Set-Valued Var. Anal (2012) 20:467–475
DOI 10.1007/s11228-011-0196-6

Relaxed Derivatives and Extremality Conditions in Optimal Control

Jack Warga

Received: 15 September 2011 / Accepted: 16 September 2011 /
Published online: 6 October 2011

Abstract In previous work dating back to the early 1970's F.H. Clarke and the author had independently derived necessary conditions for minimum including a maximum principle for optimal control problems defined by ordinary differential equations in which the right hand side $f(t, \cdot, r)$ and functions defining side conditions are Lipschitz continuous in their dependence on the state variable. Our results, though not the methods, were similar in the formulation of the maximum principle in which the nonexisting derivative $f_v(t, v, \sigma)$ was replaced by an unknown element of Clarke's generalized Jacobian but differed in handling some side conditions. In the present paper we exhibit a maximum principle in which the dual variables and the related functions are limits of appropriate subsequences of computable sequences.

Keywords Optimal control · Maximum principle · Relaxation · Nonsmooth functions · Generalized derivatives

Mathematics Subject Classifications (2010) 49J15 · 49J52

1 Introduction

Our purpose is to derive extremality conditions for the optimal control problem defined by a differential equation in R^n

$$y(\tau; \sigma) = a_0 + \int_{\tau_0}^{\tau} \tilde{f}(s, y(s; \sigma), \sigma(s))ds \; \forall \tau \in \tilde{T} = [\tau_0, \tau_1] \tag{1}$$

This is the last paper by Professor Warga. It was mathematically completed by him right before his death in June 2011 and prepared for publication by Q. J. Zhu. Please send any correspondence to zhu@wmich.edu.

J. Warga (✉)
Department of Mathematics, Northeastern University, Boston, MA 02115, USA
e-mail: qiji.zhu@wmich.edu

and consisting in finding a relaxed control $\bar{\sigma}$ and the corresponding solution $\bar{y}(\tau) = y(\tau; \bar{\sigma})$ that minimize $h_0(y(\tau_1; \sigma))$ subject to the restriction

$$h_1(y(\tau_1; \sigma)) = 0, \tag{2}$$

where σ is a relaxed control and the functions $\tilde{f}(t, v, \sigma)$, $h_0(v)$ and $h_1(v)$ are lipschitz in their dependence of v. Among the basic assumptions are the existence of an integrable function $\chi(\tau)$ such that both the norm $|\tilde{f}|$ of $\tilde{f}$ and its time dependent Lipschitz constant $L_{\tilde{f}}(\tau)$ are bounded by $\chi(\tau)$. At an early stage we replace the independent variable τ by $t(\tau) = \int_{\tau_0}^{\tau} \chi(s)ds$ which transform $\tilde{f}$ to a function f whose norm and Lipschitz constant are bounded by 1.

While well known proofs of existence of an optimal control $\bar{\sigma}$ remain valid, we search to establish extremality conditions (essentially a maximum principle) for our problem. Such conditions are well established for the case where f, h_0 and h_1 are C^1 in their dependence on the state variable v. However, when this dependence is only Lipschitzian, the only maximum principle of which we are aware, e.g. [1, 5, 7], are derived by replacing the nonexistent derivatives with not otherwise specified functions whose pointwise values are elements of some convex sets of which the corresponding Clarke's generalized Jacobian turned out to be the smallest [6]. Our aim in what follows is to derive computable sequences whose appropriate subsequences converges to the dual variables and the corresponding coefficients of the maximum principle. We don't provide direct substitutes for the nonexisting derivatives $f_v, h_{0,v}$ and $h_{1,v}$ but deal with the integrals involving their temporary substitutes. These temporary substitutes are derivatives of mollified approximations to the Lipschitzian functions and the corresponding sequences converging to the appropriate dual function appearing in the maximum principle.

This explains why it seemed impossible to present a theorem resembling the one available for the case of f, h_0 and h_1 being C^1 in their dependence on v. The main results we derive here, in Theorem 3.8, involves mollified functions and thus clearly involve concepts introduced by Sobolev [3] and Friedrichs [2] for the study of partial differential equations but applied here to our specific problems. The form of Theorem 3.8 appears to lie half way between the results for problems with C^1 data and the results previously obtained for Lipschitzian data. While the latter results indicated that the desired data lie in specified convex sets, the results here prove that known sequence have subsequences converging to the desired data. We could get more definitive results if we could replace nonexisting derivative by derivatives but we won't be able to do that.

Section 2 is devoted to the mollification process and associated relaxed derivatives, Section 3 to a maximum principle for a "mollified" problem and to the statement of our final results, and Section 4 to proofs.

2 Relaxed Derivatives

Let R be the real line, $T = [\tau_0, \tau_1] \subset R$, N the set of natural numbers, $R^{m'}$ endowed with the Euclidean norm and the Lebesgue measure $\mu_{m'}$ for all $m' \in N$, $B^o_{m'}$ the open and $B_{m'}$ the closed unit ball in $R^{m'}$, $\tilde{V}$ an open convex subset of R^n and V a compact convex subset of $\tilde{V}$ such that $V + B_n \subset \tilde{V}$. We assume that the function $(\tau, v) \mapsto \phi(\tau, v) : \tilde{T} \times \tilde{V} \mapsto R^m$ in $L^1(T, C(\tilde{V})^m)$ and integrable function $\tau \mapsto \chi(\tau) : \tilde{T} \mapsto R^+$ and $\tau \mapsto L_\phi(\tau) : \tilde{T} \mapsto R^+$ satisfy the inequalities

$$|\phi(\tau, v)| \leq \chi(\tau),\ L_\phi(\tau) \leq \chi(\tau)\ \forall \tau \in \tilde{T}, v \in \tilde{V}$$

and

$$|\phi(\tau, v_1) - \phi(\tau, v_2)| \leq L_\phi(\tau)|v_1 - v_2|\ \forall \tau \in \tilde{T}, v_1, v_2 \in \tilde{V}.$$

Obviously, we may assume that $\chi(\tau) \geq 1\ \forall\ \tau \in \tilde{T}$.

In anticipation of the arguments to follow, we shall simplify our problem, without loss of generality, by changing the variable τ, which will be treated as an integration variable in the expression $\int_{\tilde{T}} \phi(\tau, v)d\tau$, by replacing it with the variable $t = t(\tau)$ defined by $t(\tau) := \int_{\tau_0}^{\tau} \chi(s)ds\ \forall \tau \in [\tau_0, \tau_1]$, defining its inverse $\tau(t)$ and setting $\psi(t, v) = [\chi(\tau(t))]^{-1}\phi(\tau(t), v)$. Let $\kappa := t(\tau_1)$ and $T = [0, \kappa]$. Then $\int_{\tilde{T}} \phi(\tau, v)d\tau = \int_T \psi(t, v)dt$. This results in great simplification because the functions like $v \mapsto \psi(t, v)$ that we shall be dealing with from now on have a norm and a Lipschitz constant L_ψ independent of t and, in fact, less than or equal to 1.

We shall use mollifiers $p^j : j^{-1}B_n \mapsto R\ \forall j \in N$ that are C^1 on $j^{-1}B_n$ and are approximations to the Dirac measure at 0, with $p^j(w)$ positive on $j^{-1}B_n^o$ and with $\int_{j^{-1}B_n} p^j(w)\mu_n(dw) = 1$. To be specific, we might use the well known C^∞ mollifier (in which ϵ is replaced by $1/j$) defined by the functions

$$\pi(w) = \exp\left(-1/\left(1 - |w|^2\right)\right)\ \forall w \in B_n^o,\ \pi(w) = 0 \forall w \in R^n \backslash B_n^o$$

$$\bar{\pi}(w) = \pi(w) / \int_{B_n^o} \pi(w')\mu_n(dw'),\ p^j(w) = j^n \bar{\pi}(jw).$$

We next consider, for all $(t, v) \in T \times V$ and $j \in N$ a sequence of C^1 functions

$$\psi^j(t, v) := \int_{j^{-1}B_n} p^j\left(w'\right)\psi\left(t, v - w'\right)\mu_n(dw') \tag{3}$$

Replacing the integration variable w' with $w = v - w'$ we obtain the relation

$$\psi^j(t, v) := \int_{v + j^{-1}B_n} p^j(v - w)\psi(t, w)\mu_n(dw). \tag{4}$$

Since by Rademacher's theorem, the derivative $\psi_v(t, w)$ exists a.e., we also deduce from relation (3) that

$$\psi_v^j(t, v) := \int_{j^{-1}B_n} p^j(w')\psi_v(t, v - w')\mu_n(dw'). \tag{5}$$

We observe that the integrals above are defined because the function ψ is defined on all of $V + B_n$. Furthermore, relation (4) shows that, because the function p^j are C^1 for each $j \in N$ and because v is in the interior of the domain of integration, the function $v \mapsto \psi^j(t, v)$ is C^1 and its (partial) derivative $\psi_v^j(t, v)$ is for each $j \in N$ and $t \in T$ uniformly continuous on the compact space V. This derivative, with t held constant, is represented by a matrix M with m rows and n columns that we shall treat as an element of the set $R_{m,n}$ of all such matrices endowed with the norm of M treated as a linear operator i.e. $|M| := \sup\{|Mx| \mid |x| \leq 1\}$.

Now let $t \mapsto \eta(t) : T \mapsto V$ be a continuous function and consider the function $t \mapsto \psi_v^j(t, \eta(t)) : T \mapsto R_{m,n}$. Since ψ_v^j is measurable in t and continuous in v, it follows that, for every $j \in N$, $t \mapsto \psi_v^j(t, \eta(t)) : T \mapsto R_{m,n}$ is measurable and, being dominated in the $R_{m,n}$ norm by $L_\psi \leq 1$, is itself integrable. Thus, $\Psi^j := \int_T \psi_v^j(t, \eta(t))\mu_T(dt)$ exists and is contained in the compact set $\kappa R_{m,n}$, where $\kappa = \int_{\tilde{T}} \chi(\tau)d\tau$. It follows then that there exists a sequence $J \subset N$ such that $\Psi := \lim_{j \in J} \Psi^j$ exists and belongs to $\kappa R_{m,n}$.

We shall later require

Lemma 1 *For all* $(t, v) \in T \times V$ *we have*

$$L_{\psi^j} = L_\psi, \ \left|\psi_v^j(t, v)\right| \leq L_\psi, \ \left|\psi^j(t, v) - \psi(t, v)\right| \leq L_\psi / j. \tag{6}$$

Proof By (3),

$$\psi^j(t, v) - \psi^j(t, v_1) = \int_{j^{-1} B_n} p^j(w')[\psi(t, v - w') - \psi(t, v_1 - w')]\mu_n(dw');$$

hence

$$\begin{aligned}
&\left|\psi^j(t, v) - \psi^j(t, v_1)\right| \\
&= \int_{j^{-1} B_n} p^j\left(w'\right)\left|\psi\left(t, v - w'\right) - \psi(t, v_1 - w')\right| \mu_n\left(dw'\right) \\
&\leq L_\psi |v - v_1|.
\end{aligned}$$

Since the norm of the derivative of a Lipschitzian function, whenever it exists, is bounded by its Lipschitz constant and the functions p^j define a probability measure, it follows from (5) that $|\psi_v^j(t, w)| \leq L_\psi \ \forall t \in T, \ w \in V$. Finally, again by (3),

$$\begin{aligned}
&|\psi^j(t, v) - \psi(t, v)| \\
&= \int_{j^{-1} B_n} p^j\left(w'\right)|\psi(t, v - w') - \psi(t, v)| \mu_n\left(dw'\right) \\
&\leq L_\psi \int_{j^{-1} B_n} p^j(w')\left|w'\right| \mu_n(dw') \leq j^{-1} L_\psi.
\end{aligned}$$

□

3 Extremality Conditions

Let $\mathcal{U}$ be the set of *original control functions* which are measurable functions from $\tilde{T} = [\tau_0, \tau_1] \subset R$ to a compact metric space U. We shall denote by $\mathcal{S}$ the set of relaxed control function $\sigma : \tilde{T} \mapsto rpm(U)$, as described in [4, Chapter IV], where $rmp(U)$ is the set of Radon probability measures on U. Each $\sigma \in \mathcal{S}$ is measurable with $\mathcal{S}$ identified as a subset of $L^1(T, C(U))^*$ and endowed with its weak star topology which is metrizable by its weak norm $|\ |_w$ [4, IV.1.5, p266]. Furthermore, $\mathcal{S}$ is a compact convex subset $\tilde{T}$ of $L^1(T, C(U))^*$. $\mathcal{U}$ is treated as a subset of $\mathcal{S}$ by identifying $u(\tau)$ for each $u \in \mathcal{U}$ with $\delta_{u(\tau)}$, where δ is the Dirac measure. If $\varphi : U \mapsto R^m$ for some $m \in N$ is continuous and ζ is a bounded measure on U we write $\varphi(\zeta)$ for $\int_U \varphi(r)\zeta(dr)$. As before, we define $V \subset R^n$ as a compact subset of an open V.

We shall consider the relaxed optimal control problem defined for $\sigma \in \mathcal{S}$ by the differential equation in R^n

$$y(\tau;\sigma) = a_0 + \int_{\tau_0}^{\tau} \tilde{f}(s, y(s;\sigma), \sigma(s))\, ds \; \forall \tau \in \tilde{T} = [\tau_0, \tau_1] \tag{7}$$

and consisting in finding a relaxed control $\bar{\sigma}$ and the corresponding solution $\bar{y}(\tau) = y(\tau;\bar{\sigma})$ that minimize $h_0(y(\tau_1;\sigma))$ subject to the restriction

$$h_1(y(\tau_1;\sigma)) = 0. \tag{8}$$

We make the following

Assumption 2

(i) The function $(\tau, v, r) \mapsto \tilde{f}(\tau, v, r) : \tilde{T} \times V \times R \mapsto R^n$ is measurable in τ, Lipschitz continuous in v with a Lipschitz constant $L_{\tilde{f}}(\tau)$ independent of τ, and continuous in r.
(ii) There exists a nonnegative function $\chi(\tau)$ integrable on $\tilde{T}$ and such that $|\tilde{f}(\tau, v, r)|$ and $L_{\tilde{f}}(\tau)$ are both bounded by $\chi(\tau)$ for all $\tau \in \tilde{T}$.
(iii) $h_0 : V \mapsto R$ and $h_1 : V \mapsto R^m$ are Lipschitz and have Lipschitz constants L_{h_0} and L_{h_1}.
(iv) There exists a subset $\mathcal{S}_1$ of $\mathcal{S}$ such that the solution $y(\tau) = y(\tau;\sigma)$ of equation (7) for $\sigma \in \mathcal{S}_1$ satisfies the conclusion of Lemma 1.

It easily follows that equation (7) has, for all $\sigma \in \mathcal{S}$, unique uniformly bounded and absolutely continuous solutions $y(\tau) = y(\tau;\sigma)$ for all $\sigma \in \mathcal{S}$ and we may therefore assume that V was defined so that every solution $y(\tau;\sigma)$ has its range in V.

Assumption (2) above satisfy the requirement of [4, Theorem VI.1.1, p348] and are therefore sufficient to ensure the existence of a minimizing control function $\bar{\sigma}$. We observe that, as a consequence of Assumption 2, the function $\tilde{f}(\tau, v, \sigma(\tau))$ satisfies, for each $\sigma \in \mathcal{S}$, the assumptions imposed on the function $\phi(\tau, v)$ in Section 2 as do the functions h_0 and h_1 that are independent of τ. We may therefore assume, without loss of generality, that after the change of the integration variable from τ to t, the transformed function f corresponding to ψ of Section 2 has a norm and a Lipschitz constant L_f with respect to v bounded by 1; that $\tilde{T} = [\tau_0, \tau_1]$ is replaced by $T = [0, \kappa]$, where $\kappa = \int_{\tau_0}^{\tau_1} \chi(\tau) d\tau$; that the relaxed control function σ is redefined as a function on T; that relations (7) and (8) are replaced by

$$y(t;\sigma) = a_0 + \int_0^t f(s, y(s;\sigma), \sigma(s)) ds \; \forall t \in T; \tag{9}$$

and

$$h_1(y(\kappa;\sigma)) = 0, \tag{10}$$

and that the control problem defined by the function h_0 and relations (9) and (10) has an optimal control function $\bar{\sigma}(t)$ and corresponding function $\bar{y}(t) := y(t;\bar{\sigma})$.

Our first step in the search for the extremality conditions relating to the problem defined by relations (9) and (10) will be to define, for each $\sigma \in \mathcal{S}$ and $j \in N$ the mollified functions $f^j(t, v, \sigma(t))$, $h_0^j(v)$ and $h_1^j(v)$ by replacing $\psi(t, v)$ in relation (3) with $f(t, v, \sigma(t)), h_0(v)$ and $h_1(v)$, respectively, and replacing equation (9) with equation

$$y^j(t;\sigma) = a_0 + \int_0^t f^j\left(s, y^j(s;\sigma), \sigma(s)\right) ds \; \forall t \in T. \tag{11}$$

which has a unique solution $y^j(t;\sigma)$.

We shall require the following results:

Lemma 3 *Let*

$$\kappa := \int_{\tau_0}^{\tau_1} \chi(\tau) d\tau, c_y := 1 + \kappa e^{\kappa}, c_{h_i} := L_{h_i}(c_y + 1) \; \forall i = 0, 1, c_f := c_y + 1.$$

Then, for each $\sigma \in \mathcal{S}$, we have

(i) $w^j(t) := |y^j(t;\sigma) - y(t;\sigma)| \le c_y/j$,
(ii) $|f^j(t, y^j(t;\sigma), \sigma(t)) - f(t, y(t;\sigma), \sigma(t))| \le c_f/j$,
(iii) $|h_i^j(y^j(\kappa;\sigma)) - h_i(y(\kappa;\sigma))| \le c_{h_i}/j \; \forall i = 0, 1$.

We next define a function $H_1^j : V \times B^j \mapsto R^{m_1}$ by $H_1^j(v, b) = h_1^j(v) - b^j$, where $b^j \in B^j$ is a control parameter and $B^j = j^{-1} c_{h_1} B_{m_1}$.

We can now formulate our "mollified" problem: let $y^j(t;\sigma)$ denote the unique solution of equation (11) for the choice of any $\sigma \in \mathcal{S}$. We wish to determine a relaxed control $\bar{\sigma}^j$ and a control parameter $\bar{b}^j \in B^j$ that minimize $h_0^j(y^j(\kappa;\sigma))$ subject to the restriction

$$H_1^j\left(y^j(\kappa;\sigma), b\right) = h_1^j\left(y^j(\kappa;\sigma)\right) - b^j = 0. \tag{12}$$

The proof that each of these new problems, indexed by j, has a minimizing solution will follow from the same arguments as in the existence proof for the original problem defined by relations (9) and (10) once we verify that there exists a choice of $(\sigma^j, b^j) \in \mathcal{S} \times B^j$ satisfying relation (12). We claim that the control $(\sigma^j, b^j) = (\bar{\sigma}, h_1^j(y^j(\kappa;\bar{\sigma})))$, where $\bar{\sigma}$ is the optimal control function for the problem defined by relations (9) and (10), satisfying these relations. Indeed, by (ii) in Lemma 3, we have $|h_1^j(y^j(\kappa;\bar{\sigma})) - h_1(y(\kappa;\bar{\sigma}))| = |h_1^j(y^j(\kappa;\bar{\sigma}))| \le c_{h_1}/j$ and therefore there exists an element $b^j \in B^j$ equal to $h_1^j(y^j(\kappa;\bar{\sigma}))$. Thus the restriction in (12) is satisfied by the above choice of b^j combined with $\sigma = \bar{\sigma}$.

The corresponding optimal control problems satisfy extremality conditions [4, Theorem VI.2.3, p360–361] and, in particular, of [4, Step 2 of Theorem VI.2.3.] which asserts that there exist functions

$$Z^j : T \mapsto R_{n,n}, k^j : T \mapsto R^n$$

and vectors $l^j = (l_0^j, l_1^j) \in R^{m_1+1}$ with $l_0^j \ge 0$ such that, setting $\bar{y}^j(t) := y^j(t;\bar{\sigma}^j)$,

$$Z^j(t) = I + \int_t^{\kappa} Z^j(s) f_v^j\left(s, \bar{y}^j(s), \bar{\sigma}^j(s)\right) ds \tag{13}$$

and

$$k^j(t)^{tr} = \left[l_0^j h_{0;v}^j\left(\bar{y}^j(\kappa)\right) + \left(l_1^j\right)^{tr} h_{1;v}^j\left(\bar{y}^j(\kappa)\right)\right] Z^j(t), \tag{14}$$

we have

$$l_0^j + \left| l_1^j \right| > 0 \tag{15}$$

and

$$\int_0^\kappa k^j(t)^{tr} f^j\left(t, \bar{y}^j(t), \sigma^j(t)\right) dt \geq \int_0^\kappa k^j(t)^{tr} f^j\left(t, \bar{y}^j(t), \bar{\sigma}^j(t)\right) dt \; \forall \sigma \in \mathcal{S}. \tag{16}$$

Theorem 4 *Let $\bar{\sigma} \in \mathcal{S}$ yield a an optimal solution to the problem of minimizing $h_0(y(t;\sigma))$ subject to relations (9) and (10). Then there exist a subsequence J of N and continuous functions $Z : T \mapsto (1+\kappa e^\kappa) R_{n,n}$ and $k : T \mapsto R^n$ such that we have the* **maximum principle**

$$k^j(t)^{tr} f^j\left(t, y(t;\bar{\sigma}), \bar{\sigma}(t)\right) = \min_{r\in U} k^j(t)^{tr} f^j\left(t, y(t;\bar{\sigma}), r\right) \; a.e., \tag{17}$$

where

$$Z(t) = \lim_{j\in J} Z^j \text{ uniformaly on } T,$$

$$k(t)^{tr} = \left[l_0 \mathcal{H}_0 + l_1^{tr} \mathcal{H}_1\right] Z(t), \;\; (l_0, l_1) = \lim_{j\in J} \left(l_0^j, l_1^j\right),$$

$$\mathcal{H}_0 = \lim_{j\in J} h_{0;v}^j\left(\bar{y}^j(\kappa)\right) \in L_{h_0} B_n, \;\; \mathcal{H}_1 = \lim_{j\in J} h_{1;v}^j\left(\bar{y}^j(\kappa)\right) \in L_{h_1} B_{m,n}.$$

4 Proofs

Proof of Lemma 3 To prove relation (i) we subtract equation (9) from equation (11) for each fixed $\sigma \in \mathcal{S}$ yielding

$$\begin{aligned} w^j(t) &:= \left| y^j(t;\sigma) - y(t;\sigma) \right| \\ &\leq \int_{t_0}^t \Big[\left| f^j(s, y^j(s;\sigma), \sigma(s)) - f(s, y^j(s;\sigma), \sigma(s)) \right| \\ &\quad + \left| f(s, y^j(s;\sigma), \sigma(s)) - f(s, y(s;\sigma), \sigma(s)) \right| \Big] ds. \end{aligned} \tag{18}$$

Since $L_f \leq 1$, it follows, by Lemma 1, that the first part of the integrand is bounded by $1/j$ and the second part by $w^j(t)$. It then follows by Gronwall's inequality that relation (i) is valid.

The left hand side of statement (ii) is bounded by the integrand in (18) which itself is bounded by $1/j + w^j(t)$, hence in view of (i), by c_f/j. The proof of (iii) is like the proof of (ii), with h_i replacing f. □

Proof of Theorem 4 We shall first prove that the solutions Z^j of the differential equations (13) for $j \in N$ form an equicontinuous and bounded set in $R_{n,n}$ and there exists a sequence $J_1 \subset N$ such that $Z^j(t)$ converge uniformly over J_1 to an absolutely continuous function $Z(t)$. This is because, by Lemma 1, we have

$|f_v^j(s,\bar{y}^j(s),\bar{\sigma}^j(s))| \le L_f \le 1$, and therefore $|Z^j(t)| \le \int_0^\kappa |Z^j(s)|ds + 1$. This implies, by Gronwall's inequality, that

$$\left|Z^j(s)\right| \le 1 + \kappa e^\kappa := \kappa_Z.$$

Thus all functions $t \mapsto Z^j(t) : T \mapsto R$ are uniformly bounded. Furthermore, for each $j \in N$, we have

$$\left|(Z^j)'(t)\right| \le \left|Z^j(t)\right| \le \kappa_Z \ \forall t \in T$$

and this shows that the closure of the set $\{Z^j(\cdot) \mid j \in N\}$ is compact in $C([0,\kappa], \kappa_Z B_{n,n})$ and there exists a sequence $J_1 \subset N$ such that $\lim_{j\in J_1} Z^j(t) = Z(t) \in \kappa_Z B_{n,n}$ uniformly on T.

We next consider equation (14) defining $k^j(t)$. We observe that dividing l_0^j and l_1^j by $l_0^j + |l_1^j|$, which results in dividing $k^j(t)\ \forall t \in T$ by the same factor, does not affect the validity of relations (14)–(16) and justifies assuming that

$$l_0^j \ge 0,\ l_0^j + |l_1^j| = 1 \forall j \in N.$$

Thus $0 \le l_0^j \le 1$, $|l_1^j| \le 1$ and, by Lemma 1, $|h_{0;v}^j(\bar{y}^j(\kappa))| \le L_{h_0}$ and $|h_{1;v}^j(\bar{y}^j(\kappa))| \le L_{h_1}$. Since $\lim_{j\in J_1} Z^j(t) = Z(t)$ uniformly on T, it follows that there exist

$$J_2 \subset J_1, l_0 \ge 0, l_1 \in B_m, \mathcal{H}_0 \in L_{h_0} B_n, \mathcal{H}_1 \in L_{h_1} B_{m,n}$$

such that the limits, as $j \mapsto \infty$ over J_2, of

$$l_0^j, l_1^j, h_{0;v}^j(\bar{y}^j(\kappa)), h_{1;v}^j\left(\bar{y}^j(\kappa)\right)$$

are, respectively,

$$l_0, l_1, \mathcal{H}_0, \mathcal{H}_1,$$

with $l_0 \ge 0$, $l_0 + |l_1| = 1$ and

$$k(t)^{tr} := \lim_{j\in J_2} k^j(t)^{tr} = \left[l_0\mathcal{H}_0 + l_1^{tr}\mathcal{H}_1\right] Z(t)$$

uniformly on T.

We next observe that, by Lemma 3 (i), $|y^j(t;\bar{\sigma}^j) - y(t;\bar{\sigma}^j)| \le c_y/j \mapsto 0$ and the set Y of all the functions $t \mapsto y(t;\bar{\sigma}^j)$ has a compact closure in $C(T, R^n)$ because it is bounded and equicontinuous. Therefore the argument of [4, theorem VI.1.1, p348] proves that there exist $\bar{\sigma} \in \mathcal{S}$ and a subsequences J of J_2 such that

$$\lim_{j\in J} y^j\left(t;\bar{\sigma}^j\right) = \lim_J y\left(t;\bar{\sigma}^j\right) = y(t;\bar{\sigma}). \quad (19)$$

We now consider relation (16). We first observe that, since $\lim_{j\in J} k^j(t) = k(t)$, $\lim_{j\in J}\bar{\sigma}^j = \bar{\sigma}$ and $\lim_{j\in J} y^j(t;\bar{\sigma}^j) = y(t;\bar{\sigma})$, it follows from [4, Theorem IV.2.9, p278–279] that

$$\lim_{j\in J}\int_T k^j(t)^{tr} f\left(t, y^j(t;\bar{\sigma}^j), \bar{\sigma}^j\right)dt = \int_T k(t)^{tr} f(t, y(t;\bar{\sigma}), \bar{\sigma})\,dt.$$

Since $|f^j(t,v,r) - f(t,v,r)| \le 1/j \mapsto 0$, it follows from relation (16) that

$$\int_0^\kappa k(t)^{tr} f(t,\bar{y}(t),\sigma(t))dt \ge \int_0^\kappa k(t)^{tr} f(t,\bar{y}(t),\bar{\sigma}(t))dt \ \forall\sigma \in \mathcal{S}.$$

Applying an argument of [4, Step 2 of Theorem VI.2.3, p360–361] we deduce from the above maximum principle in integral form the pointwise maximum principle

$$k(t)^{tr} f\left(t, y(t;\bar{\sigma}), \bar{\sigma}(t)\right) = \min_{r\in U} k(t)^{tr} f(t, y(t;\bar{\sigma}), r)\ a.e.$$

By relation (12), we have $h_1^j(y^j(t_1;\bar{\sigma}^j)) - \bar{b}^j = 0$. Since $(\bar{\sigma}^j, \bar{b}^j)$ is an optimal control for the problem indexed by j, we have $h_0^j(y^j(\kappa;\bar{\sigma}^j)) \le h_0^j(y^j(\kappa;\sigma^j))$ for all σ such that $h_1^j(y^j(\kappa;\sigma)) = \bar{b}^j$. Furthermore, by the definition of B^j we have $\lim_{j\in N} \bar{b}^j = 0$ and , since $\lim_{j\in J} \bar{\sigma}^j = \bar{\sigma}$ this implies, by (19) that $\lim_{j\in J} y^j(t;\bar{\sigma}^j) = y(t;\bar{\sigma})$. Therefore, $\lim_{j\in J} h_1^j(y(t;\bar{\sigma})) = 0$ and thus

$$\lim_{j\in J} h_0^j\left(y^j(\kappa;\bar{\sigma}^j)\right) = h_0(y(\kappa;\bar{\sigma})) \le \lim_{j\in J} h_0^j\left(y^j(\kappa;\sigma)\right) = h_0(y(\kappa;\sigma))$$

for all $\sigma \in \mathcal{S}$ such that $h_1(y(\kappa;\sigma)) = 0$. We conclude that $\bar{\sigma}$ is an optimal control for the problem defined by the function h_0 and the equations (9) and (10) and that it satisfies the maximum principle (17). □

References

1. Clarke, F.H.: Optimization and nonsmooth analysis. Wiley Interscience, New York (1983)
2. Friedrichs, K.O.: The identity of weak and strong extensions of differential operators. Trans. Am. Math. Soc. **55**(1), 132–151 (1944)
3. Soboleff, S.L.: Sur un thèoréme d'analyse fonctionnelle. Matematicheskii Sbornik **4**(46)(3), 472–497, Zbl 0022.14803 (1938) (in Russian, with French summary)
4. Warga, J.: Optimal control of differential and functional equations. Academic, New York (1972)
5. Warga, J.: Necessary conditions without differentiability assumptions in optimal control. J. Differ. Equ. **18**, 41–62 (1975)
6. Warga, J.: Fat homeomorpgisms and unbounded derivate containers. J. Math. Anal. Appl. **81**, 545–560 (1981); Errata. **90**, 582–583 (1982)
7. Warga, J.: Controllability, estremality and abnormality in nonsmooth optimal control. J. Optim. Theory Appl. **41**, 239–260 (1983)