\hat{\mu}/2), \lambda(c-c^*)\in clB(0,\hat{\mu}/2)$. From this inequality, (68) follows, since $|H(a^*+\lambda(a-a^*))-c^*-\lambda(c-c^*)|\leq \lambda W^a|a-a^*|+\lambda|c-c^*|$.

In case of Lipschitz continuity of $v\to D^tH(a)(v)$ uniformly (i.e. same rank) for all $a\in A$, then the set $A\cap clB(a,1)$ in (67) can be replaced by $A\cap clB(a^*,1)$, compare the relationship between (66) and (65).

Corollary 9. In the situation of Corollary 8, if $a\to(H(a),\eta(a))$ has a d.d. trap denoted $D^t(H,\eta)(a^*)(v)$ at a^*, for all $v\in\hat{A}$, then (68) implies that for each $(a,c)\in A\times C$,

$$\inf\{\omega+(16KW_{a^*}/\check{\mu})|v-c+H(a^*)|:(v,\omega)\in D^t(H,\eta)(a^*)(a-a^*)\}\leq 0.$$

In particular, if $D(H,\eta)(a^*)(v)$ is a directional derivative, written $D^d(H,\eta)$, then

$$D^d\eta(a^*)(a-a^*)+(16KW_{a^*}/\check{\mu})|D^dH(a^*)(a-a^*)-c+H(a^*)|\leq 0$$

Proof . Denote the limsup in (68) by $\vartheta_{a,c}$ and let $\chi_{a,c}(\lambda):=\eta(a^*+\lambda(a-a^*))+16\check{\mu}^{-1}KW_{a^*}|H(a^*+\lambda(a-a^*))-c^*-\lambda(c-c^*)|$. Evidently, $\vartheta_{a,c}$ belongs to the contingent derivative $D^*\chi_{a,c}(0)$ (1). By (68), all elements in this contingent derivative has to be non-positive for all $(a,c)\in A\times C$. By general rules for calculating d.d. traps, $D^*\chi_{a,c}(0)$ (1) $\subset cl\{\omega+(16KW_{a^*}/\check{\mu})|v-c+H(a^*)|:(v,\omega)\in D(H,\eta)(a^*)(a-a^*)\}$.

In the next Corollary, the use of d.d. traps is replaced by the use of contingent derivatives.

Corollary 10. Corollary 8 holds also if (67) is replaced by: For some $\check{\mu}\in(0,1),\mu^*>0,\mu'>0$,

for all $a\in A\cap clB(a^*,\mu^*)$, for all $v\in Y$ with $|v|=\mu'$, there exist elements $a''\in A\cap clB(a,1)$, $c''\in C\cap clB(H(a^*),1)$, and $v''\in D^*H(a)(a''-a)$, such that $|v''-v-(c''-H(a^*))|\leq (1-\check{\mu})\mu'$. (68')

In particular, if (H,η) has a nonempty contingent derivative $D^*(H,\eta)(a^*)(a-a^*)$ at a^* for all $a\in A$, then

$$\sup\{\omega+(16KW_{a^*}/\check{\mu})|v-c+H(a^*)|:(v,\omega)\in D^*(H,\eta)(a^*)(a-a^*)\}$$

$$\leq 0 \tag{68"}$$

for all $(a, c) \in A \times C$.

Proof: Again (54') follow for $\mu'' = \check{\mu}/2, K' = 1/\mu'$.

D. Consequences of the controllability condition

Consider the case where $z^* = 0, C = \{0\}$. In this case, (67) reduces to: For some $\mu \in (0, 1], \mu^* > 0, \mu' > 0$,

for all $a \in A \cap clB(a^*, \mu^*)$ for all $v \in Y$ with $|v| = \mu'$, there exist an $a'' \in A \cap clB(a, 1)$ such that for all $v'' \in DH(a)(a'' - a)$,
(*) $\quad |v'' - v| \leq (1 - \mu)\mu'$. (69)

Note that when (69) holds, it also holds for any other μ', if necessary by adjusting μ, (but keeping it in $(0, 1)$). First, let $\alpha \in (0, 1)$: Then, by replacing a'' by $\hat{a}'' := \alpha a'' + (1 - \alpha)a$, from (*) in (69), we get that for all $v'' \in DH(a)(\hat{a}'' - a), |v'' - \alpha v| \leq (1 - \mu)\alpha\mu'$. Hence, μ' can be replaced by $\alpha\mu'$. This also holds if $\alpha > 1$, provided μ is replaced by μ/α : From (*) in (69), we get $|\alpha v - v''| \leq |\alpha v - v + v - v''| \leq (\alpha - 1)|v| + |v - v''| \leq (1 - \mu)|v| + (\alpha - 1)|v| \leq (\alpha - \mu)|v| = (1 - \mu/\alpha)\alpha\mu'$.

Let us now assume Lipschitz continuity of $v \to DH(a)(v)$, of rank $\leq k$, uniformly in $a \in A$. Write $a'' = a_a$. Define $k' := k'(a) := \inf\{|v''| : v'' \in DH(a)(a_a - a)\}, k'' := k''(a) := \sup\{|v''| : v'' \in DH(a)(a_a - a)\} \leq k$. Then, (*) in (69) implies an upper bound on k'' : Note that (*) implies $|v| \geq |v''| - (1 - \mu)\mu'$, so $\mu' \geq k'' - (1 - \mu)\mu'$, or $(2 - \mu)\mu' \geq k''(a)$. The last inequality is implied by $\mu' \geq k$. Now, (*) in (69) also implies a lower bound on k' : $|v| \leq |v''| + (1 - \mu)\mu'$, hence $\mu'\mu \leq |v''|$, so $\mu'\mu \leq k'(a)$. The property that, for some $\varepsilon > 0, \varepsilon \leq \inf\{|v''| : v'' \in DH(a)(a'' - a)\}$ for all $a \in A \cap clB(a^*, \mu^*)$, for some $a'' \in A \cap clB(a^*, 1)$ might be called uniform nonsingularity of $DH(a)(.)$. It is necessary for (69) to hold. If a^* is an interior point in A, and $DH(a)(.)$ is a linear operator (with bound k), and Y is finite dimensional, then this nonsingularity is sufficient for (69) to hold.

E. Comments on the control system

In the setting of the control problem, a certain type of d.d. containers naturally arises in the problem, once f, G, and ϕ have d.d. containers with respect to $x(.)$ and x, respectively.

Postulating nonempty contingent derivatives in the control theory setting would be to make assumptions of a more ad hoc nature. (We have omitted the discussion of this approach.) It is not possible to express the contingent derivative of a composition of two functions in terms of the contingent derivatives of the two functions. The solution of a differential equation may be viewed as resulting from the composition of many (in fact an infinite number) of functions. When using contingent derivatives of $\delta \to G(x^{\tilde{f}_\delta})$, and $\delta \to (G(x^{\tilde{f}_\delta}), \eta(x^{\tilde{f}_\delta}))$, (for the notation, see (28)), the results would be more general, (the controllability condition less demanding, the necessary conditions sharper), and would be closely parallel to the sharper results obtainable when using contingent derivatives in the mathematical programming setting above.

It is possible to generalize the boundedness conditions on the f's in F : Assume in the situation of Theorem 1 that the first inequality in (7) fails to hold. Assume instead that for any $f \in F$, for some $\hat{M}^f$, $|f(t,s,x^*(.))| \leq \hat{M}^f$ for all t,s. Then, for all $\hat{x}(.) \in B(0,\varsigma)$, $|f(t,s,x^*(.)+\hat{x}(.))| \leq \hat{M}^f + M^f|\hat{x}(.)| \leq \hat{M}^f + M^f\varsigma =: M_f$. Even all functions $\tilde{f} \in coF$, satisfies $|\tilde{f}(t,s,x^*(.)+\hat{x}(.))| \leq M_{\tilde{f}}$, for all t, s, $\hat{x}(.) \in B(0,\varsigma)$, for some $M_{\tilde{f}}$. Define $F_n := \{f \in F : \sup_{t,\hat{x}(.)\in B(0,\varsigma)} |f(t,s,x^*(.)+\hat{x}(.))| \leq \max[M_{f^*}, n] \text{ for a.e. } s\}$ Then F_n is essentially $\sigma-$closed and closed under switching.) Assume, in the situation of Theorem 1, that for some $\hat{n}$, and some triple $(\mu, \hat{\mu}, \mu')$, (15) is satisfied for F replaced by $F_{\hat{n}}$. Define $\Lambda_n = 64 M_\phi n \exp(M^{f^*}T)\mu^{-1}\max\{1/\mu', 1 + 1/\breve{\mu}\}$, for $\breve{\mu}$ see (17). Then, evidently, for $\Lambda = \Lambda_n$, for any $n \geq \max\{M_{f^*}, \hat{n}\}$, the conclusion $0 \geq \inf \Omega_{c,\tilde{f}}$ in Theorem 1 holds, for any $\tilde{f} \in coF_n, c \in C$. Here, $\Lambda = \Lambda_n$ enters the definition of $\Omega_{c,\tilde{f}}$, see (16), so we write $\Omega_{c,\tilde{f}} = \Omega_{c,\tilde{f},n}$. In fact, $0 \geq \inf \Omega_{c,\tilde{f},\max\{M_{f^*},\hat{n},M_{\tilde{f}}\}}$, for any $\tilde{f} \in coF, c \in C$.

A generalization to the case where $\hat{M}^f$ and M^f are integrable functions of s, rather than constants, is evidently possible.

F. Controllability results

There are actually local controllability results connected to the situation in Lemma 1. In the following two lemmas, ϕ does not appear and the optimality of $(x^*(.), f^*)$ is irrelevant:

Lemma 14. (No d.d. containers, no switching) Consider the system (4),(6),(7) and let ζ be as defined in (14). Assume that $G(x)$ is continuous in $B(x^*(T),\varsigma)$ and let c^* be a given point in C. Assume the existence of a quintuple $(K, z^*, \mu'', \hat{\mu}, \mu') \in (0,\infty) \times Y \times (0,1) \times (0,\zeta] \times (0,\infty)$ with the property that, for all $f \in clB(f^*,\hat{\mu})$, all $c \in C \cap clB(c^*,\hat{\mu})$, all $v \in Y$ with

$|v-z^*|=\mu'$, all $r>0$, there exists a triple $(\hat{f},c'',\hat{\delta})$, $\hat{f}\in clB(f^*,\zeta)$, $c''\in C$, and $\hat{\delta}\in(0,r]$, such that

$$|G(x^{\hat{f}}(T))-G(x^f(T))-\hat{\delta}v-\hat{\delta}(c''-c)|\le(1-\mu'')\hat{\delta}\mu'|v|/(|z^*|+\mu'),$$
$$\sigma(\hat{f},f)\le\hat{\delta}K|v|, \text{ and } |c''-c|\le K|v|. \tag{70}$$

Then for all $z\in clB(G(x^*(T))-c^*,\ \mu''\mu'\hat{\mu}/4K(|z^*|+\mu'))$, there exists a triple $(f,\alpha,c), f\in clB(f^*,\hat{\mu}\gamma/2)\subset F, \alpha\in[0,\hat{\mu}\gamma/2K(|z^*|+\mu')]$, $c\in C\cap clB(c^*,\hat{\mu}\gamma/2)$, such that $z+\alpha z^*=G(x^f(T))-c$, where $\gamma:=4K(|z^*|+\mu')|G(x^*(T))-c^*-z|/\mu''\mu'\hat{\mu}\le 1$.

Proof. Combine Corollary 2 with arguments in the proof of Lemma 1 and Corollary 4.

Lemma 15. (d.d. containers, switching) Consider the system (4),(6)-(11) and let ζ be as defined in (14). Let c^* be a given point in C. Assume the existence of a quintuple $(K,z^*,\mu,\hat{\mu},\mu')\in(0,\infty)\times Y\times(0,1)\times(0,\zeta]\times(0,\infty)$ with the property that, for all $f\in clB(f^*,\hat{\mu})$, all $c\in C\cap clB(c^*,\hat{\mu})$, all $v\in Y$ with $|v-z^*|=\mu'$, there exists a triple $(\tilde{f},c'',\tilde{\delta}),\tilde{f}\in coF$, $c''\in C\cap clB(c,K|v|)$, $\tilde{\delta}>0$, such that $\tilde{f}\in coB(f,\tilde{\delta}K|v|)$, and

$$\sup\{|\tilde{\delta}v-z+\tilde{\delta}(c''-c)|:z\in DG(x^f(T))(Q(T,\tilde{f},f))\}\le$$
$$(1-\mu)\tilde{\delta}\mu'|v|/(|z^*|+\mu'). \tag{71}$$

Then for all $z\in clB(G(x^*(T))-c^*,\ \mu\mu'\hat{\mu}/8K(|z^*|+\mu'))$, there exists a triple $(f,\alpha,c), f\in clB(f^*,\hat{\mu}\gamma/2)\subset F,\alpha\in[0,\hat{\mu}\gamma/2K(|z^*|+\mu')]$, $c\in clB(c^*,\hat{\mu}\gamma/2)$, such that $z+\alpha z^*=G(x^f(T))-c$, where $\gamma:=8K(|z^*|+\mu')|G(x^*(T))-c^*-z|/\mu\mu'\hat{\mu}\le 1$.

Proof: From (71), using the arguments for the implications (42)⇒(43)⇒ (27), it follows that, for all $r\in(0,(0,\varsigma/8TM\exp\{TM^f\}]$, for all $\delta>0$ small enough,

$$|G(x^{f^\delta}(T))-G(x^f(T))-\delta\tilde{\delta}v-\delta\tilde{\delta}(c''-G(x^*(T)))|\le$$
$$\delta\tilde{\delta}|v|\mu'(1-\mu/2)/(|z^*|+\mu'),$$

where $f^\delta:=\delta\tilde{f}+(1-\delta)f,\ \tilde{f}=\sum\lambda_i f_i, f_i\in F$. Next, considering only δ belonging to $(0,\zeta/2\tilde{\delta}K|v|]$, a switching combination $\hat{f}$ of f and the f_i's (with weights $(1-\sum\delta\lambda_i$ and $\delta\lambda_i)$, can be constructed such that $\sigma(\hat{f},f)\le\delta\tilde{\delta}K|v|\le\zeta/2$,$(\Rightarrow\sigma(\hat{f},f^*)\le 3\zeta/2)$, and

$$|G(x^{f^\delta}(T)) - G(x^{\hat f}(T))| \le \delta\tilde\delta|v|\mu'\mu/4(|z^*| + \mu'),$$

compare the arguments for the implication (27) $\Rightarrow$ (25) in the proof of Lemma 7 above. Using the two inequalities involving μ' above, then (70) follows for $\hat\delta = \delta\tilde\delta$ and $\mu'' = \mu/4$.

G. Time-pointwise necessary conditions, nonretarded case

A time-pointwise version of Theorem 1 can easily be given in the nonretarded case.

Then the following definition is needed: For any integrable function $z(t) : J \to Y$, (Y a Banach space), a left regular point $\tau \in (0, T]$ of $z(.)$ is a point for which $\lim_{\delta\searrow 0}\delta^{-1}\int_{[\tau-\delta,\tau]}|z(\tau) - z(s)|ds = 0$.

In this section G, the f's are functions from $J \times J \times X \to X$. Moreover, it is assumed that

> for each s, at each $x \in B(x^*(s), \varsigma) \subset X$, the map $\check x \to f(., s, \check x)$, $B(x^*(s), \varsigma) \to C(J, X)$ has a closed d.d. container $D_3 f(., s, x)(q)$ for each $q \in X$, and $q \to D_3 f(., s, x(s))(q)$ is Lipschitz continuous with rank $\le M^f$ on X. For each $x(.) \in B(x^*(.), \varsigma) \subset C(J, X)$, each $q(.) \in C(J, X), s \to D_3 f(., s, x(s))(q(s))$ is measurable and essentially separably valued, (i.e. for some separable set $X_{f,x(.),q(.)}$ $\subset C(J, X), D_3 f(., s, x(s))(q(s)) \subset X_{f,x(.),q(.)}$ for a.e. s). (72)

The following condition will also be needed: For all $f \in F$, for all $x(.) \in B(x^*(.), \varsigma)$,

> for each $s, D_3 f(., s, x(s))(q)$ is uniform in q in the set $\tilde Q(s, \tilde f, f, x(.)) :=$ $\{q(s)(s) : q(.) \in \hat Q(\tilde f, f, f, x(.))\}$ for all $\tilde f \in coF$, (for $\hat Q$, see (10)). (72')

The conditions (72), (72') are specializations of (10) to the present case. For any given $f \in F$, let Q_f be the set of finite collections $A := \{(\upsilon_i, s_i, f_i)\}_{i=1,\dots,i^*}$, $(\upsilon_i, s_i, f_i) \in (0, \infty)\times(0, T)\times F, \sum \upsilon_i \le 1$, where each s_i is a left regular point of $s \to f_i(., s, x^f(s)) - f(., s, x^f(s)) : J \to C(J, X)$, and let $Q^*(f, A)$ be the set of piecewise continuous functions $q(.) : J \to C(J, X)$ jumping only at the time points s_i, with the left and right limits of $q(.)$ at s_i satisfying

$$q(s_i+) - q(s_i-) = h^f(s_i+) - h^f(s_i-), \tag{73}$$

where

$$h^f(s) := \textstyle\sum_{s_i\le s} \upsilon_i(f_i(., s_i, x^f(s_i)) - f(., s_i, x^f(s_i))), \tag{74}$$

and where, for each $i = 0, ..., i^*$, $q(.)$ is antidifferentiable on (s_i, s_{i+1}) and satisfies

$$dq(s)/ds \in D_3 f(., s, x^f(s))(q(s)(s))ds \text{ a.e.} \tag{75}$$

and $q(0+) = 0, (s_0 = 0,\ s_{i^*+1} = T)$.

Define $Q^*(T, f, A) := \{q(T)(T) : q(.) \in Q^*(f, A)\}$. The collections A and functions $q(.)$ are used in the necessary conditions below. The equation (75) does not immediately generalize to the retarded case, at least if we want to stick to $C(J, X)$ as a (sort of) state space, (the "perturbations" $q(.)$ are outside $C(J, X)$).

The following version of Theorem 1 holds:

Theorem 4. Consider problem (4)-(9),(11),(72),(72'). Assume that $(x^*(.), f^*)$ is an optimal admissible pair in the problem. Assume that there exists a triple $(\mu'', \hat{\mu}, \check{\mu}'), \mu'' \in (0, 1), \hat{\mu} \in (0, \zeta], \check{\mu}' \in (0, \infty)$, with the property that for any $f \in clB(f^*, \hat{\mu})$, for any $v \in Y$ with $|v| = \check{\mu}'$, there exists a pair $(A, c'') \in Q_f \times (C \cap clB(G(x^*(T)), 1)$ such that

$$\sup\{|v - z + c'' - G(x^*(T))| : z \in DG(x^f(T))(Q^*(T, f, A))\} \leq (1 - \mu'')\check{\mu}'. \tag{76}$$

Then the following necessary condition holds: For any $A = \{v_i, s_i, f_i\} \in Q_{f^*}$, and for any $c \in C$, the inequality $0 \geq \inf \Omega_{c,A}$ holds, where $\Omega_{c,A} :=$

$$\{w - \Lambda^*|v| : (w, v) \text{ belongs to a triple } (w, v, z) \text{ satisfying } z \in Q^*(T, f^*, A), w \in D\phi(x^*(T))(z), v \in DG(x^*(T)(z) - c + G(x^*(T))\}, \tag{77}$$

and where $\Lambda^* := 128 M_\phi M \exp(M^{f^*} T)\mu''^{-1} \max\{1/\check{\mu}', 1 + 1/\tilde{\mu}\}$, $\tilde{\mu} := \min\{\hat{\mu}, \mu''\check{\mu}'/16\}$.

To prove Theorem 4 it is necessary to show that (75) implies (15). Two lemmas are needed:

Lemma 16. Assume in Lemma 10, that the initial value of $q(0)$ is specified to be different from $q_0(0)$. Then there exists an antidifferentiable function $q(.)$ with the specified initial value $q(0)$ such that $dq(t)/dt \in A(t, q(.))$ a.e., and such that $|q_0(t) - q(t)| \leq \xi^*(t)$, where $\xi^*(t) := 2(|q(0) - q_0(0)| + \int_0^t \lambda_0(\rho)d\rho) \exp(\int_0^t 2\kappa(\rho)d\rho))$, provided $\xi^*(t) < \varepsilon$ for all t.

Proof. For $t \in [-1, 0]$, define $A(t, q(.)) = \{0\}$, $\kappa(t) = 0, \lambda_0(t) = |q(0) - q_0(0)|, q_0(t) = q(0) + (q_0(0) - q(0))(t + 1)$, and apply Lemma 10 for $[0, T]$ replaced by $[-1, T]$.

Lemma 17. Let $f \in clB(f^*, \hat{\mu})$ and $A = \{\upsilon_i, s_i, f_i\}_{i=1,...,i^*} \in Q_f$ be given. For all $\varepsilon > 0$, there exist a $\delta_A > 0$ and a $f^A \in coF$, such that for all $\check{q}(.) \in Q(f^A, f)$, there exists a $q(.) \in Q^*(f, A)$ such that $|\delta_A q(T) - \check{q}(T)| < \delta_A \varepsilon$.

Proof: Let $\varepsilon > 0$. Let $s^k, k = 1, ..., k^*$ be the set of distinct time points in $\{s_i\}$. Define $\bar{\upsilon}^k = \sum_{i \in \{j: s_j = s^k\}} \upsilon_i$, $\tilde{f}_k = \sum_{i \in \{j: s_j = s^k\}} (\upsilon_i/\bar{\upsilon}_k) f_i \in coF, f_\delta^A :=$ $f(1 - \sum_k 1_{[s^k - \delta\bar{\upsilon}^k, s^k]}) + \sum_k \tilde{f}_k 1_{[s^k - \delta\bar{\upsilon}^k, s^k]}, h_\delta(\rho) =$ $\delta^{-1}[f_\delta^A(., \rho, x^f(\rho)) - f(., \rho, x^f(\rho))]$. For $\delta^A > 0$, δ^A small enough, f_δ^A belongs to co$F$ for $\delta \in (0, \delta^A]$, (the intervals become disjoint). By left regularity, for $\delta := \delta_A \in (0, \delta^A]$ small enough,

$$|h^f(s^k+) - h^f(s^k-) - \textstyle\int_{s^{k-1}}^{s^k} h_\delta(s)ds| \leq \varepsilon. \tag{78}$$

Let $\check{q}(.) \in Q(f_\delta^A, f)$ and observe that (*) in (10) implies that $|\check{q}(s)| \leq \int_{s^1 - \delta\bar{\upsilon}^1}^s (|\delta h_\delta(\rho)| + M^f |\check{q}(.)|_\rho) d\rho$, $s \in (s^1 - \delta\bar{\upsilon}^1, s^1), (h_\delta(s) = 0$ and $|\check{q}(s) = 0$ on $(0, s^1 - \delta\bar{\upsilon}^1])$. Now, $|\delta h_\delta(s)|ds \leq 2M$, and $\int_{s^1 - \delta\bar{\upsilon}^1}^{s^1} |\delta h_\delta(s)|ds \leq \delta\bar{\upsilon}^1 2M$. By Gronwall's inequality, $|\check{q}(s)| \leq (\int_{s^1 - \delta\bar{\upsilon}^1}^s |\delta h_\delta(\rho)|d\rho) \exp(M^f s^1) \leq$ $2M\delta\bar{\upsilon}^1 \exp(M^f s^1)$. Hence, $|\check{q}(s)/\delta - \int_{s^1 - \delta\bar{\upsilon}^1}^s h_\delta(\rho)d\rho| \leq 2M\bar{\upsilon}^1 \exp(M^f s^1) +$ $2M\bar{\upsilon}^1 =: \gamma$. Define $q_{k-1}(s) := \check{q}(s)/\delta - \int_{s^{k-1}}^s h_\delta(\rho)d\rho$, $s \in (s^{k-1}, s^k)$, $(s^0 = 0, s^{k^*+1} = T)$, and note that, for $s \leq s^1, |q_0(s)| \leq \varepsilon$, if $\delta \leq \varepsilon/\gamma$. Assume that the δ introduced in connection with (78) was chosen so small that $\delta \in (0, \varepsilon/\gamma]$. By (10) and (12) (and linear homogeneity), $(d/ds)q_{k-1}(s) \in D_3 f(., s, x^f(s))(q_{k-1}(s))$ on (s^{k-1}, s^k), Define $q(.) = 0$ on $(0, s^1)$. By induction on k, assume that $q(.) \in Q^*(f, A)$ is defined such that $|q(s) - q_{k'-1}(s)| \leq \varepsilon k' \exp(M^f s)$ on $(s^{k'-1}, s^{k'}), k' \leq k$. (This inequality holds on $(0, s^1)$.) Define $q(s^k+) = q(s^k-) + h^f(s^k+) - h^f(s_k-)$. Then, by (78), $|q(s^k+) - q_k(s^k)| =$ $|q(s^k+) - \check{q}(s^k)/\delta| \leq |q(s^k-) + h^f(s^k+) - h^f(s_k-) - \check{q}(s^k)/\delta]| \leq |q(s^k-)+$

$\int_{s^{k-1}}^{s^k} h_\delta(\rho)d\rho - \check{q}(s^k)/\delta]|+\varepsilon = |q(s^k-)-q_{k-1}(s^k-)|+\varepsilon \leq \varepsilon+\varepsilon k\exp(M^f s^k) \leq \varepsilon(k+1)\exp(M^f s^k)$.

By Lemma 16, (letting $q_k(.)$ play the role of $q_0(.)$), there exists a $q(.) \in Q^*(f,A)$ defined on (s^k, s^{k+1}) with initial value $q(s^k+)$ as specified above, such that $|q(s)-q_k(s)| \leq \varepsilon(k+1)\exp(M^f s^k)\exp\{M^f(s-s^k)\} = \varepsilon(k+1)\exp\{M^f s\}$, and the induction is complete. In particular,
$|q(s^{k^*+1}) - \check{q}(s^{k^*+1})/\delta| = |q(s^{k^*+1}) - \check{q}(s^{k^*+1})/\delta + \int_{s^{k^*}}^{s^{k^*+1}} h_\delta(\rho)d\rho| \leq \varepsilon(k^*+1)\exp(M^f s^{k^*+1})$, ($h_\delta(s)$ vanishes for $s > s^{k^*}$). □

Lemma 17 implies that for all $\varepsilon > 0$, there exists a $\delta := \delta_A > 0$, such that $\delta^{-1}Q(T,f^A,f) \subset Q^*(T,f,A) + B(0,\varepsilon/M')$, where $M' := \max\{M_G, M_\phi\}$. By Lipschitz continuity, this implies that both for $D^f(.) = DG(x^f(T))(.)$ and $D^f(.) = D\phi(x^f(T))(.)$,

$$D^f(\delta^{-1}Q(T,f^A,f)) \subset D^f(Q^*(T,f,A)) + B(0,\varepsilon), \tag{79}$$

Thus, $DG(x^f(T))(Q(T,f^A,f)) \subset DG(x^f(T))(\delta Q^*(T,f,A)) + B(0,\delta\varepsilon)$. For

$\varepsilon = \mu''\check{\mu}'/2$, from this inclusion and (76), (15) follows, for $\mu = \mu''/2, \mu' = \check{\mu}', \tilde{\delta} = \delta := \delta_A$. Observe that if Γ is the map $(w,v) \to w - \Lambda|v|$, then, by (79) and Lipschitz continuity of Γ of rank $\leq 1+\Lambda$, it follows that (for $f = f^*), \Gamma^* \subset \Gamma^{**} + B(0,(1+\Lambda)\varepsilon)$, where
$\Gamma^* := \cup_{w\in\delta^{-1}Q(T,f^A,f^*))}\Gamma(D\phi(x^*(T))(w), DG(x^*(T))(w))$ and
$\Gamma^{**} := \cup_{w\in Q^*(T,f^*,A)}\Gamma(D\phi(x^*(T))(w), DG(x^*(T))(w))$

The conclusion in Theorem 4 follows from this inclusion and the arbitrariness of ε: $\inf\{\Gamma^{**} + B(0,(1+\Lambda)\varepsilon)\} = \inf\Gamma^{**} - (1+\Lambda)\varepsilon \leq \inf\Gamma^* \leq 0$. □

Instead of the "uniformities" in (72') and (11), alternatively, the following assumptions also yield the conclusion of Theorem 4.

> Given any finite collection $\hat{A} := \{v_i, s_i, w_i\}_{i=1,\ldots,i^*}, s_i \in (0,T)$, $w_i \in X, v_i > 0, \sum_i v_i \leq 1$, let $\tilde{Q}(f,\hat{A})$ be the set of piecewise continuous functions $y(.) : J \to C(J,X)$ jumping only at the s_i's, $y(0) = 0, y(s_i+) - y(s_i-) = \sum_{s_j\leq s_i} v_j w_j - \sum_{s_j<s_i} v_j w_j$, $y(.)$ antidifferentiable between the jump points, with $dy(s)/ds \in D_3 f(.,s,x(s))(y(s)(s))ds$ a.e.
> For each collection $\hat{A}$, each $\check{x} \in B(0,\varsigma)$, and each s, it is assumed that $D_3 f(.,s,x^*(s)+\check{x})(y)$ is uniform in y in the set

$$\{y(s)(s) : y(.) \in \tilde{Q}(f, \hat{A})\}. \tag{80}$$

For each f and collection $\hat{A}$, at each $x' \in B(x^*(T), \varsigma)$,
$x \to G(x)$ and $\to \phi(x)$ are assumed to have d.d. containers
in all directions v, being uniform in
$v \in \{y(T)(T) : y(.) \in \tilde{Q}(f, \hat{A})\}$ (81)

Placing slightly stronger conditions on the perturbations points s_i, (still, in a sense, for any given f, they will exist a.e., see Seierstad (1997)), a variant of Theorem 4 is obtainable. Redefine Q_f to consist of collections $\{(\upsilon_i, s_i, f_i)\}_{i=1,...,i^*}, (\upsilon_i, s_i, f_i) \in (0, \infty) \times (0, T) \times F, \sum \upsilon_i \leq 1$, such that each s_i is a left regular point of the indicator function of the set $N_i := \{s : f_i(., s, x^f(s)) \neq f(., s, x^f(s))\}$ and such that s_i is a point of continuity of the function $s \to f_i(., s, x^f(.)) - f(., s, x^f(.))$ restricted to N_i. Also redefine Λ^* to equal $32 M_\phi M \exp(M^{f^*} T) \mu''^{-1} \max\{1/\breve{\mu}', 1 + 1/\tilde{\mu}\}$, $\tilde{\mu} := \min\{\hat{\mu}, \mu''\breve{\mu}'/4\}$.

Theorem 5. For Q_f and Λ^* as just redefined, Theorem 4 holds for (72'),(11) replaced by (80),(81).

Proof: The proof of this theorem can - and will - be reduced to the case of ordinary differential equations, (see above).

A lemma is needed:

Lemma 18. Let $g(t, x) : J \times X \to X$ be measurable in t for each $x \in X$. Let $(s^i, x^i) \in (0, T) \times X, i = 1, ..., k$, be given, $s^1 < s^2 < ... < s^k$, and let $t \to x(t; s^1, ..., s^k, x^1, ..., x^k) =: x(t; ...)$ be piecewise continuous with jumps at the s^i's satisfying

$$x(s^i+; ...) - x(s^i-; ...) = x^i \tag{83}$$

and such that $t \to x(t; ...)$ is antidifferentiable on each (s^i, s^{i+1}), satisfying

$$dx(s)/ds = g(s, x(s; ...)) \text{ a.e.}, \; x(0) = x_0, \tag{83'}$$

(so $s \to g(s, x(s; ...))$ is integrable, by assumption). Assume also that for some $\lambda > 0$, for all $t, x \to g(t, x(t; ...) + x)$ is Lipschitz continuous of rank $\leq \kappa_g(t)$ in $B(0,\lambda) \subset X$, $\kappa_g(.)$ integrable. Moreover, assume that $x \to g(t, x(t; ...) + x)$ has a closed d.d. container $D_2g(t, x(t; ...))(q)$ at 0 for all t, all $q \in X$, which is Lipschitz continuous in q of rank $\leq \kappa_g(t)$, and that for each function $q(.) \in C(J, X), t \to D_2g(t, x(t; ...))(q(t))$ is measurable and essentially separably valued. Let $Q(s^1, ..., s^k, v^1 ..., v^k)$ be the set of piecewise continuous functions $q(.)$, such that $q(.)$ is antidifferentiable on each (s^i, s^{i+1}),with

$$dq(s)/ds \in D_2g(s, x(s; ...))(q(s)) \text{ a.e.,}$$
$$q(s^i+) - q(s^i-) = v^i \in X, q(0) = 0. \quad (83\text{''})$$

Assume that for a.e. t, $D_2g(t, x(t; ...))(q)$ is uniform in $q \in \{q(t) : q(.) \in Q(s^1, ..., s^k, v^1 ..., v^k)\}$. Then, a piecewise continuous solution $t \to x(t; s^1, ..., s^k, z^1, ..., z^k)$ (antidifferentiable between the jump points s^i) through $(0, x_0)$ of (83),(83') also exists for jump values z^i near x^i, $i = 1, ..., k$, (i.e. for x^i in (83) replaced by such z^i's), and for $t > s^k, (z^1, ..., z^k) \to x(t; s^1, ..., s^k, z^1 ..., z^k)$ has a d.d. container with respect to the vector $(z^1 ..., z^k) \in X^k$ in direction $(v^1, ..., v^k)$, denoted $D_{(z^1,...,z^k)}x(t; s^1, ..., s^k, x^1, ..., x^k)(v^1, ..., v^k)$ that equals $\{q(t) : q(.) \in Q(s^1, ..., s^k, v^1 ..., v^k)\}$

Proof: The proof follows by applying Lemma 11 to the following system, with state space X^{k+1} :

$$dx_i/dt = g(t, \textstyle\sum_{j \leq i} x_j(t))1_{[s^i, s^{i+1})}, x_i(0) = z^i, \; i = 0, ..., k, \quad (84)$$

$(s^0 = 0, s^{k+1} = T, z^0 = x_0)$. Then, for $\sum_{j=0}^k x_j(t)1_{[s^j,T]}(t) =:$ $x(t;\; s^1, ..., s^k, z^1, ..., z^k) =: x(t)$, on $(s^i,\; s^{i+1}), dx/dt = dx_i/dt = g(t, x(t))$, and $x(s^i+) - x(s^i-) = \sum_{j \leq i} x_j(s^j) - \sum_{j \leq i-1} x_j(s^j) = x_i(s^i) = z^i$. □

We now turn back to the proof of Theorem 5, by first proving that,

> for any $f \in clB(f^*, \hat{\mu})$, for any $A = \{(\upsilon_i, s_i, f_i)\} \in Q_f$, for all $\varepsilon > 0$, there exists a $\tilde{\delta} \in (0, \zeta/2]$, such that for any $\delta \in (0, \tilde{\delta}]$, there exists a $f_{A,\delta} \in F$, $f_{A,\delta} \in clB(f, \delta)$, such that $\delta^{-1}(G(x^{f_{A,\delta}}(T)) - G(x^f(T)), \phi(x^{f_{A,\delta}}(T)) - \phi(x^f(T))) \in \Psi^* + B(0, \varepsilon)$, where $\Psi^* := \cup_{y \in Q^*(T,f,A)} DG(x^f(T))(y) \times D\phi(x^f(T))(y)$ (85)

(For Q^*, see (73)-(75)). To prove (85), let (f, v, A, c'') be a quadruple for which (76) holds, $A = \{(\upsilon_i, s_i, f_i)\}_{i=1,...,i^*}$. Define $w^k := \bar{\upsilon}^k(\tilde{f}_k(s^k, x^f(s^k)) - f(s^k, x^f(s^k)))$, (for $\bar{\upsilon}^k, s^k$ and $\tilde{f}_k$, see the proof of Lemma 17). Note first that,

by Seierstad (1997), Section 5, for any $\varepsilon > 0$, for some $\delta^* > 0$, there exists a function $f_{A,\delta} \in clB(f,\delta) \cap F$, such that $|x(T; s^1, ..., s^k, \delta\bar{\upsilon}^1 w^1, ..., \delta\bar{\upsilon}^{k^*} w^{k^*}) - x^{f_{A,\delta}}(T)|/\delta \leq \varepsilon/2M', M' = \max\{M_G, M_\phi\}$, when $\delta \in (0, \delta^*]$. From the preceding Lemma, by shrinking δ^* if necessary, for all $\delta \in (0, \delta^*]$,
$[G(x(T; s^1, ..., s^k, \delta\bar{\upsilon}^1 w^1, ..., \delta\bar{\upsilon}^{k^*} w^{k^*})) - G(x^f(T))]/\delta \in$
$DG(x^f(T))(Q^*(T, f, A)) + B(0, \varepsilon/2)$, and
$[\phi(x(T;\ s^1, ..., s^k, \delta\bar{\upsilon}^1 w^1, ..., \delta\bar{\upsilon}^{k^*} w^{k^*})) - \phi(x^f(T))]/\delta) \in$
$D\phi(x^f(T))(Q^*(T, f, A)) + B(0, \varepsilon/2)$. Using the last inequality, (85) follows.

Below, we shall apply Lemma 5 (Lemma 2), for $\mu = \mu''/2, K' = 1/\mu', \mu' = \breve{\mu}'$. Evidently, (25) follows from (76), by letting $\varepsilon = \mu''\breve{\mu}'/2$ and using (85).

Let us next prove that, for any given $A = \{\upsilon_i, s_i, f_i\}_{i=1,...,i^*}$ in Q_{f^*},

for all $\varepsilon' > 0$, there exists a $\delta'' > 0$, such that $f := f_{A,\delta}$ satisfies
$[\bar{\iota}_{f,\delta c+(1-\delta)c^*} - \bar{\iota}_{f^*,c^*}]/\delta \leq \varepsilon'$, for all $\delta \in (0, \delta''], c \in C$. (86)

where $c^* := G(x^*(T))$ and $\bar{\iota}_{f,c} := \phi(x^f(T)) - \Lambda^*|G(x^f(T)) - c|$, ($\Lambda^*$ as defined in connection with Theorem 5.)

To prove (86), let $A = \{\upsilon_i, s_i, f_i\}_{i=1,...,i^*} \in Q_{f^*}$ be given and define $\tilde{M} := \max\{M^{f^*}, \max\{M^{f_i}\}\}$. We can let $\beta^{f_{A,\delta}} = \tilde{M}\sigma(f_{A,\delta}, f^*) + (T - \sigma(f_{A,\delta}, f^*))M^{f^*}$. Let $\varepsilon' > 0$ and let $\Theta := M_G 2M \exp(TM^{f^*}) + |c - G(x^*(T))|$ and $\Phi := 32\mu''^{-1}M_\phi M \max\{1/\breve{\mu}', 1 + 1/\tilde{\mu}\}$. Note that $0 \leq \beta^{f_{A,\delta}} - TM^{f^*} = (\tilde{M} - M^{f^*})\sigma(f_{A,\delta}, f^*) \leq \tilde{M}\delta$. When $\delta \in (0, \delta'], \delta' := \ln\{1 + \varepsilon'/\tilde{M}\Phi\Theta \exp(TM^{f^*})\}$, then $\exp(\beta^{f_{A,\delta}}) - \exp(TM^{f^*}) = \exp(TM^{f^*})[\exp(\beta^{f_{A,\delta}} - TM^{f^*}) - 1] \leq \varepsilon'/\Phi\Theta$. Now, for $f = f_{A,\delta}$, Lemma 5 (Lemma 2) provides the inequality $0 \geq [\iota^*(f, \delta c + (1-\delta)c^*) - \iota^*(f^*, c^*)]/\delta$, $(\delta \in (0, \breve{\mu}/2]$. Furthermore, note that for $\delta \in (0, \zeta]$, (24)(i) gives, for $f \in clB(f^*, \delta)$, that $|G(x^f(T)) - G(x^*(T))|/\delta \leq M_G 2M \exp(TM^{f^*})$ and hence $|G(x^f(T)) - G(x^*(T)) - \delta c + \delta G(x^*(T))|/\delta \leq \Theta$. Thus, replacing the term $\beta^f = \beta^{f_{A,\delta}}$ in $\iota^*(f, \delta c + (1-\delta)c^*)$ by TM^{f^*} introduces an error smaller than $\Phi\Theta(\exp(\beta^{f_{A,\delta}}) - \exp(TM^{f^*}))$, hence yields the weak inequality in (86), (let $\delta'' = \min\{\delta', \breve{\mu}/2\}$).

The following argument shows that the conclusion of Theorem 5 follows from (85) and (86). Let A and c be given. By (86), for any $\varepsilon > 0$, for $\delta > 0$, small enough, the inequality $[\bar{\iota}_{f,\delta c+(1-\delta)c^*} - \bar{\iota}_{f^*,c^*}]/\delta \leq \varepsilon/3$ holds. By (85), for δ small enough, for some $z \in Q^*(T, f, A), v' \in DG(x^*(T))(z)$ and $w \in D\phi(x^*(T))(z)$, the inequalities $\Lambda^*|v' - [G(x^{f_{A,\delta}}(T)) - G(x^*(T))]/\delta| \leq \varepsilon/3$ and $|w - [\phi(x^{f_{A,\delta}}(T)) - \phi(x^*(T))]/\delta)| \leq \varepsilon/3$ hold. The three inequalities involving $\varepsilon/3$ yield $w - \Lambda^*|v' - c + G(x^*(T))| \leq \varepsilon$.

H. Necessary conditions under simplifying assumptions

Under the Lipschitz continuity assumptions of (7) and (8), the following condition implies (10) and (11):

$$x \to (\phi(x), G(x)) \text{ and } x(.) \to f(t,s,x(.)) \text{ have directional derivatives in all directions, for all } x \in B(x^*(T),\varsigma), \text{ respectively, all } x(.) \in B(x^*(.),\varsigma). \tag{87}$$

When (87) holds, let $q_{\tilde{f},f}(.)$ be the continuous solution of

$$q(t) = \int_0^t \tilde{f}(t,s,x^f(\to s)) - f(t,s,x^f(\to s))ds + \int_0^t D^d f(t,s,x^f(\to s))(q(\to s))ds,$$

(we write D^d for directional derivatives).

Consider next the condition that for some pair $(\lambda,\varepsilon), \lambda > 0, \varepsilon > 0$,

$$clB(0,\varepsilon) \subset cl\{D^d G(x^f(T))(q_{\tilde{f},f}(T)) - (c - G(x^*(T))) : \tilde{f} \in coF, c \in C \cap clB(G(x^*(T)),1)\} \text{ for all } f \in clB(f^*,\lambda). \tag{88}$$

Evidently, (88) implies (15), and the inequality $0 \geq \inf \Omega_{c,\tilde{f}}$ in Theorem 1 then yields the conclusion in the following Corollary.

Corollary 11. Let $(x^*(.), f^*)$ be an optimal pair in problem (4)-(9),(87), (88). Then, for any $\tilde{f} \in coF$, and for any $c \in C$, for Λ as in (17),

$$D^d\phi(x^*(T))(q_{\tilde{f},f^*}(T)) - \Lambda|D^d G(x^*(T))(q_{\tilde{f},f^*}(T)) - c + G(x^*(T))| \leq 0. \tag{89}$$

The following corollary evidently follows from Corollary 1 in Section 4.

Corollary 12. Assume that all f in F are independent of t and that $f(s,x(\to s)) = f(s,x(s))$, i.e. there is no history dependence. (Then (4) is equivalent to an ordinary differential equation in X, in which case solutions $x(t)$ by definition are required to be antidifferentiable.) Assume also that (4)-(9) hold and that all $f \in F$, for all s, have bounded linear Gâteaux derivatives with respect to x at each point in $B(x^*(t),\varsigma)$. Assume, furthermore, that ϕ and G have bounded linear Gâteaux derivatives in $B(x^*(T),\varsigma)$, and that (18) and (19) are satisfied. Assume also that $x \to f_x(t, x^*(t)+x)[q]$

is continuous in $B(0,\varsigma) \subset X$, for any $t, f \in F, q \in X$, and similarly, that $x \to G_x(x^*(T)+x)[q']$ is continuous in $B(0,\varsigma) \subset X$ for any $q' \in X$. Then the following maximum principle holds: For some non-negative number λ_0, some bounded linear functional $\lambda^* \in Y^*$, $(\lambda^*, \lambda_0) \neq 0$, for any $f \in F$, for t not in a null set N_f,

$$\langle f(t, x^*(t)), p(t)\rangle \leq \langle f^*(t, x^*(t)), p(t)\rangle. \tag{91}$$

Here $p(t)$ is a continuous, weak* solution in X^*, satisfying

$$dp(t)/dt = -[f_x^*(t, x^*(t))]^* p(t) \text{ a.e.}, \tag{92}$$

and

$$p(T) = [\phi_x(x^*(T))]^*\lambda_0 + [G_x(x^*(T))]^*\lambda^*,\ \langle C - G(x^*(T)), \lambda^*\rangle \leq 0, \tag{93}$$

where $[\ ]^*$ means adjoint. (That $p(.)$ is a weak* solution means that $\langle \hat{x}, p(t)\rangle$ is absolutely continuous for each $\hat{x} \in X$, and $d\langle \hat{x}, p(t)\rangle/dt =$ $\langle -f_x(t, x^*(t))\hat{x}, p(t)\rangle$ a.e. for all $\hat{x} \in X$).

I. Comments on the measurability condition in (9)

To see that (9) allows applications to weakly nonlinear evolution equations, note the following result:

Let $Q(t), t \geq 0$ be a strongly continuous semigroup in X, and let $h(s)$ be a measurable function from J into X. Define $Q(t) = I$, for $t < 0$. Let $h^*(s) := t \to Q(t-s)h(s) : J \to X$, so $h^*(.) : J \to C(J, X)$. Then $h^*(.)$ is measurable.

To show this, for any $\varepsilon > 0$, first choose a real-valued step function $\alpha(s)$ such that meas$(J_\varepsilon) < \varepsilon$, where $J_\varepsilon := \{s : |h(s) - h(\alpha(s))| > \varepsilon/2M\}$, $M := \sup_{t\in J} |Q(t)|$. Next, choose $\delta > 0$ such that $|(Q(\sigma) - I)x| \leq \varepsilon/2M$ for all $\sigma \in [0, \delta]$, all x in the finite set $h(\alpha(J))$. Define the piecewise constant function $\beta(s)$ by $\beta(s) = n\delta$ on $(n\delta, (n+1)\delta]$, $n = 0, 1, 2, ...$ .Then $\delta \geq s - \beta(s) \geq 0$. Define $\sigma := \max\{0, t-\beta(s)\} - \max\{0, t-s\}, t, s \in J$. Then σ belongs to $[0, \delta]$, (check the cases $t - \beta(s) > 0, t - s \leq 0$ and $t-\beta(s) > 0, t-s > 0$). Then, for all $t, s \in J$, for all x in the finite set $h(\alpha(J))$, $|[Q(t-\beta(s)) - Q(t-s)]x| = |[Q(\max\{0, t-\beta(s)\}) - Q(\max\{0, t-s\})]x| =$ $|(Q(\sigma) - I)Q(\max\{0, t-s\})x| < \varepsilon/2$. From this it follows that for all $t \in J, s \in J\backslash J_\varepsilon$,we have $|Q(t-s)h(s) - Q(t-\beta(s))h(\alpha(s))| \leq |Q(t-s)h(s)-$ $Q(t-s)h(\alpha(s))| + |Q(t-s)h(\alpha(s)) - Q(t-\beta(s))h(\alpha(s))| \leq \varepsilon/2 + \varepsilon/2$, and the claimed measurability follows.

Assume that $a(s,x)$ is measurable in $s \in J$, and locally Lipschitz continuous in $x \in X$. Assume also that the directional derivative $D^d_x a(s,x)(v)$ exists for all s, x, v. Then the function $b(s,x) : J \times X \to C(J,X)$ given by $b(s,x)(t) {:=:} t \to Q(t-s)a(s,x)$ has the directional derivative
$Q(.-s)D^d_x a(s,x)(v)$, and for any continuous $x(.)$, $D^d_x a(s,x(s))(v)$ and
$Q(.-s)D^d_x a(s,x(s))(v)$ are measurable in s.

Note that the measurability condition in (9) follows from the following assumption.

$$\begin{array}{l}\text{For any } x(.) \text{ in } C(J,X) \text{ and any } f \in F,\\ f(t,s,x(.)) \text{ is continuous in } t \text{, uniformly in } s \text{, i.e. for all } t \in J,\\ \text{for all } \varepsilon > 0 \text{, there exists a } \delta > 0 \text{, such that if } t' \in J \text{ and } |t'-t| < \\ \delta \text{, then } |f(t',s,x(.)) - f(t,s,x(.))| < \varepsilon \text{ for all } s \in J.\end{array} \tag{94}$$

Let us prove that for any $y(.) \in C(J,X)$,$f(.,s,y(.)) : J \to C(J,X)$ is measurable: For any natural number n, by the equicontinuity in (9), (which actually is uniform in t in the compact set J), there exists a piecewise constant function $\beta_n(t) : J \to J$, such that $\sup_t |f(\beta_n(t),s,y(.)) - f(t,s,y(.))| \leq 1/3n$ for all s. Moreover, a step function $\alpha_n(s) : J \to J$, exists such that $meas(J \backslash A^n) \leq 1/n$, where
$A^n := \{s \in J : \sup_t |f(\beta_n(t),\alpha_n(s),y(.)) - f(\beta_n(t),s,y(.))| \leq 1/3n\}$.
Note that $\sup_t |f(\beta_n(t),\alpha_n(s),y(.)) - f(t,\alpha_n(s),y(.))| \leq 1/3n$ for all s. Using $f(t,\alpha_n(s),y(.)) - f(t,s,y(.)) = f(t,\alpha_n(s),y(.)) - f(\beta_n(t),\alpha_n(s),y(.)) +$ $f(\beta_n(t),\alpha_n(s),y(.)) - f(\beta_n(t),s,y(.)) + f(\beta_n(t),s,y(.)) - f(t,s,y(.))$, it follows that, for all $s \in A^n$, $\sup_t |f(t,\alpha_n(s),y(.)) - f(t,s,y(.))| \leq 1/n$. So the sequence of step functions of s, $f(.,\alpha_n(s),y(.))$ converges to $f(.,s,y(.))$ in measure in J, in sup-norm in $C(J,X)$, which yields measurability.

J. Examples of d.d. containers

(i) Let $g : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz continuous with Lipschitz rank K, and let $f : C(J,\mathbb{R}^n) \to C(J,\mathbb{R})$ be defined by $f(x(.)) := s \to g(x(s)), x(.) \in C(J,\mathbb{R}^n)$. Let $h(\lambda,s) := [f(x(s)+\lambda z(s)) - f(x(s))]/\lambda$. Define, for any $\delta > 0$, $\beta_\delta(s) := \mathrm{limsup}_{\lambda \searrow 0} \beta_{\delta,\lambda}(s)$, where $\beta_{\delta,\lambda}(s) := \sup\{h(\lambda,s') : |s'-s| \leq \delta\}$, and $\alpha_\delta(s) := \mathrm{liminf}_{\lambda \searrow 0} \alpha_{\delta,\lambda}(s)$, where $\alpha_{\delta,\lambda}(s) := \inf\{h(\lambda,s') : |s'-s| \leq \delta\}$. Let J' be a finite set such that $J \subset J' + B(0,\delta)$, and choose a function $\phi(s') : J \to J'$, such that $|s' - \phi(s')| < \delta$ for all s'. Then, for any $\varepsilon > 0$, for some $\delta' > 0$, $\beta_{\delta,\lambda}(s) \leq \beta_\delta(s) + \varepsilon$, for $\lambda \in (0,\delta']$, all $s \in J'$. Moreover, for all $s' \in J$, for all $\lambda \in (0,\delta']$, $h(\lambda,s') \leq \beta_{\delta,\lambda}(\phi(s')) \leq \beta_\delta(\phi(s')) + \varepsilon$.

Similarly, for all $\varepsilon > 0$, for some $\delta'' > 0$, for all $s' \in J$, for all $\lambda \in (0, \delta'']$, $h(\lambda, s') \geq \alpha_{\delta,\lambda}(\phi(s')) \geq \alpha_\delta(\phi(s')) - \varepsilon$. Let $\alpha^\delta(s)$ and $\beta^\delta(s)$ be any pair of continuous functions such that $\alpha^\delta(s) \leq \alpha_\delta(\phi(s))$ and $\beta_\delta(\phi(s)) \leq \beta^\delta(s)$ for all s, and let $Df(x(.))(z(.))$ be the set of continuous functions $c(.)$ such that $\alpha^\delta(s) \leq c(s) \leq \beta^\delta(s)$. By the above results, $Df(x(.))(z(.))$ is a d.d. container.

(ii) If we change the example in (i) by working in L_p-space, $p \in [1, \infty)$, (i.e. $x(.), z(.)$ and all $c(.)$ belong to $L_p(J, \mathbb{R}^n)$, $f : L_p(J, \mathbb{R}^n) \to L_p(J, \mathbb{R})$), the inequalities in the definition of $Df(x(.))(z(.))$ can be sharper: Now, we can take $Df(x(.))(z(.)) := \{c(.) \in L_p(J, \mathbb{R}) : \hat{\alpha}(s) \leq c(s) \leq \hat{\beta}(s)\}$, where $\hat{\alpha}(s) :=\text{liminf}_{\lambda \searrow 0} h(\lambda, s)$ and $\hat{\beta}(s) :=\text{limsup}_{\lambda \searrow 0} h(\lambda, s)$. (Note that $\lambda \to h(\lambda, s)$ is continuous, so $\hat{\alpha}(s)$ and $\hat{\beta}(s)$ are measurable.) The set $Df(x(.))(z(.))$ is a d.d. container for the following reason: Note that $|h(\lambda, s)| \leq |z(s)|K|$ and $|\hat{\beta}(s)| \leq |z(s)|K$. For all s, $\max\{0, h(\lambda, s) - \hat{\beta}(s)\} \searrow 0$ when $\lambda \searrow 0$. Then, by dominated convergence, $\int_J [\max\{0, h(\lambda, s) - \hat{\beta}(s)\}]^p ds \searrow 0$,when $\lambda \searrow 0$. Hence, for any $\varepsilon' > 0$, for some $\delta' > 0$, $h(\lambda, s) \leq \hat{\beta}(s) + \hat{v}_\lambda(s)$, when $\lambda \in (0, \delta']$, $|\hat{v}_\lambda(.)|_p < \varepsilon'$, where $\hat{v}_\lambda(s) := \max\{0, h(\lambda, s) - \hat{\beta}(s)\}$. Similarly, for some $\delta'' > 0, h(\lambda, s) \geq \hat{\alpha}(s) + \check{v}_\lambda(s)$, when $\lambda \in (0, \delta'']$, for some nonpositive L_p-function $\check{v}_\lambda(.)$, with $|\check{v}_\lambda(.)|_p < \varepsilon'$.

(iii) Let C' be a compact subset of $\mathbb{R}^n$ and let $L(x, y) : \mathbb{R}^m \times C' \to \mathbb{R}$ be continuous in y and Lipschitz continuous in x with rank $\leq K$, uniformly in y. Given any elements $\hat{x}$ and z in $\mathbb{R}^m$, if the family of functions of $y, \{\hat{b}(\lambda, y)\}_{\lambda \in (0,1]}$, $\hat{b}(\lambda, y) := [L(\hat{x} + \lambda z, y) - L(\hat{x}, y)]/\lambda$ is equicontinuous in y, then $E(x) := L(x, .) : \mathbb{R}^m \to C(C', \mathbb{R})$ has a d.d. container in $C(J, X)$, at $\hat{x}$, given by $D_1^* L(\hat{x}, .)(z)$, where $D_1^* L(x, y)$ is the contingent derivative of $x \to L(x, y)$. Actually, $D_1^* L(\hat{x}, .)(z)$ is a set valued directional derivative (by the sup-norm compactness, it is actually a directional multiderivative). For example, if $L(x, y) = a(x)c(y)$, where $a(x)$ is Lipschitz continuous and $c(y)$ is continuous, this equicontinuity holds. If a continuous function $F(x, y)$: $\mathbb{R}^2 \to \mathbb{R}$ has a partial derivative $F_1(x, y)$ which is continuous in (x, y), then, provided the above equicontinuity holds, $x \to F(L(x, .), .)$ has the directional multiderivative $F_1(L(\hat{x}, .), .)D_1^* L(\hat{x}, .)(z)$. (Note that $\lambda^{-1}[F(L(\hat{x}+\lambda z, y), y) - F(L(\hat{x}, y), y)]$
$= \lambda^{-1}[L(\hat{x} + \lambda z, y) - L(\hat{x}, y)] \int_0^1 F_1(L(\hat{x}, y) + r[L(\hat{x} + \lambda z, y) - L(\hat{x}, y)], y) dr$. The integrand is bounded, and for each r, is continuous in y, uniformly in λ.)

Next, let continuous functions $\hat{x}(.)$ and $z(.)$ in $C(C', \mathbb{R}^m)$ be given. If the family of functions of $y, \{b(\lambda, y)\}_{\lambda \in (0,1]}$, $b(\lambda, y) := [L(\hat{x}(y) + \lambda z(y), y) - L(\hat{x}(y), y)]/\lambda$, is equicontinuous in y, then $E(x(.)) := L(x(.), .) : C(C', \mathbb{R}^m) \to$

$C(C',\mathbb{R})$, has a d.d. container in $C(J,X)$, at $\hat{x}(.)$, given by $D_1^*L(\hat{x}(.),.)(z(.))$. Actually, $D_1^*L(\hat{x}(.),.)(z(.))$ is a directional multiderivative. The example $x(.) \to |x(.)| : C([0,1],\mathbb{R}) \to C([0,1],\mathbb{R})$ shows that such an equicontinuity can be difficult to obtain: If $\hat{x}(.)$ is any continuous function with both positive and negative values, no directional derivative exists at $\hat{x}(.)$ in $C([0,1],X)$, and more generally, equicontinuity does not hold.

K. Modification of (15)

If $\tilde{\delta}$ in (15) is required to belong to (ε,∞) rather than to $(0,\infty)$, for some given $\varepsilon > 0$ then in case of (18), in (15) we can replace $Q(T,\tilde{f},f)$ by $\{q(T)(T) : q(.) \in Q(\tilde{f},f^*,f^*,x^f(.))\}$. A proof is obtained from the proof of Remark 2 in Section 11, by choosing $\hat{\mu}$ also such that $4M\hat{\mu}\exp(2TM^{f^*}) \le \varepsilon\check{\mu}\mu'/4 \le \tilde{\delta}\check{\mu}\mu'/4$, and noting that $dq(t)/dt \in Df^*(s,x^f(.))(q(.))+$ $clB(0,1_{C_f}(\tilde{\delta}KM^{f^*}+2M) + \tilde{f}(t,x^f(.)) - f^*(t,x^f(.))$, Then, (15) follows for $\mu = \check{\mu}/4$.

L. Nonconvexity of C.

The following variant of Lemma 2 holds even in the case where C is nonconvex:

Lemma 19. (Exact penalization. No d.d. containers. No switching.) Let $(x^*(.),f^*)$ be optimal in problem (4)-(7), and let ζ be as defined in (14). Assume that $G(x)$ is continuous - and $\phi(x)$ Lipschitz continuous - in $B(x^*(T),\varsigma)$, the Lipschitz rank of ϕ being M_ϕ. Assume, moreover, that there exists a quadruple $(K,\mu,\hat{\mu},\mu') \in (0,\infty)\times(0,1)\times(0,\zeta]\times(0,\infty)$ with the property that for all $(f,c) \in clB(f^*,\hat{\mu}) \times (C \cap clB(G(x^*(T)),\hat{\mu}))$, all $v \in Y$ with $|v| = \mu'$, all $r \in (0,1]$, there exists a triple $(\hat{f},c',\hat{\delta})$, $\hat{f} \in clB(f^*,\zeta) \subset F$, $c' \in C$, and $\hat{\delta} \in (0,r]$, such that

$$\begin{aligned}&|G(x^{\hat{f}}(T)) - G(x^f(T)) - \hat{\delta}v - (c'-c)| \le (1-\mu)\hat{\delta}\mu',\\ &\sigma(\hat{f},f) \le \hat{\delta}K\mu', \text{ and } |c'-c| \le \hat{\delta}K\mu'.\end{aligned} \tag{95}$$

Let $c^* = G(x^*(T))$. Then, (f^*,c^*) maximizes

$$\iota^{**}(f) := \phi(x^f(T)) - 4\mu^{-1}KMM_\phi \mathrm{dist}(G(x^f(T)),C)\exp(\beta^f) \tag{96}$$

for f in some ball $clB(f^*, \tilde{\mu})$ in F, where $\beta^f := M^f\sigma(f, f^*)+$
$M^{f^*}(T - \sigma(f, f^*))$ and $\tilde{\mu} \in (0, \hat{\mu}/2]$.

Proof: The conclusion in Corollary 6, Section 10, can be stated as follows: a^* minimizes $a \to \eta(a) + 4KW_a\mu^{-1}\text{dist}(H(a), C \cap clB(H(a^*), \hat{\mu}/2))$ in $clB(a^*, \hat{\mu}/2)$. From this observation, the conclusion in the lemma follows, once the following identifications is made. Let $A = clB(f^*, \zeta)$, $a = f, H(a) = G(x^f(T)), \eta(f) = -\phi(x^f(T)), a^* = f^*$, and $W_a = M_\phi 2M\exp(\beta^f)$, see (23)(i). Observe that $\sigma(f, f^*) \leq \tilde{\mu} \leq \zeta$ implies $|x^f(T) - x^*(T)| \leq \tilde{\mu}2M\exp(TM^{f^*})$, so for $\tilde{\mu}$ small enough, $|G(x^f(T)) - c^*| \leq \hat{\mu}/4$, when f satisfies $\sigma(f, f^*) \leq \tilde{\mu}$. The next to last inequality yields $\hat{\mu}/4 \geq \text{dist}(G(x^f(T)), C)$ $= \text{dist}(G(x^f(T)), C \cap clB(c^*, \hat{\mu}/2))$.

M. Gronwall's inequality

Let non-negative continuous functions $a(t), b(t), c(t)$ satisfy $a(t) \leq b(t) + \int_0^t a(s)c(s)ds$, for all t in $[0, T]$. Then, for all t,

$a(t) \leq b(t) + \int_0^t b(s)c(s)\exp(\int_s^t c(r)dr)ds.$

When $b(s)$ is nondecreasing, the right hand side is $\leq$
$b(t) + \int_0^t b(t)c(s)\exp(\int_s^t c(r)dr)ds = b(t)\exp(\int_0^t c(r)dr)$.

References

Aubin, J-P., and I. Ekeland, (1984), *Applied Nonlinear analysis,* J.Wiley, New York.

Aubin, J-P., and H. Frankowska, (1990), *Set-valued analysis*, Birkhauser, 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.

Burnap, C. and M.A.Kazemi, (1999), " Optimal control of a system governed by nonlinear Volterra integral equations with delay", *IMA J. Math. Control Inf.* 16, No.1, 73-89.

Cichon, M. (1996), "Differential inclusions and abstract control problems", *Bull.Aust.Math.Soc.*, 1, 109-122.

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

Clarke, F.H. and P.R. Wolenski, (1996), "Necessary conditions for functional differential Inclusions", *Applied Math. and Optim.* 34, 51-78.

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., (1996), "Optimal control probems with state constraints for semilinear distributed-parameter systems". *J.Optim. Theory Appl.* 88, 25-59.

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.

Neustadt, L.W., (1976), *Optimization. A theory of Necessary Conditions,* Princeton University Press, Princeton, NJ.

Pogorzelski, W. (1966), *Integral equations and their applications,* Vol I, Pergamon Press, London.

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. (1995), "Directional derivatives and nonsmooth optimal control problems", *Optimization*, 35, 61-75.

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

Tiba, Dan, (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.

Xiang, X. and N.U. Ahmed, (1997), "Infinite dimensional functional differential inclusions and necessary conditions", *Nonlinear Anal Theory Methods Appl.* 30, 429-435.

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.