

NECESSARY CONDITIONS FOR OPTIMAL CONTROL PROBLEMS WITH STATE CONSTRAINTS

R. B. VINTER AND H. ZHENG

ABSTRACT. Necessary conditions of optimality are derived for optimal control problems with pathwise state constraints, in which the dynamic constraint is modelled as a differential inclusion. The novel feature of the conditions is the unrestrictive nature of the hypotheses under which these conditions are shown to be valid. An Euler Lagrange type condition is obtained for problems where the multifunction associated with the dynamic constraint has values possibly unbounded, nonconvex sets and satisfies a mild ‘one-sided’ Lipschitz continuity hypothesis. We recover as a special case the sharpest available necessary conditions for state constraint free problems proved in a recent paper by Ioffe. For problems where the multifunction is convex valued it is shown that the necessary conditions are still valid when the one-sided Lipschitz hypothesis is replaced by a milder, local hypothesis. A recent ‘dualization’ theorem permits us to infer a strengthened form of the Hamiltonian inclusion from the Euler Lagrange condition. The necessary conditions for state constrained problems with convex valued multifunctions are derived under hypotheses on the dynamics which are significantly weaker than those invoked by Loewen and Rockafellar to achieve related necessary conditions for state constrained problems, and improve on available results in certain respects even when specialized to the state constraint free case.

Proofs make use of recent ‘decoupling’ ideas of the authors, which reduce the optimization problem to one to which Pontryagin’s maximum principle is applicable, and a refined penalization technique to deal with the dynamic constraint.

1. INTRODUCTION

The derivation of necessary conditions for optimal control problems in which the dynamic constraint takes the form of a differential inclusion has been an area of active research since the early 1970’s. A framework which is frequently adopted for work in this field is the optimization problem

$$\begin{aligned} &\text{Minimize } g(x(0), x(1)) \\ &\text{over absolutely continuous arcs } x : [0, 1] \rightarrow R^n \text{ which satisfy} \\ &\dot{x}(t) \in F(t, x(t)), \quad \text{a.e.,} \\ &(x(0), x(1)) \in C. \end{aligned}$$

Received by the editors May 6, 1996.

1991 *Mathematics Subject Classification*. Primary 49K24.

Key words and phrases. Euler Lagrange condition, Hamiltonian inclusion, state constraint, nonconvex differential inclusion, nonsmooth analysis.

This research was carried out with financial support provided by EPSRC.

Here $g : R^n \times R^n \rightarrow R$ is a given function, $F : [0, 1] \times R^n \rightrightarrows R^n$ is a given multifunction, and $C \subset R^n \times R^n$ is a given set.

Let \bar{x} be a minimizing arc. Currently available sets of necessary conditions typically assert the existence of an “adjoint arc” p which satisfies a boundary condition involving g and C (the transversality condition), some sort of differential inclusion and a condition involving the Hamiltonian function

$$H(t, x, p) := \max_{e \in F(t, x)} e \cdot p.$$

What distinguishes one set of conditions from another is usually the nature of the differential inclusion which p is required to satisfy. In Clarke’s Hamiltonian inclusion [2], it is

$$(-\dot{p}(t), \dot{\bar{x}}(t)) \in \text{co}\partial H(t, \bar{x}(t), p(t)), \quad \text{a.e.}$$

Here ∂ denotes the limiting subdifferential with respect to x and p for fixed t (see below), and co the operation of taking the convex hull. In the Euler Lagrange inclusion (early versions of which were derived by Clarke [1], Mordukhovich [11] and Loewen and Rockafellar [7]) the above relationship is replaced by

$$(1) \quad (\dot{p}(t), p(t)) \in \text{co}N_{\text{Gr}F(t, \cdot)}(\bar{x}(t), \dot{\bar{x}}(t)), \quad \text{a.e.}$$

In much of the literature F is assumed to be convex valued. For a more thorough discussion than we can provide here of necessary conditions and their historical background, see [4], [14], [11].

In recent years the necessary conditions have been improved, and the hypotheses under which they can be shown to be valid have been relaxed. Conditions are now available, valid for possibly nonconvex F ’s, involving an extended Euler Lagrange condition, of the type

$$(2) \quad \dot{p}(t) \in \text{co}\{\eta : (\eta, p(t)) \in N_{\text{Gr}F(t, \cdot)}(\bar{x}(t), \dot{\bar{x}}(t))\}, \quad \text{a.e.}$$

and

$$(3) \quad p(t) \cdot \dot{\bar{x}}(t) = H(t, \bar{x}(t), p(t)), \quad \text{a.e.}$$

(3) is implied by (2) in the convex case, but in the absence of convexity hypotheses it is a distinct condition.) Notice that condition (2) improves on (1) because it requires p to satisfy a relationship involving convexification in only one variable. Versions of these results have been provided by Mordukhovich [11] (without (3)), Ioffe [4] and Vinter and Zheng [14]. The most general necessary conditions along these lines (which allow for non-convex, unbounded F ’s satisfying a “one-sided” Lipschitz continuity condition) were proved by Ioffe [4]; two steps are involved in the proof:

1) Derive an analogous set of necessary conditions for a certain calculus of variations problem with finite Lagrangian (this step was taken earlier by Ioffe and Rockafellar [5], whose proof techniques are based on an analysis of subdifferentials, and associated fuzzy calculus, of integral functions on infinite dimensional spaces).

2) Obtain necessary conditions for the original problem by applying the conditions of step 1 to a sequence of related problems via a new penalization procedure, and passing to the limit.

Vinter and Zheng [14] provided a simple derivation of the required results in step 1, based on an application of the Pontryagin Maximum Principle to a related “decoupled” problem and well-known properties of limiting subdifferentials in finite dimensional spaces.

Another recent development, Rockafellar's dualization theorem [12], was to show that, when $F(t, \cdot)$ is convex valued, given n -vectors p', p, x and v , then (for fixed t) the conditions

$$(4) \quad p' \in \text{co}\{\eta : (\eta, p) \in N_{\text{Gr}F(t, \cdot)}(x, v)\}$$

and

$$(5) \quad p' \in \text{co}\{-\eta : (\eta, v) \in \partial H(t, x, p)\}$$

are equivalent under mild hypotheses on the multifunction $F(t, \cdot)$. (Subsequently Ioffe [4] established, as a partial dualization, that (4) implies (5), under somewhat weaker hypotheses.) This means, if we can show (2) and we know that F is convex valued, we automatically have that the adjoint arc p also satisfies the condition

$$(6) \quad \dot{p}(t) \in \text{co}\{-\eta : (\eta, \dot{x}(t)) \in \partial H(t, \bar{x}(t), p(t))\}.$$

This is a strengthened form of the Hamiltonian inclusion involving convexification in only one variable, proved earlier by Loewen and Rockafellar. (See [7] and the references therein.)

The significance of the extended Euler Lagrange condition (2) now becomes clear. It is the endproduct of an extensive research effort into refinements of the Euler Lagrange equation which are valid for nonconvex F 's, and which for convex F 's are equivalent to a sharpened version of rival conditions in the form of the Hamiltonian inclusion.

We are now ready to describe the goals of the present paper. They are to provide advances parallel to those described above, when the optimal control problem considered incorporates a unilateral state constraint of the form

$$h(t, x(t)) \leq 0 \quad \text{for all } t \in [0, 1],$$

where $h : [0, 1] \times R^n \rightarrow R$ is a given upper semi-continuous function such that $h(t, \cdot)$ is Lipschitz continuous "near" the minimizing arc, uniformly with respect to t .

These goals are accomplished; the centrepiece is an extended Euler Lagrange condition for state constrained problems, involving nonconvex differential inclusions, which is derived under the mild hypotheses (unbounded F 's satisfying merely a one-sided Lipschitz continuity condition, etc.) invoked by Ioffe in the state-constraint-free case. With the help of the partial dualization theorem, we derive also a refined Hamiltonian condition for convex valued F 's under a number of alternative hypotheses, some of which reduce to those employed by Ioffe for no state constraints, and which are significantly weaker in some respects than those under which the refined Hamiltonian inclusion for state-constrained problems was derived earlier by Loewen and Rockafellar.

The proof techniques employ the same steps 1 and 2 described earlier, but adapted to allow for the state constraint. In step 1 we use the decoupling technique of Vinter and Zheng [14], which can be simply modified to accommodate the state constraint, while in step 2 we use a variant on Ioffe's penalization procedure (which again allows for state constraints) to obtain the final necessary conditions under very unrestrictive hypotheses on the data.

To conclude this introduction we clarify some notation and define some basic concepts.

$|\cdot|$ denotes the Euclidean norm throughout. The closed Euclidean unit ball is written B . $d_C(x)$ denotes the Euclidean distance of the point $x \in R^n$ from the set $C \subset R^n$.

$\text{epi}\{f\}$ is the epigraph set of the function $f : R^k \rightarrow R \cup \{+\infty\}$, i.e.

$$\text{epi}\{f\} = \{(x, \alpha) \in R^k \times R : \alpha \geq f(x)\}.$$

$\Psi_A : R^k \rightarrow R \cup \{+\infty\}$ is the indicator function of the set $A \subset R^k$, which takes the value 0 on A and $+\infty$ elsewhere.

Given a continuous arc $p : [0, 1] \rightarrow R^n$, $\|p\| = \max_{t \in [0, 1]} |p(t)|$. For a non-negative measure μ on the Borel subsets of $[0, 1]$, we define $\|\mu\| := \int_{[0, 1]} \mu(ds)$.

$W^{1,1}([0, 1]; R^n)$, or briefly $W^{1,1}$, denotes the space of R^n valued absolutely continuous functions on $[0, 1]$. The norm on $W^{1,1}$ is

$$\|x\|_{W^{1,1}} := |x(0)| + \int_0^1 |\dot{x}(t)| dt.$$

We shall refer to an arc \bar{x} which is a $W^{1,1}$ local minimizer of the optimization problem under currently consideration. This means that there exists some ϵ such that \bar{x} achieves the minimum value of the cost function over arcs $x \in W^{1,1}$ satisfying the constraints of the problem and also $\|x - \bar{x}\|_{W^{1,1}} \leq \epsilon$.

Use will be made of the following two constructs from nonsmooth analysis.

Definition 1. Take a closed set $A \subseteq R^k$ and points $x \in A$, $p \in R^k$. We say that p is a *limiting normal* to A at x if and only if there exist $p_i \rightarrow p$ and $x_i \rightarrow x$ in A such that, for each i , $p_i \cdot (x - x_i) \leq o(|x - x_i|)$ for all $x \in A$, where $o(\cdot) : R^+ \rightarrow R^+$ is some function such that $o(\alpha)/\alpha \rightarrow 0$ as $\alpha \downarrow 0$ (i.e. limiting normals are limits of vectors which support A at points near x , to first order). The limiting normal cone to A at x , written $N_A(x)$, is the set of all limiting normals to A at x .

Definition 2. Given a lower semicontinuous function $f : R^k \rightarrow R \cup \{+\infty\}$ and a point $x \in R^k$ such that $f(x) < +\infty$, the *limiting subdifferential* of f at x , written $\partial f(x)$, is

$$\partial f(x) := \{\xi : (\xi, -1) \in N_{\text{epi}\{f\}}(x, f(x))\}.$$

We shall frequently call upon standard results in nonsmooth calculus (to be found, for example, in [2], [6], [10]).

2. THE BOLZA PROBLEM WITH FINITE LAGRANGIAN

We begin by deriving necessary conditions for a problem in which a state constraint is present but no restrictions are placed on the velocity:

$$(7) \quad \begin{aligned} &\text{Minimize } J(x) := l(x(0), x(1)) + \int_0^1 L(t, x(t), \dot{x}(t)) dt \\ &\text{over arcs } x \text{ which satisfy } h(t, x(t)) \leq 0 \text{ for all } t. \end{aligned}$$

Here $l : R^n \times R^n \rightarrow R \cup \{+\infty\}$, $L : [0, 1] \times R^n \times R^n \rightarrow R$ and $h : [0, 1] \times R^n \rightarrow R$ are given functions. Note that, since l is extended real valued, this formulation incorporates an “endpoint constraint”, namely

$$(x(0), x(1)) \in \{(\alpha, \beta) : l(\alpha, \beta) < +\infty\}.$$

This is because an arc x which violates this constraint has infinite cost and is therefore not a candidate for being a minimizer. This is termed a “finite Lagrangian” problem because the extended Lagrangian $L_e(t, x, v)$, which coincides with $L(t, x, v)$ on $F(t, x)$ and take the value $+\infty$ off $F(t, x)$, is finite valued.

These necessary conditions provide an important step in the treatment of problems with a dynamic constraint. But necessary conditions for finite Lagrangian problems are of independent interest because of the unrestrictive nature of the hypotheses under which they are valid (apart of course from the fact that they do not allow for a dynamic constraint). These hypotheses, which relate to a $W^{1,1}$ local minimizer \bar{x} , are

- (H1) l is lower semicontinuous.
 (H2) $L(\cdot, x, \cdot)$ is $\mathcal{L} \times \mathcal{B}$ measurable for each x , where \mathcal{L} and \mathcal{B} denote the Lebesgue subsets of $[0, 1]$ and the Borel subsets of R^n respectively. $L(t, \cdot, \cdot)$ is lower semicontinuous for a.e. t .
 (H3) For every $N > 0$ there exist $\delta > 0$ and $k \in L^1$ such that
- $$|L(t, x', v) - L(t, x, v)| \leq k(t)|x' - x|, \quad L(t, \bar{x}(t), v) \geq -k(t)$$
- for all $x', x \in \bar{x}(t) + \delta B$ and $v \in \dot{\bar{x}}(t) + NB$, a.e. $t \in [0, 1]$.
 (H4) h is upper semicontinuous near $(t, \bar{x}(t))$ for all t , and there exists a constant k_h such that

$$|h(t, x') - h(t, x)| \leq k_h|x' - x|$$

for all t in $[0, 1]$ and all $x', x \in \bar{x}(t) + \delta B$.

Theorem 3. *Let \bar{x} be a $W^{1,1}$ local minimizer for (7), for which $J(\bar{x}) < \infty$. Assume that hypotheses (H1)–(H4) are satisfied. Then there exist an arc $p \in W^{1,1}$, a non-negative constant λ , a non-negative measure μ and a μ -integrable function $\gamma : [0, 1] \rightarrow R^n$ such that*

- (i) $\lambda + \|p\| + \|\mu\| = 1$,
 (ii)

$$\begin{aligned} \dot{p}(t) \in \text{co}\{\eta : (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds), -\lambda) \\ \in N_{\text{epi}\{L(t, \cdot, \cdot)\}}(\bar{x}(t), \dot{\bar{x}}(t), L(t, \bar{x}(t), \dot{\bar{x}}(t)))\} \quad \text{a.e.}, \end{aligned}$$

- (iii) $(p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)], -\lambda) \in N_{\text{epi}\{l\}}(\bar{x}(0), \bar{x}(1), l(\bar{x}(0), \bar{x}(1)))$,

(iv)

$$\begin{aligned} (p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \cdot \dot{\bar{x}}(t) - \lambda L(t, \bar{x}(t), \dot{\bar{x}}(t)) \\ \geq (p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \cdot v - \lambda L(t, \bar{x}(t), v) \end{aligned}$$

for all $v \in R^n$ a.e.,

- (v) $\gamma(t) \in \partial_x^> h(t, \bar{x}(t))$ μ -a.e. and $\text{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}$. Here

$$\partial_x^> h(t, x) := \text{co}\{\lim_i \xi_i : \exists t_i \rightarrow t, x_i \rightarrow x \text{ such that}$$

$$h(t_i, x_i) > 0 \text{ and } \xi_i \in \partial_x h(t_i, x_i) \text{ for all } i\}.$$

Comments. (1) If the state constraint “ $h(t, x(t)) \leq 0$ ” is dropped from Problem 1, then, under hypotheses (H1)–(H3), we may deduce from Theorem 3 the existence of $p \in W^{1,1}$ such that

(ii)' $\dot{p}(t) \in \text{co}\{\eta : (\eta, p(t)) \in \partial L(t, \bar{x}(t), \dot{\bar{x}}(t))\}$ a.e., in which ∂L denotes the limiting subdifferential of $L(t, \cdot, \cdot)$,

(iii)' $(p(0), -p(1)) \in \partial l(\bar{x}(0), \bar{x}(1))$,

(iv)' $(p(t)) \cdot \dot{\bar{x}}(t) - L(t, \bar{x}(t), \dot{\bar{x}}(t)) \geq (p(t)) \cdot v - L(t, \bar{x}(t), v)$ for all $v \in R^n$ a.e.

We thereby recover [14, Thm. 4] as a special case. To see this, set $h \equiv -1$ (so that the state constraint is automatically satisfied). Condition (v) implies $\mu \equiv 0$. We must have $\lambda > 0$ (for if $\lambda = 0$, then $\|p\| = 1$ by (i), yet (iv) implies $p \equiv 0$, a contradiction). Scaling the multipliers so that $\lambda = 1$ and noting Definition 2, we arrive at (ii)'–(iv)'.

(2) The need to formulate conditions (ii) and (iii) in terms of limiting normal cones to epigraph sets arises because $L(t, \cdot, \cdot)$ and l need not be Lipschitz continuous. If we further hypothesize that l is Lipschitz continuous near $(\bar{x}(0), \bar{x}(1))$ and $L(t, \cdot, \cdot)$ is Lipschitz continuous near $(\bar{x}(t), \dot{\bar{x}}(t))$ for a.e. t , the assertions of Theorem 3 are true with conditions (ii) and (iii) strengthened to

$$(ii)'' \quad \dot{p}(t) \in \text{co}\{\eta : (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \in \lambda \partial L(t, \bar{x}(t), \dot{\bar{x}}(t))\} \text{ a.e.},$$

$$(iii)'' \quad (p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)]) \in \lambda l(\bar{x}(0), \bar{x}(1)).$$

This is a consequence of the fact that, if $f : R^k \rightarrow R \cup \{+\infty\}$ is finite and Lipschitz continuous on a neighbourhood of $x \in R^k$, then

$$N_{\text{epi}\{f\}}(x, f(x)) = \{(\alpha\xi, -\alpha) : \alpha \geq 0, \xi \in \partial f(x)\}.$$

Proof. Fix $N > 0$ and let $k(\cdot)$ and δ be the corresponding bounds and constant of (H3). Define

$$\tilde{L}(t, w, v) := L(t, \bar{x}(t) + w, \dot{\bar{x}}(t) + v),$$

$$\tilde{l}(x, y) := l(\bar{x}(0) + x, \bar{x}(1) + y),$$

$$\tilde{h}(t, x) := h(t, \bar{x}(t) + x).$$

Choose a positive sequence $\epsilon_i \rightarrow 0$. For each i define a subset of $R^n \times L^1 \times L^1$ by

$$W := \{(\xi, v, w) : |v(t)| \leq N, \tilde{h}(t, x_{\xi,v}(t)) \leq 0, \text{ for all } t \in [0, 1], \|x_{\xi,v}\|_{W^{1,1}} \leq \epsilon\},$$

where $x_{\xi,v}(t) := \xi + \int_0^t v(s)ds$. Define the norm $\|\cdot\|_k$ on W by

$$\|(\xi, v, w)\|_k := |\xi| + \|kv\|_{L^1} + \|kw\|_{L^1}$$

and define the functional \tilde{J}_i on W

$$\begin{aligned} \tilde{J}_i(\xi, v, w) &:= \tilde{l}(x_{\xi,v}(0), x_{\xi,v}(1)) + \int_0^1 \tilde{L}(t, w(t), v(t))dt \\ &\quad + \epsilon_i^{-1} \int_0^1 k(t)|x_{\xi,v}(t) - w(t)|^2 dt. \end{aligned}$$

It can be shown that $(W, \|\cdot\|_k)$ is a complete metric space and, for each i , the functional \tilde{J}_i is lower semicontinuous on $(W, \|\cdot\|_k)$. Furthermore, there exists a positive sequence $\alpha_i \rightarrow 0$ such that for each i ,

$$\tilde{J}_i(0, 0, 0) \leq \inf_W \tilde{J}_i(\xi, v, w) + \alpha_i^2.$$

Since for each i , $(0, 0, 0)$ is an “ α_i^2 minimizer” for \tilde{J}_i over W , Ekeland’s variational principle tells us that there exists $(\xi_i, v_i, w_i) \in W$ which minimizes

$$\tilde{J}_i(\xi, v, w) + \alpha_i \|(\xi, v, w) - (\xi_i, v_i, w_i)\|_k$$

over W , and

$$\|(\xi_i, v_i, w_i)\|_k \leq \alpha_i.$$

Write $x_i = x_{\xi_i, v_i}$. We may then arrange by subsequence extraction that $(v_i, w_i) \rightarrow 0$ in L^1 and a.e., $x_i \rightarrow 0$ in $W^{1,1}$, and

$$\begin{aligned} \lim_{i \rightarrow \infty} \tilde{l}(x_i(0), x_i(1)) &= \tilde{l}(0, 0), \\ \lim_{i \rightarrow \infty} \tilde{L}(t, w_i(t), v_i(t)) &= \tilde{L}(t, 0, 0) \quad \text{a.e.} \end{aligned}$$

The arguments justifying these assertions are exactly the same as in [14]. Define

$$\begin{aligned} \tilde{L}_i(t, x, v, w) &:= \tilde{L}(t, v, w) + \epsilon_i^{-1} k(t) |x - w|^2 + \alpha_i k(t) |v - v_i(t)| + \alpha_i k(t) |w - w_i(t)|, \\ \tilde{l}_i(x, y) &:= \tilde{l}(x, y) + \alpha_i |x - x_i(0)|. \end{aligned}$$

The minimizing property of (ξ_i, v_i, w_i) can be stated as follows: $((x_i, y_i, z_i), (v_i, w_i))$ is a local minimizer for the optimal control problem

$$\begin{aligned} &\text{minimize } z(1) + \int_0^1 \tilde{L}_i(t, x(t), w(t), v(t)) dt, \\ &\text{subject to } \begin{cases} \dot{x}(t) = v(t), \dot{z}(t) = 0 & \text{a.e.}, \\ v(t) \in NB, w(t) \in R^n, \\ \tilde{h}(t, x(t)) \leq 0 \text{ for all } t \in [0, 1], \\ (x(0), x(1), z(1)) \in \text{epi}\{\tilde{l}_i\}, \end{cases} \\ &\text{and } \begin{cases} \dot{y}(t) = |v(t)| & \text{a.e.}, \\ y(0) = 0, |x(0)| + y(1) \leq \epsilon. \end{cases} \end{aligned}$$

Here $x_i(t) := x_{\xi_i, v_i}(t)$, $y_i(t) := \int_0^t |v_i(s)| ds$, and $z_i(t) := \tilde{l}_i(x_i(0), x_i(1))$.

This is an optimal control problem to which the maximum principle [2, Thm. 5.2.1] is applicable. A straightforward refinement to the proof in [2] permits us to express the transversality condition in terms of the limiting normal cone of the endpoint constraint set (in place of its convex hull, cf. [7]) and also to employ the subdifferential $\partial_x^> h(t, x)$ (defined above) in place of $\bar{\partial}_x h(t, x)$ of [2]; we shall use this version of the maximum principle.

Notice that, because the right endpoint constraint on y is inactive at $y = y_i(t)$ and because of the “decoupled” structure of the cost and dynamics in y and (x, z) , the costate arc component associated with y must be zero; we therefore drop it from the relationships.

The optimality conditions tell us that there exist an arc p_i , a constant $\lambda_i \geq 0$, a non-negative measure μ_i , and a μ_i -integrable function γ_i such that

- (A) $\|p_i\| + \lambda_i + \|\mu_i\| = 1$,
- (B) $-\dot{p}_i(t) = -2\lambda_i \epsilon_i^{-1} k(t)(x_i(t) - w_i(t))$,
- (C)

$$\begin{aligned} &(p_i(0), -[p_i(1) + \int_{[0,1]} \gamma_i(s) \mu_i(ds)], -\lambda_i) \\ &\in N_{\text{epi}\{\tilde{l}_i\}}(x_i(0), x_i(1), \tilde{l}_i(x_i(0), x_i(1))), \end{aligned}$$

- (D) $(v, w) \mapsto (p_i(t) + \int_{[0,t]} \gamma_i(s) \mu_i(ds)) \cdot v - \lambda_i \tilde{L}_i(t, x_i(t), w, v)$ achieves its maximum at $(v_i(t), w_i(t))$ over all $(v, w) \in NB \times R^n$, a.e.,
- (E) $\gamma_i(t) \in \partial_x^> \tilde{h}(t, x_i(t))$ μ_i -a.e. and $\text{supp}\{\mu_i\} \subset \{t : \tilde{h}(t, x_i(t)) = 0\}$.

(B) and (D) imply that

$$(8) \quad \begin{aligned} & (\dot{p}_i(t), p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)) \\ & \in \lambda_i \partial[\tilde{L}(t, w_i(t), v_i(t)) + \Psi_{NB}(v_i(t))] + \lambda_i \alpha_i k(t)(B \times B). \end{aligned}$$

Fix $v = v_i(t)$. From (B) and (D),

$$(9) \quad \dot{p}_i(t) \in \lambda_i \partial_w \tilde{L}(t, w_i(t), v_i(t)) + \lambda_i \alpha_i k(t)B.$$

Fix $w = w_i(t)$. (D) implies that

$$(10) \quad \begin{aligned} & (p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)) \cdot v_i(t) - \lambda_i \tilde{L}(t, w_i(t), v_i(t)) \\ & \geq (p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)) \cdot v - \lambda_i \tilde{L}(t, w_i(t), v) - \lambda_i \alpha_i k(t)|v - v_i(t)| \end{aligned}$$

for all $v \in NB$.

Since $\tilde{L}(t, \cdot, v)$ is Lipschitz continuous with rank $k(t)$ for all $v \in NB$, (9) implies that $|\dot{p}_i(t)| \leq 2k(t)$ for sufficiently large i . We now extract a subsequence such that $p_i \rightarrow p$ uniformly and $\dot{p}_i \rightarrow \dot{p}$ weakly in L^1 for some arc p . We can further extract a subsequence such that $\lambda_i \rightarrow \lambda$, $\mu_i \rightarrow \mu$ weakly*, $\gamma_i d\mu_i \rightarrow \gamma d\mu$ weakly* for some constant $\lambda \geq 0$, non-negative measure μ and μ -integrable function γ . Furthermore, $\gamma(t) \in \partial_x^> \tilde{h}(t, 0)$ μ -a.e. by (E) and Lemma 4.5 in [13]. We note that from (9) and from (10) we get

$$(11) \quad \dot{p}(t) \in \lambda \partial_w \tilde{L}(t, 0, 0),$$

$$(12) \quad -\lambda \tilde{L}(t, 0, 0) \geq (p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \cdot v - \lambda \tilde{L}(t, 0, v) \text{ for all } v \in NB.$$

Assertion (i) of the theorem is valid because $\|p_i\| + \|\mu_i\| + \lambda_i = 1$.

To show (v) we only need to prove $\text{supp}\{\mu\} \subset \{t : \tilde{h}(t, 0) = 0\}$. Assume it were not true; then there would exist an $\epsilon > 0$ such that $\mu(E) > 0$, where $E = \{t : \tilde{h}(t, 0) < -\epsilon\}$. Since $\tilde{h}(t, \cdot)$ is Lipschitz continuous with constant rank k_h and $x_i \rightarrow 0$ uniformly, for i sufficiently large (independent of t) we would have

$$\{t : \tilde{h}(t, 0) < -\epsilon\} \subset \{t : \tilde{h}(t, x_i(t)) \leq -\epsilon/2\}.$$

So $\mu_i(E) = 0$. But E is a relatively open set (since h is upper semicontinuous and \bar{x} is a continuous arc) and $\mu_i \rightarrow \mu$ weakly*. [13, Lemma 4.4] then tells us that $\mu(E) = \lim_{i \rightarrow \infty} \mu_i(E) = 0$, which is a contradiction. (v) of the theorem is proved.

We now confirm properties (ii) and (iii) of the theorem. The following conditions may be obtained from Mazur's and Carathéodory's theorems. For each i there exist integers $0 \leq k_{i0} \leq \dots \leq k_{in}$ and a convex combination $\{\alpha_{i0}, \dots, \alpha_{in}\}$ such that $\sum_{j=0}^n \alpha_{ij} \dot{p}_{i+k_{ij}}(t)$, $i = 1, 2, \dots$, converges strongly in L^1 to $\dot{p}(t)$. We can arrange by subsequence extraction that convergence is a.e. Fix t outside a suitable measure zero set. For each j , α_{ij} , $i = 1, 2, \dots$, and $\dot{p}_{i+k_{ij}}(t)$, $i = 1, 2, \dots$, are bounded. A further subsequence extraction ensures that (for each j) they have limits α_j and $\bar{p}_j(t)$. Since $v_i(t), w_i(t) \rightarrow 0$, $\tilde{L}(t, w_i(t), v_i(t)) + \Psi_{NB}(v_i(t)) \rightarrow \tilde{L}(t, 0, 0)$ a.e.,

$p_i(t) \rightarrow p(t)$, and $\tilde{l}_i(x_i(0), x_i(1)) \rightarrow \tilde{l}(0, 0)$, we have

$$(\tilde{p}_j(t), p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \in \lambda\partial\tilde{L}(t, 0, 0) \text{ for } j = 0, \dots, n,$$

$$(p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)]) \in \lambda\partial\tilde{l}(0, 0),$$

if $\lambda > 0$, or

$$(\tilde{p}_j(t), p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \in \partial^\infty\tilde{L}(t, 0, 0) \text{ for } j = 0, \dots, n,$$

$$(p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)]) \in \partial^\infty\tilde{l}(0, 0),$$

if $\lambda = 0$. For either case we deduce that

$$(\tilde{p}_j(t), p(t) + \int_{[0,t)} \gamma(s)\mu(ds), -\lambda) \in N_{\text{epi}\{\tilde{L}(t, \cdot, \cdot)\}}(0, 0, \tilde{L}(t, 0, 0)) \text{ for } j = 0, \dots, n,$$

$$(p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)], -\lambda) \in N_{\text{epi}\{\tilde{l}\}}(0, 0, \tilde{l}(0, 0)).$$

But the α_j 's define a convex combination, so

$$\dot{p}(t) = \sum_{j=0}^n \alpha_j \tilde{p}_j(t) \in \text{co}\{\eta : (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds), -\lambda) \in N_{\text{epi}\{\tilde{L}(t, \cdot, \cdot)\}}(0, 0, \tilde{L}(t, 0, 0))\} \text{ a.e.}$$

These are precisely the assertions of the theorem, except that in the condition (iv), the inequality holds only for $v \in \dot{x}(t) + NB$.

Take $N_i \rightarrow \infty$. Let $(p_i, \lambda_i, \mu_i, \gamma_i)$ denote the multipliers when $N = N_i$. We deduce from (11) that the \dot{p}_i 's are uniformly integrably bounded. We can then show as before that along a subsequence $p_i \rightarrow p$ uniformly and $\dot{p}_i \rightarrow \dot{p}$ weakly in L^1 to some arc p . Arguing as before, we arrive at a strengthened form of (12) for some $\lambda \geq 0$, a non-negative measure μ and a μ -integrable function γ in which R^n replaces NB . This is condition (iv). The proof is complete. \square

3. PROBLEMS WITH DYNAMIC AND ENDPOINT CONSTRAINTS

This section provides the main results of the paper. These concern the optimization problem

$$\begin{aligned} &\text{Minimize } g(x(0), x(1)) \\ &\text{over absolutely continuous arcs } x : [0, 1] \rightarrow R^n \text{ which satisfy} \\ \text{(P)} \quad &\dot{x}(t) \in F(t, x(t)), \quad \text{a.e.}, \\ &(x(0), x(1)) \in C, \\ &h(t, x(t)) \leq 0 \text{ for all } t. \end{aligned}$$

Here $g : R^n \times R^n \rightarrow R$ and $h : [0, 1] \times R^n \rightarrow R$ are given functions, $F : [0, 1] \times R^n \rightrightarrows R^n$ is a given multifunction, and $C \subset R^n \times R^n$ is a given set.

The first theorem, an extension of [4] to allow for state constraints, asserts the validity of the extended Euler Lagrange, Weierstrass, and transversality conditions for problems in which F is possibly non-convex valued.

Theorem 4. Let \bar{x} be a $W^{1,1}$ local minimizer for (P). Suppose that the following hypotheses are satisfied

- (G1) g is Lipschitz continuous on a neighbourhood of $(\bar{x}(0), \bar{x}(1))$ with rank k_g .
- (G2) $F(t, x)$ is a nonempty, closed set for each (t, x) . The graph of $F(t, \cdot)$ is closed for a.e. t . F is $\mathcal{L} \times \mathcal{B}$ measurable. There exists $k \in L^1$ for some $\epsilon > 0$ and $\beta \geq 0$, such that

$$(13) \quad F(t, x') \cap (\dot{\bar{x}}(t) + NB) \subset F(t, x) + (k(t) + \beta N)|x' - x|B$$

for all $N \geq 0$ and $x, x' \in \bar{x}(t) + \epsilon B$ a.e.

- (G3) h is upper semicontinuous on $\{(t, x) : t \in [0, 1], x \in \bar{x}(t) + \epsilon B\}$ and there is a constant k_h such that $|h(t, x) - h(t, x')| \leq k_h|x - x'|$ for all t in $[0, 1]$ and all $x, x' \in \bar{x}(t) + \epsilon B$.

Then there exist an arc p , a non-negative number λ , a non-negative measure μ , and a μ -integrable function γ , such that

- (i) $\lambda + \|p\| + \|\mu\| = 1$,
- (ii) $\dot{p}(t) \in \text{co}\{\eta : (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \in N_{\text{Gr}F(t,\cdot)}(\bar{x}(t), \dot{\bar{x}}(t))\}$ a.e.,
- (iii) $(p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)]) \in \lambda \partial g(\bar{x}(0), \bar{x}(1)) + N_C(\bar{x}(0), \bar{x}(1))$,
- (iv) $(p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \cdot \dot{\bar{x}}(t) \geq (p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \cdot v$ for all $v \in F(t, \bar{x}(t))$ a.e.,
- (v) $\gamma(t) \in \partial_x^> h(t, \bar{x}(t))$ μ -a.e. and $\text{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}$.

Comments. (1) The one-sided Lipschitz continuity hypothesis (G2) (invoked by Loewen and Rockafellar [7] and others) is a refinement of the Lipschitz continuity hypothesis

$$(14) \quad F(t, x') \subset F(t, x) + k(t)|x' - x|B$$

commonly employed in the earlier literature. (G2) is satisfied by many unbounded F 's of interest (defined for instance in terms of epigraph sets), whereas (14) is not. This point is discussed in [7].

(ii) The necessary conditions of Theorem 4 are of interest only when the state constraint is non-degenerate, in the sense that $0 \notin \partial_x^> h(t, \bar{x}(t))$ for all t such that $h(t, \bar{x}(t)) = 0$. This is because the necessary conditions are satisfied along any arc \bar{x} satisfying

$$0 \in \partial_x^> h(t', \bar{x}(t')) \quad \text{and} \quad h(t', \bar{x}(t')) = 0,$$

for some time t' , with the choice of multipliers $\mu = \delta_{\{t'\}}$ (the unit measure concentrated on $\{t'\}$), $p(t) \equiv 0$ and $\lambda = 0$; unless we exclude a priori the existence of such arcs, the necessary conditions convey no useful information about the minimizer.

(iii) In the problem formulation adopted here pathwise restrictions on the state variable take the form of a functional inequality constraint

$$h(t, x(t)) \leq 0 \quad \text{for all } t \in [0, 1].$$

Our formulation allows a number of commonly occurring state constraint descriptions as special cases. For example, problems with multiple state constraints

$$h_1(t, x(t)) \leq 0, \dots, h_k(t, x(t)) \leq 0$$

are subsumed in the "single" inequality constraint of problem (P) by setting

$$h(t, x(t)) := \max_i \{h_i(t, x(t))\}.$$

If the constraint operates only on a closed subset $A \subset [0, 1]$, i.e.

$$\tilde{h}(t, x(t)) \leq 0 \text{ for all } t \in A,$$

we can again treat this as a constraint of the form

$$h(t, x(t)) \leq 0 \text{ for all } t \in [0, 1]$$

by setting

$$h(t, x) := \begin{cases} \tilde{h}(t, x) & \text{for } t \in A, \\ -K & \text{for } t \notin A, \end{cases}$$

for suitably large K . Notice the significance of allowing h to be upper semicontinuous (but possibly discontinuous), so that the h 's constructed in this way are covered by the theorem.

(iv) The reason for adopting the hybrid generalized gradient $\partial_x^> h(t, x)$ is to deal with certain problems involving a state constraint specified as

$$(15) \quad x(t) \in X(t).$$

These too we can reduce to Problem (P) by setting

$$(16) \quad h(t, x) := d_{X(t)}(x).$$

For a large class of $X(t)$'s $0 \notin \partial_x^> h(t, x)$, and so the degeneracy remarked on in comment (i) is avoided. The standard choice of generalized gradient $\text{co}\partial h(t, \cdot)$ is not suitable in this context because $0 \in \text{co}\partial_x h(t, \cdot)$ on $X(t)$.

In some of the literature on state constrained problems (see e.g. Loewen & Rockafellar's treatment of problems with convex valued F 's), necessary conditions are derived directly for problems with state constraints specified as (15). While of course there is a great deal of overlap, the functional and set inclusion approaches to describing the state constraints both have individual advantages: there are examples of h 's satisfying hypothesis (G3) for which $X(t) := \{x : h(t, x) \leq 0\}$ fails to satisfy the hypotheses of [7], and examples of $X(t)$'s for which the necessary conditions of [7] are nondegenerate but the choice of h in (16) gives rise to degenerate necessary conditions.

Theorem 4 will be proved in Section 4.

If we now impose the hypothesis that F is convex valued, to what extent can we relax the hypotheses in other respects so that the assertions of Theorem 4 remain valid? Because in other areas of nonlinear analysis specialization to the convex case often means that hypotheses of a global nature can be weakened to local ones, we might hope in the convex case to replace (13) in hypothesis (G2) of Theorem 4 by the following condition:

There exists $\epsilon > 0$ such that, for all $x, x' \in \bar{x}(t) + \epsilon B$,

$$(17) \quad F(t, x') \cap (\dot{\bar{x}}(t) + \epsilon B) \subset F(t, x) + k(t)|x - x'|B \quad \text{a.e.}$$

Whether this is an adequate substitute for (13) in the convex case remains an open question. However the following theorem tells us that a slight strengthening of (17) at least will serve. In fact the theorem provides two modified versions of (17) with the desired consequences. The first condition (G2)' is a new criterion for validity of the optimality conditions even in the absence of state constraints. Choosing the second, (G2)'' below, and specializing to the case when no state constraint operates, we arrive at necessary conditions previously derived by Ioffe [4]. Notice that (G2)' is strictly weaker than (G2), while the Ioffe-type condition is not.

Theorem 5. *Let \bar{x} be a $W^{1,1}$ local minimizer for (P). Assume that $F(t, x)$ is a convex set for each $(t, x) \in [0, 1] \times R^n$. Assume further that hypotheses (G1) and (G3) of Theorem 4 are satisfied in addition to either of the following “local” versions of (G2):*

(G2)' *There exists $k \in L^1$ and $\epsilon > 0$ such that for a.e. t*

$$F(t, x') \cap (\dot{\bar{x}}(t) + \epsilon k(t)B) \subset F(t, x) + k(t)|x - x'|B \quad \text{for all } x, x' \in \bar{x}(t) + \epsilon B,$$

or

(G2)'' *There exist $\bar{K} > 0$ and $\epsilon > 0$ such that the following two conditions are satisfied for a.e. t :*

$$F(t, x') \cap (\dot{\bar{x}}(t) + \epsilon B) \subset F(t, x) + k(t)|x - x'|B, \quad \text{for all } x, x' \in \bar{x}(t) + \epsilon B,$$

and

$$\inf\{|v - \dot{\bar{x}}(t)| : v \in F(t, x)\} \leq \bar{K}|x - \bar{x}(t)| \quad \text{for all } x \in \bar{x}(t) + \epsilon B.$$

Then there exist an arc p , a constant $\lambda \geq 0$, a nonnegative measure μ and a μ -integrable function γ such that

- (i) $\|p\| + \|\mu\| + \lambda = 1$,
- (ii) $\dot{p}(t) \in \text{co}\{\eta : (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \in N_{G \circ F(t, \cdot)}(\bar{x}(t), \dot{\bar{x}}(t))\}$ a.e.,
- (iii) $(p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)]) \in \lambda \partial g(\bar{x}(0), \bar{x}(1)) + N_C(\bar{x}(0), \bar{x}(1))$,
- (iv) $\gamma(t) \in \partial_x^> h(t, \bar{x}(t))$ μ -a.e. and $\text{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}$.

Under the stated hypotheses, (ii) implies that (p, μ, γ) also satisfy the conditions

- (a) $\dot{p}(t) \in \text{co}\{-\xi : (\xi, \dot{\bar{x}}(t)) \in \partial H(t, \bar{x}(t), p(t) + \int_{[0,t)} \gamma(s)\mu(ds))\}$ a.e., where ∂H denotes the limiting subdifferential of the Hamiltonian

$$H(t, x, p) = \max\{p \cdot v : v \in F(t, x)\}$$

in the (x, p) variables, and

- (b) $(p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \cdot \dot{\bar{x}}(t) \geq (p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \cdot v$ for all $v \in F(t, \bar{x}(t))$ a.e.

The necessary conditions take a somewhat different form from those provided by Loewen and Rockafellar [7], which also apply to problems with state constraints when F is convex valued, because in [7] the state constraint is formulated as a set inclusion. Distinctions between the two formulations were discussed following the statement of Theorem 4. The hypotheses (G2)' (or in some respects (G2)'' on the dynamics are however significantly weaker than the hypotheses governing the x dependence of F in [7], and the Euler Lagrange inclusion of [7] is replaced here by its sharper “extended” version.

We mention that the hypotheses of Theorem 5 are somewhat weaker than those under which Rockafellar’s dualization theorem [12] holds, and it is not known whether conditions (a) and (ii) are equivalent. However we can infer from Ioffe’s partial dualization result [4], which is valid under the weaker hypotheses operating here, that (ii) implies (a); the role of partial dualization is precisely to deal with applications of this nature.

Theorem 5 is proved in Section 5.

4. PROOF OF THEOREM 4

A crucial role in the proof of the theorem is played by the “distance function” associated with F ,

$$\rho_F(x, v) := \inf\{|v - y| : y \in F(x)\},$$

and it is convenient at the outset to explore some of its properties. The following two lemmas are based on related results of Ioffe [4], but supply more detailed information about the neighbourhoods and Lipschitz constants involved.

Lemma 6. *Take a multifunction $F : R^n \rightrightarrows R^n$ and a point $(\bar{x}, \bar{v}) \in \text{Gr}F$. Assume that F has values closed sets and that there exist $\epsilon > 0$, $r > 0$, $k > 0$ and $K > 0$ such that*

$$(18) \quad F(x') \cap (\bar{v} + rB) \subset F(x) + k|x' - x|B$$

for all $x', x \in \bar{x} + \epsilon B$ and

$$(19) \quad \rho_F(x, \bar{v}) \leq K|x - \bar{x}|$$

for all $x \in \bar{x} + \epsilon B$. Then

- (a) For each $x \in R^n$, $\rho_F(x, \cdot)$ is Lipschitz continuous with Lipschitz constant 1 and, for each $v \in \bar{v} + (r/3)B$, we have that $\rho_F(\cdot, v)$ is Lipschitz continuous on $\bar{x} + \min\{\epsilon, r/(3K)\}B$ with Lipschitz constant k .
- (b) Take any $(x, v) \in R^n \times R^n$ such that $|v - \bar{v}| < r/3$ and $|x - \bar{x}| < \min\{\epsilon, r/(3K)\}$ and $(w, p) \in R^n \times R^n$ such that

$$(w, p) \in \partial\rho_F(x, v).$$

Then $|w| \leq k$ and $|p| \leq 1$. Furthermore, $v \in F(x)$ implies $(w, p) \in N_{\text{Gr}F}(x, v)$ and $v \notin F(x)$ implies $p = (v - u)/|v - u|$ for some $u \in F(x)$ such that $|v - u| = \min\{|v - y| : y \in F(x)\}$.

Proof. (a) For fixed x , $\rho_F(x, \cdot)$ is the Euclidean distance function for the set $F(x)$ and, by the properties of distance functions, is Lipschitz continuous with Lipschitz constant 1. Choose any $x', x \in \bar{x} + \min\{\epsilon, r/(3K)\}B$. Let u' be the closest point to v in $F(x')$, i.e., $|v - u'| = \rho_F(x', v)$. Then by (19)

$$\begin{aligned} |u' - \bar{v}| &\leq |v - \bar{v}| + |v - u'| \leq |v - \bar{v}| + |v - \bar{v}| + \rho_F(x', \bar{v}) \\ &\leq 2|v - \bar{v}| + K|x' - \bar{x}| \leq r. \end{aligned}$$

We have shown that $u' \in \bar{v} + rB$. By (18), then there exists $u \in F(x)$ such that $|u' - u| \leq k|x' - x|$. It follows that

$$\rho_F(x, v) \leq |v - u| \leq |v - u'| + |u' - u| \leq \rho_F(x', v) + k|x' - x|.$$

Since the variables x' and x are interchangeable, we conclude that

$$|\rho_F(x, v) - \rho_F(x', v)| \leq k|x - x'|.$$

This is the desired Lipschitz continuity property.

(b) Apply the results of part (a). These tell us that ρ_F is Lipschitz continuous on the closed neighbourhood \mathcal{N} of (\bar{x}, \bar{v}) , where

$$\mathcal{N} := (\bar{x}, \bar{v}) + (\min\{\epsilon, r/(3K)\}B) \times (r/3B)$$

Scrutiny of the proof of (a) also reveals that

$$(20) \quad (x, v) \in \mathcal{N} \text{ implies } |u - \bar{v}| \leq r \text{ for any closest point } u \text{ to } v \text{ in } F(x).$$

Fix (x, v) in the interior of \mathcal{N} . Since (as we have shown) ρ_F is Lipschitz continuous on a neighbourhood of (x, v) , it is meaningful to talk of the limiting sub-differential $\partial\rho_F(x, v)$. Take any $(w, p) \in \partial\rho_F(x, v)$. We know that there exist $(x_i, v_i) \rightarrow (x, v)$, $(w_i, p_i) \rightarrow (w, p)$, $\epsilon_i \downarrow 0$ and a sequence of positive numbers M_i such that

$$(21) \quad w_i \cdot (x' - x_i) + p_i \cdot (v' - v_i) \leq \rho_F(x', v') - \rho_F(x_i, v_i) + M(|x' - x_i|^2 + |v' - v_i|^2)$$

for all $(x', v') \in (x_i, v_i) + \epsilon_i B$.

By part (a), $\rho_F(\cdot, v_i)$ and $\rho_F(x_i, \cdot)$ are Lipschitz continuous (on appropriate neighbourhoods of (x_i, v_i)) with Lipschitz constants k and 1 respectively, for sufficiently large i . We easily deduce from (21) that $|w_i| \leq k$ and $|p_i| \leq 1$. Passing to the limit as $i \rightarrow \infty$ gives

$$|w| \leq k \text{ and } |p| \leq 1.$$

Suppose that $v \notin F(x)$. It follows from (18) that $v_i \notin F(x_i)$ for all sufficiently large i . Let u_i be a closest point to v_i in $F(x_i)$. In the inequality (21) set $x' = x_i$ and $v' = v_i + \delta(u_i - v_i)$. We see that for $\delta > 0$ sufficiently small

$$\delta p_i \cdot (u_i - v_i) \leq (1 - \delta)|v_i - u_i| - |v_i - u_i| + M_i \delta^2 |u_i - v_i|^2.$$

Divide the inequality by δ and pass to the limit as $\delta \downarrow 0$. This gives

$$p_i \cdot (v_i - u_i) \geq |v_i - u_i|.$$

Since $|p_i| \leq 1$ and $v_i - u_i \neq 0$ for i sufficiently large, this relationship implies

$$p_i = (v_i - u_i)/|v_i - u_i|.$$

Now $\{u_i\}$ is a bounded sequence (see (21)). By restricting attention to a subsequence we can arrange that $u_i \rightarrow u$ for some $u \in \bar{v} + rB$. From (18), it follows that $u \in F(x)$. Since $v \notin F(x)$, it follows that $p_i \rightarrow p$, where

$$(22) \quad p := (v - u)/|v - u|.$$

Let \tilde{u} be a closest point to v in $F(x)$. Then by (18) there exist points $\tilde{u}_i \in F(x_i)$, $i = 1, 2, \dots$, such that $\tilde{u}_i \rightarrow \tilde{u}$. Since u_i is a closest point to v_i in $F(x_i)$ it follows that $|u_i - v_i| \leq |v_i - \tilde{u}_i|$ for all i . Passing to the limit, we obtain

$$|v - u| \leq |v - \tilde{u}| = \min\{|v - y| : y \in F(x)\}.$$

Recalling that $u \in F(x)$, we see that u is a closest point to v in $F(x)$. (22) now provides the desired representation of p .

Next suppose that $v \in F(x)$. For each i , again take u_i to be a closest point to v_i in $F(x_i)$ (we allow the possibility that $v_i \in F(x_i)$ for some values of i). It follows from (18) and (20) that $u_i \rightarrow v$ as $i \rightarrow \infty$. Fix i sufficiently large. Choose any (x', u') in $\text{Gr}F$ close to (x_i, u_i) . Now insert $(x', v' = u' + v_i - u_i)$ into (21). We arrive at

$$\begin{aligned} w_i \cdot (x' - x_i) + p_i \cdot (u' - u_i) \\ \leq \rho_F(x', u' + v_i - u_i) - \rho_F(x_i, v_i) + M_i(|x' - x_i|^2 + |u' - u_i|^2). \end{aligned}$$

Notice however that

$$\rho_F(x', u' + v_i - u_i) \leq |u' + v_i - u_i - u'| = |v_i - u_i| = \rho_F(x_i, v_i).$$

It follows that

$$w_i \cdot (x' - x_i) + p_i \cdot (u' - u_i) \leq M_i(|x' - x_i|^2 + |u' - u_i|^2)$$

for all (x', u') in $\text{Gr}F$ sufficiently close to (x_i, u_i) . Hence $(w_i, p_i) \in N_{\text{Gr}F}(x_i, u_i)$. However, $\text{Gr}F \cap ((x, v) + \alpha B)$ is a closed set for some $\alpha > 0$. Recalling that $(x_i, u_i) \rightarrow (x, v)$, we deduce from the closure properties of the normal cone that $(w, p) \in N_{\text{Gr}F}(x, v)$. This is what we set out to prove. \square

Lemma 7. *Take a multifunction $F : R^n \rightrightarrows R^n$ and a point $(\bar{x}, \bar{v}) \in \text{Gr}F$. Assume that F has values closed sets and there exist $\epsilon > 0$, $\beta > 0$ and $k > 0$ such that*

$$(23) \quad F(x') \cap (\bar{v} + NB) \subset F(x) + (k + \beta N)|x' - x|B$$

for all $x', x \in \bar{x} + \epsilon B$ and all $N > 0$. Then, for any $v \in R^n$, $\rho_F(\cdot, v)$ is Lipschitz continuous on $\bar{x} + \epsilon B$ with Lipschitz constant $k(1 + \beta\epsilon) + 2\beta|v - \bar{v}|$.

Proof. Take any $x', x \in \bar{x} + \epsilon B$ and any $v \in R^n$. Let u' be a closest point to v in $F(x')$. By (23) there exist $w' \in F(x')$ such that $|w' - \bar{v}| \leq k|x' - \bar{x}|$. It follows from the triangle inequality that

$$|v - u'| \leq |v - w'| \leq |v - \bar{v}| + |\bar{v} - w'| \leq |v - \bar{v}| + k|x' - \bar{x}| \leq |v - \bar{v}| + k\epsilon,$$

and hence

$$|u' - \bar{v}| \leq |u' - v| + |v - \bar{v}| \leq 2|v - \bar{v}| + k\epsilon.$$

A further appeal to (17) now gives existence of $u \in F(x)$ such that

$$|u - u'| \leq (k + \beta(2|v - \bar{v}| + k\epsilon))|x - x'|.$$

From the triangle inequality

$$\rho_F(x, v) \leq |v - u| \leq |v - u'| + |u - u'| \leq \rho_F(x', v) + (k + \beta(2|v - \bar{v}| + k\epsilon))|x' - x|.$$

Since the roles of x' and x can be reversed, we conclude that

$$|\rho_F(x', v) - \rho_F(x, v)| \leq (k + \beta(2|v - \bar{v}| + k\epsilon))|x' - x|.$$

\square

We are now ready to prove Theorem 4. By reducing the size of ϵ if necessary we can arrange that \bar{x} is minimizing in relation to arcs x satisfying $\|x - \bar{x}\|_{W^{1,1}} \leq \epsilon$, and that g is Lipschitz continuous on $(\bar{x}(0), \bar{x}(1)) + \epsilon(B \times B)$ with Lipschitz constant k_g . Define \mathcal{S} , the set of “admissible” arcs close to \bar{x} , to be

$$\mathcal{S} := \{x \in W^{1,1} : \dot{x}(t) \in F(t, x(t)) \text{ a.e., } h(t, x(t)) \leq 0, \text{ for all } t, \\ (x(0), x(1)) \in C, \|x - \bar{x}\|_{W^{1,1}} \leq \epsilon\}$$

and define the functional $\tilde{g} : W^{1,1} \rightarrow R$ by

$$\tilde{g}(x) := g(x(0), x(1)).$$

Evidently, \tilde{g} is Lipschitz continuous on \mathcal{S} , with Lipschitz constant $2k_g$, with respect to the $W^{1,1}$ norm. We know that \bar{x} is a minimizer for

$$\text{Minimize } \tilde{g}(x) \text{ over } x \in \mathcal{S} \subset \{x \in W^{1,1} : h(t, x(t)) \leq 0, \|x - \bar{x}\|_{W^{1,1}} \leq \epsilon\}.$$

It follows from [3, p. 51] that \bar{x} is a minimizer for the problem

$$\begin{aligned} &\text{Minimize } \tilde{g}(x) + 2k_g \inf\{\|x - y\|_{W^{1,1}} : y \in \mathcal{S}\} \\ &\text{over arcs } x \text{ which satisfy } h(t, x(t)) \leq 0 \text{ for all } t \text{ and } \|x - \bar{x}\|_{W^{1,1}} \leq \epsilon. \end{aligned}$$

Following [4], we now identify the two possible situations which can arise:

(a) There exist $\epsilon' \in (0, \epsilon)$ and $K > 0$, such that for any $x \in W^{1,1}$ satisfying $h(t, x(t)) \leq 0$ for all $t \in [0, 1]$ and $\|x - \bar{x}\|_{W^{1,1}} \leq \epsilon'$ we have

$$\inf\{\|x - y\|_{W^{1,1}} : y \in \mathcal{S}\} \leq K \left[\int_0^1 \rho_F(t, x(t), \dot{x}(t)) dt + d_C(x(0), x(1)) \right].$$

(b) There exists a sequence of arcs \bar{x}_i such that $h(t, \bar{x}_i(t)) \leq 0$ for all $t \in [0, 1], i = 1, 2, \dots, \bar{x}_i \rightarrow \bar{x}$ in $W^{1,1}$ as $i \rightarrow \infty$, and

$$(24) \quad J(\bar{x}_i) < (2i)^{-1} \inf\{\|\bar{x}_i - y\|_{W^{1,1}} : y \in \mathcal{S}\}$$

for $i = 1, 2, \dots$, where

$$J(x) := \int_0^1 \rho_F(t, x(t), \dot{x}(t)) dt + d_C(x(0), x(1)).$$

We deal first with the more straightforward case (a). Here \bar{x} is a $W^{1,1}$ local minimizer of $g(x(0), x(1)) + 2k_g K J(x)$ over $x \in W^{1,1}$ satisfying $h(t, x(t)) \leq 0$. This is the finite Lagrangian problem (7) discussed in Section 2 with

$$\begin{aligned} L(t, x, v) &:= 2k_g K \rho_F(t, x, v), \\ l(x, y) &:= g(x, y) + 2k_g K d_C(x, y). \end{aligned}$$

The hypotheses are satisfied for existence of an arc p , a non-negative number λ , a non-negative measure μ , and a μ -integrable function γ satisfying (i)-(v) of Theorem 3. (i) and (v) of Theorem 4 are clearly satisfied. (iv) is obtained by noticing that $\rho_F(t, \bar{x}(t), v) = 0$ if $v \in F(t, \bar{x}(t))$. That (ii) and (iii) are also satisfied is clear from the comments to be found after Theorem 3 and Lemma 7.

It remains to address case (b). Notice first of all that the functional J is continuous on

$$W := \{x \in W^{1,1} : \|x - \bar{x}\|_{W^{1,1}} \leq \epsilon, h(t, x(t)) \leq 0\}$$

with respect to the strong $W^{1,1}$ topology. Indeed if $x_i \rightarrow x$ in W , then $x_i \rightarrow x$ uniformly and $\dot{x}_i \rightarrow \dot{x}$ strongly in L^1 . Consequently, in view of Lemma 7

$$\begin{aligned} |J(x_i) - J(x)| &\leq \int_0^1 |\rho_F(t, x_i, \dot{x}_i) - \rho_F(t, x, \dot{x})| dt \\ &\quad + |d_C(x_i(0), x_i(1)) - d_C(x(0), x(1))| \\ &\leq ((1 + \beta)\|k\|_{L^1} + 2\beta\|\dot{x} - \dot{\bar{x}}\|_{L^1})\|x_i - x\|_{L^\infty} \\ &\quad + \|\dot{x}_i - \dot{x}\|_{L^1} + 2\|x_i - x\|_{L^\infty}. \end{aligned}$$

The right side has limit zero as $i \rightarrow \infty$. Continuity is confirmed.

Define

$$a_i := \inf\{\|\bar{x}_i - y\|_{W^{1,1}} : y \in \mathcal{S}\}.$$

Since the inequality (24) is strict and J is non-negative valued, $a_i > 0$ for all i . Since $\bar{x}_i \rightarrow \bar{x}$ in $W^{1,1}$ and $\bar{x} \in \mathcal{S}$, it follows that $a_i \rightarrow 0$ as $i \rightarrow \infty$. We see too that $J(\bar{x}_i) \leq (2i)^{-1} a_i$, from which we conclude that \bar{x}_i is a “ $(2i)^{-1} a_i$ -minimizer” of J

on W . According to Ekeland's theorem, for each i , there exist $x_i \in W$ such that x_i is a minimizer for

$$\begin{aligned} &\text{minimize } \int_0^1 \rho_F(t, x(t), \bar{x}(t)) dt + d_C(x(0), x(1)) \\ &\quad + i^{-1}(|x(0) - x_i(0)| + \int_0^1 |\dot{x}(t) - \dot{x}_i(t)| dt) \\ &\text{over arcs } x \text{ which satisfy } h(t, x(t)) \leq 0 \text{ for all } t \text{ and } \|x_i - \bar{x}_i\|_{W^{1,1}} \leq a_i/2. \end{aligned}$$

Recalling that $a_i := \inf\{\|\bar{x}_i - y\| : y \in \mathcal{S}\}$ we conclude from this last inequality that $x_i \notin \mathcal{S}$. It follows that, for each i , either $(x_i(0), x_i(1)) \notin C$ or $\dot{x}_i(t) \notin F(t, x_i(t))$ on a set of positive measure. Notice too that $x_i \rightarrow \bar{x}$ strongly in $W^{1,1}$, since $\bar{x}_i \rightarrow \bar{x}$ in $W^{1,1}$ and $\|x_i - \bar{x}_i\|_{W^{1,1}} \rightarrow 0$ as $i \rightarrow \infty$.

This is once again an example of the finite Lagrangian problem (7) investigated in Section 2 with

$$\begin{aligned} L(t, x, v) &:= \rho_F(t, x, v) + i^{-1}|v - \dot{x}_i(t)|, \\ l(x, y) &:= d_C(x, y) + i^{-1}|x - x_i(0)|. \end{aligned}$$

For i sufficiently large $\rho_F(t, \cdot, \cdot)$ is Lipschitz continuous "near" $(x_i(t), \dot{x}_i(t))$, in view of Lemmas 6 and 7. By Theorem 3 and the succeeding comments we can find $(p_i, \lambda_i, \mu_i, \gamma_i)$ such that

- (A) $\lambda_i + \|p_i\| + \|\mu_i\| = 1$,
- (B) $\dot{p}_i(t) \in \text{co}\{\eta : (\eta, p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)) \in \lambda_i \partial \rho_F(t, x_i(t), \dot{x}_i(t)) + \{0\} \times (\lambda_i i^{-1} B)\}$ a.e.,
- (C) $(p_i(0), -[p_i(1) + \int_{[0,1]} \gamma_i(s)\mu_i(ds)]) \in \lambda_i \partial d_C(x_i(0), x_i(1)) + \{\lambda_i i^{-1} B\} \times \{0\}$,
- (D)

$$\begin{aligned} &(p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)) \cdot \dot{x}_i(t) - \lambda_i \rho_F(t, x_i(t), \dot{x}_i(t)) \\ &\geq (p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)) \cdot v - \lambda_i \rho_F(t, x_i(t), v) - \lambda_i i^{-1}|v - \dot{x}_i(t)| \end{aligned}$$

for all $v \in R^n$ a.e.,

- (E) $\gamma_i(t) \in \partial_x^> h(t, x_i(t))$ μ_i -a.e. and $\text{supp}\{\mu_i\} \subset \{t : h(t, x_i(t)) = 0\}$.

Recalling Lemma 7, we deduce from (B) that

$$|\dot{p}_i(t)| \leq \lambda_i[(1 + \beta\epsilon)k(t) + 2\beta|\dot{x}_i(t) - \dot{\bar{x}}(t)|].$$

But $x_i \rightarrow \bar{x}$ in $W^{1,1}$ (which implies $\dot{x}_i \rightarrow \dot{\bar{x}}$ strongly in L^1) and $\|p_i\| \leq 1$. It follows that there exists an arc p such that $p_i \rightarrow p$ uniformly and $\dot{p}_i \rightarrow \dot{p}$ weakly in L^1 . By further subsequence extraction we can arrange that

$$\lambda_i \rightarrow \lambda', \quad \mu_i \rightarrow \mu \text{ weakly}^* \text{ and } \gamma_i d\mu_i \rightarrow \gamma d\mu \text{ weakly}^*$$

for some $\lambda' \geq 0$, non-negative measure μ and μ -integrable function γ satisfying

$$\gamma(t) \in \partial_x^> h(t, \bar{x}(t)) \quad \mu \text{ a.e. and } \text{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}.$$

If $0 \leq \lambda' < 1$, then $\|p\| + \|\mu\| = 1 - \lambda' > 0$ and

$$\begin{aligned} \dot{p}(t) &\in \text{co}\{\eta : (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \in \lambda' \partial \rho_F(t, \bar{x}(t), \dot{\bar{x}}(t))\} \quad \text{a.e.}, \\ (p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)]) &\in \lambda' \partial d_C(\bar{x}(0), \bar{x}(1)), \\ (p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \cdot \dot{\bar{x}}(t) &\geq (p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \cdot v \quad \text{for all } v \in F(t, \bar{x}(t)) \text{ a.e.} \end{aligned}$$

By scaling the multipliers we can arrange that $\|p\| + \|\mu\| = 1$. Since

$$\lambda' \partial \rho_F(t, \bar{x}(t), \dot{\bar{x}}(t)) \subset N_{\text{Gr}F(t, \cdot)}(\bar{x}(t), \dot{\bar{x}}(t)) \text{ and } \lambda' d_C(\bar{x}(0), \bar{x}(1)) \subset N_C(\bar{x}(0), \bar{x}(1)),$$

we have arrived at conditions (i)–(iv) of the theorem with $\lambda' = 0$.

The remaining alternative to consider is $\lambda' = 1$. We now show that this cannot arise and thereby complete the proof of the theorem. Indeed if $\lambda' = 1$, then $\|p\| + \|\mu\| = 0$. Recall however that, for each i , either $(x_i(0), x_i(1)) \notin C$ or $\dot{x}_i(t) \notin F(t, x_i(t))$ on a set of positive measure. We conclude from (C) and (D) that either

$$(25) \quad |(p_i(0), -[p_i(1) + \int_{[0,1]} \gamma_i(s)\mu_i(ds)])| \geq \lambda_i(1 - i^{-1})$$

or

$$(26) \quad |p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)| \geq \lambda_i(1 - i^{-1}) \text{ on a set of positive measure.}$$

If (25) is true for infinitely many i , then

$$|(p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)])| \geq 1.$$

This is impossible since $p = 0$ and $\mu = 0$.

On the other hand, if (26) is valid for infinitely many i , then

$$|p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)| \leq |p_i(t)| + \int_0^t |\gamma_i(s)\mu_i(ds)| \leq |p_i(t)| + k_h \|\mu_i\|, \quad \text{a.e.}$$

Hence

$$\|p_i\| \geq \max_{[0,1]} |p_i(t)| \geq \lambda_i(1 - i^{-1}) - k_h \|\mu_i\|.$$

Since $\|\mu_i\| \rightarrow 0$, we obtain in the limit $\|p\| \geq 1$. But this is not possible, since $\|p\| = 0$. Since either (25) or (26) must occur an infinite number of times, we have arrived at the desired contradiction. Q.E.D.

4. PROOF OF THEOREM 5

Let us first dispose of the concluding assertions of the theorem. Suppose that (p, μ, γ) satisfies (ii); then for a.e. $t \in [0, 1]$ there exists η such that

$$(27) \quad (\eta, q(t)) \in N_{\text{Gr}F(t, \cdot)}(\bar{x}(t), \dot{\bar{x}}(t)),$$

where $q(t) := p(t) + \int_{[0,t)} \gamma(s)\mu(ds)$. But we have assumed F is convex valued. A straightforward analysis of approximating proximal normals to $(\eta, q(t))$ permits us to deduce from (27) that

$$q(t) \in N_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t)) \quad \text{a.e.},$$

where the set on the right is the normal cone in the sense of convex analysis. This last relationship can be equivalently expressed as

$$q(t) \cdot \dot{\bar{x}}(t) \geq q(t) \cdot v \quad \text{for all } v \in F(t, \bar{x}(t)) \quad \text{a.e.}$$

This is condition (b). Condition (a) follows from a pointwise application of the partial dualization theorem [4], which tells us that, under the hypotheses currently in force,

$$\text{co}\{\xi : (\xi, q(t)) \in \partial\Psi_{\text{Gr}F(t,\cdot)}(\bar{x}(t), \dot{\bar{x}}(t))\} \subset \text{co}\{\eta : (-\eta, \dot{\bar{x}}(t)) \in H(t, \bar{x}(t), q(t))\},$$

and from the identity $\partial\Psi_{\text{Gr}F(t,\cdot)} = N_{\text{Gr}F(t,\cdot)}$.

Before proceeding, we prove a generalization of Theorem 3, which we recall provides necessary conditions for “finite Lagrangian” problems, to allow for a simple dynamic constraint which does not depend on the state variable. Consider then the optimization problem

$$(28) \quad \begin{aligned} &\text{Minimize } l(x(0), x(1)) + \int_0^1 L(t, x(t), \dot{x}(t))dt \\ &\text{over arcs } x \text{ which satisfy } \dot{x}(t) \in R(t) \text{ a.e., } h(t, x(t)) \leq 0 \text{ for all } t. \end{aligned}$$

Here $l : R^n \times R^n \rightarrow R$, $L : [0, 1] \times R^n \times R^n \rightarrow R$, and $h : [0, 1] \times R^n \rightarrow R$ are given functions and $R : [0, 1] \rightrightarrows R^n$ is a given multifunction.

Proposition 8. *Let \bar{x} be a $W^{1,1}$ local minimizer for (28). Assume*

- (G1) R has Borel measurable graph and $R(t)$ is closed for each t ,
- (G2) $L(\cdot, x, v)$ is Lebesgue measurable for each (x, v) , and there exist $\epsilon > 0$ and $k \in L^1$ such that

$$|L(t, x, v) - L(t, x', v')| \leq k(t)|x, v - (x', v')|$$

for all $(x, v), (x', v') \in (\bar{x}(t) + \epsilon B) \times (\dot{\bar{x}}(t) + \epsilon B)$,

- (G3) l is Lipschitz continuous on a neighbourhood of $(\bar{x}(0), \bar{x}(1))$, and
- (G4) h is upper semicontinuous near $(t, \bar{x}(t))$ for all t in $[0, 1]$, and there is a constant k_h such that $|h(t, x) - h(t, x')| \leq k_h|x - x'|$ for all t in $[0, 1]$ and all $x, x' \in \bar{x}(t) + \epsilon B$.

Then there exist an arc p , a constant $\lambda \geq 0$, a nonnegative measure μ and a μ -integrable function γ such that

- (i) $\|p\| + \|\mu\| + \lambda = 1$,
- (ii)' $\dot{p}(t) \in \text{co}\{\eta : (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \in \lambda\partial L(t, \bar{x}(t), \dot{\bar{x}}(t)) + \{0\} \times N_{R(t)}(\dot{\bar{x}}(t))\} \text{ a.e.},$
- (iii)' $(p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)]) \in \lambda\partial l(\bar{x}(0), \bar{x}(1))$,
- (iv) $\gamma(t) \in \partial_x^> h(t, \bar{x}(t)) \mu$ -a.e. and $\text{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}$.

Proof. Since \bar{x} is a $W^{1,1}$ local minimizer for (28), $(\bar{x}, \bar{z}(t) \equiv \int_0^t L(x, \bar{x}(s), \dot{\bar{x}}(s))ds)$ is a $W^{1,1}$ local minimizer for

$$\begin{aligned} & \text{Minimize } z(1) + l(x(0), x(1)) \\ & \text{subject to } (\dot{x}(t), \dot{z}(t)) \in \tilde{F}(t, x(t)) \quad \text{a.e.}, \\ & z(0) = 0, \\ & h(t, x(t)) \leq 0 \text{ for all } t, \end{aligned}$$

in which

$$\tilde{F}(t, x) := \{(v, \alpha) : v \in R(t) \text{ and } \alpha \geq L(t, x, v)\}.$$

The hypotheses of Theorem 4 are satisfied. We deduce the existence of an arc p , a vector $q \in R^n$, a constant $\lambda \geq 0$, a nonnegative measure μ and a μ -integrable function γ such that

$$\begin{aligned} & \|p\| + |q| + \|\mu\| + \lambda = 1, \\ & \dot{p}(t) \in \text{co}\{\eta : (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds), -q) \in N_{\text{Gr}\tilde{F}(t,\cdot)}(\bar{x}(t), \dot{\bar{x}}(t), L(t, \bar{x}(t), \dot{\bar{x}}(t)))\} \end{aligned}$$

almost everywhere,

$$\begin{aligned} & (p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)]) \in \lambda\partial l(\bar{x}(0), \bar{x}(1)) \text{ and } q = \lambda, \\ & \gamma(t) \in \partial_x^> h(t, \bar{x}(t)) \quad \mu\text{-a.e. and } \text{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}. \end{aligned}$$

But

$$\begin{aligned} \text{Gr}\tilde{F}(t, \cdot) &= \{(x, v, \alpha) : \alpha \geq L(t, x, v) + \Psi_{R(t)}(v)\} \\ &= \text{epi}\{(x, v) \rightarrow L(t, x, v) + \Psi_{R(t)}(v)\} \text{ for } t \in [0, 1]. \end{aligned}$$

It follows that “ $(w, p, -q) \in N_{\text{Gr}\tilde{F}(t,\cdot)}(\bar{x}(t), \dot{\bar{x}}(t), \dot{z}(t))$ ” implies $q \geq 0$ and

$$(w, p) \in q\partial(L(t, \bar{x}(t), \dot{\bar{x}}(t)) + \Psi_{R(t)}(\dot{\bar{x}}(t))) \subset q\partial L(t, \bar{x}(t), \dot{\bar{x}}(t)) + \{0\} \times N_{R(t)}(\dot{\bar{x}}(t)).$$

To justify the last inclusion we applied the “sum rule” [6], taking note of the Lipschitz continuity of $L(t, \cdot, \cdot)$ in a neighbourhood of $(\bar{x}(t), \dot{\bar{x}}(t))$. The differential inclusion for p may therefore be replaced by

$$\dot{p}(t) \in \text{co}\{\eta : (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \in \lambda\partial L(t, \bar{x}(t), \dot{\bar{x}}(t)) + \{0\} \times N_{R(t)}(\dot{\bar{x}}(t))\}.$$

Since $q = \lambda$ we have $\|p\| + \lambda + \|\mu\| \neq 0$. By scaling the multipliers we can arrange that $\|p\| + \lambda + \|\mu\| = 1$. The proposition is proved. \square

We may assume without loss of generality that the function k of hypothesis (G2)' (or (G2)'') satisfies $k(t) \geq 1$ a.e. Define

$$R(t) := \dot{\bar{x}}(t) + (\epsilon/3)B.$$

Reduce the size of ϵ if necessary to ensure that \bar{x} is minimizing with respect to arcs x satisfying $\|\bar{x} - x\|_{W^{1,1}} \leq \epsilon$ and that g is Lipschitz continuous on $(\bar{x}(0), \bar{x}(1)) + \epsilon(B \times B)$ with Lipschitz constant k_g . Define $\mathcal{S} \subset W^{1,1}$ and $\tilde{g} : W^{1,1} \rightarrow R$:

$$\begin{aligned} \mathcal{S} := \{x \in W^{1,1} : & \dot{x}(t) \in F(t, x(t)), \dot{x}(t) \in R(t) \quad \text{a.e.}, h(t, x(t)) \leq 0 \text{ for all } t, \\ & (x(0), x(1)) \in C, \|x - \bar{x}\|_{W^{1,1}} \leq \epsilon\} \end{aligned}$$

and

$$\tilde{g}(x) := g(x(0), x(1)).$$

Arguing as in the proof of Theorem 4, we can show that \bar{x} is a minimizer for the problem

$$\begin{aligned} & \text{Minimize } \tilde{g}(x) + 2k_g \inf\{\|x - y\|_{W^{1,1}} : y \in \mathcal{S}\} \\ & \text{subject to } \dot{x}(t) \in R(t) \text{ a.e., } h(t, x(t)) \leq 0 \text{ for all } t \in [0, 1], \quad \|x - \bar{x}\|_{W^{1,1}} \leq \epsilon. \end{aligned}$$

Again there are two possibilities:

(a) There exist $\epsilon' \in (0, \epsilon)$ and $K' > 0$, such that if $x \in W^{1,1}$ with $h(t, x(t)) \leq 0$ for all t and $\dot{x}(t) \in R(t)$ a.e. satisfies $\|x - \bar{x}\|_{W^{1,1}} \leq \epsilon'$, then

$$\inf\{\|x - y\|_{W^{1,1}} : y \in \mathcal{S}\} \leq K'[d_C(x(0), x(1)) + \int_0^1 \rho_F(t, x(t), \dot{x}(t))dt].$$

(b) There exist a sequence of arcs \bar{x}_i with $\dot{x}_i(t) \in R(t)$ a.e. and $h(t, \bar{x}_i(t)) \leq 0$ for all t such that $\bar{x}_i \rightarrow \bar{x}$ in $W^{1,1}$ and

$$\int_0^1 \rho_F(t, \bar{x}_i(t), \dot{\bar{x}}_i(t))dt + d_C(\bar{x}_i(0), \bar{x}_i(1)) < (2i)^{-1} \inf\{\|\bar{x}_i - y\|_{W^{1,1}} : y \in \mathcal{S}\}$$

for $i = 1, 2, \dots$

Consider case (a) first. Then \bar{x} is a $W^{1,1}$ local minimizer of the problem

$$\begin{aligned} & \text{Minimize } g(x(0), x(1)) + 2k_g K'[d_C(x(0), x(1)) + \int_0^1 \rho_F(t, x(t), \dot{x}(t))dt] \\ & \text{over arcs } x \text{ which satisfy } \dot{x}(t) \in R(t) \text{ a.e., } h(t, x(t)) \leq 0 \text{ for all } t. \end{aligned}$$

Now apply Lemma 6, in which for each $t \in [0, 1]$ we identify $\rho(x, v)$ with $\rho_F(t, x, v)$. This tells us that $\rho_F(t, \cdot, \cdot)$ is Lipschitz continuous with Lipschitz constant $2k(t)$ on $(\bar{x}(t), \dot{\bar{x}}(t)) + (\epsilon/3B) \times (\epsilon/3B)$ if $(G2)'$ holds and on $(\bar{x}(t), \dot{\bar{x}}(t)) + (\epsilon(3\bar{K})^{-1}B) \times (\epsilon/3B)$ if $(G2)''$ holds. (Set the parameters of Lemma 6 with $r = \epsilon k(t)$ and $K = k(t)$ if $(G2)'$ holds and $r = \epsilon$ and $K = \bar{K}$ if $(G2)''$ holds, and remember that $k(t) \geq 1$.) We check that the hypotheses under which Proposition 8 applies to this problem are met. Appealing to part (b) of Lemma 6, we deduce the existence of $p, \lambda \geq 0, \mu$ and γ satisfying conditions (i)–(iii) of the theorem. (Notice that $R(t)$ does not appear in these conditions because the constraint “ $\dot{x}(t) \in R(t)$ ” is inactive.) This deals with case (a).

Now consider case (b). Using arguments similar to those employed in the proof of Theorem 4, we show by means of Ekeland’s theorem that there exists $x_i \rightarrow \bar{x}$ strongly in $W^{1,1}$ such that, for each i , x_i is a $W^{1,1}$ local minimizer for

$$\begin{aligned} & \text{Minimize } \int_0^1 \rho_F(t, x(t), \dot{x}(t))dt + d_C(x(0), x(1)) \\ & \quad + i^{-1}(|x(0) - x_i(0)| + \int_0^1 |\dot{x}(t) - \dot{x}_i(t)|dt) \\ & \text{over arcs } x \text{ which satisfy } \dot{x}(t) \in R(t) \text{ a.e., } h(t, x(t)) \leq 0 \text{ for all } t. \end{aligned}$$

Furthermore, $x_i \notin \mathcal{S}$. This last property implies that

$$(29) \quad \begin{aligned} & (x_i(0), x_i(1)) \notin C \text{ or } \dot{x}_i(t) \notin F(t, x_i(t)) \\ & \text{for all } t\text{'s in a set of positive measure.} \end{aligned}$$

Now apply Proposition 8 to the preceding problem, with reference to the $W^{1,1}$ local minimizer x_i , for $i = 1, 2, \dots$. We conclude that there exist multipliers (p_i, λ_i, μ_i) and a Borel measurable function γ_i such that

- (A) $\lambda_i + \|p_i\| + \|\mu_i\| = 1$,
- (B) $\dot{p}_i(t) \in \text{co}\{\eta : (\eta, p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)) \in \lambda_i \partial \rho_F(t, x_i(t), \dot{x}_i(t)) + \{0\} \times (\lambda_i i^{-1}B) + \{0\} \times N_{R(t)}(\dot{x}_i(t))\}$ a.e.,
- (C) $(p_i(0), -[p_i(1) + \int_{[0,1]} \gamma_i(s)\mu_i(ds)]) \in \lambda_i \partial d_C(x_i(0), x_i(1)) + \{\lambda_i i^{-1}\} \times \{0\}$,
- (D) $\gamma_i(t) \in \partial_x^> h(t, x_i(t))$ μ_i -a.e., and $\text{supp}\{\gamma_i\} \subset \{t : h(t, x_i(t)) = 0\}$.

We deduce from Lemma 6 and (B) that $|\dot{p}_i(t)| \leq k(t)$ a.e. for $i = 1, 2, \dots$. It follows that the \dot{p}_i 's are uniformly integrably bounded. We may therefore arrange, by subsequence extraction, that $p_i \rightarrow p$ uniformly and $\dot{p}_i' \rightarrow p'$ weakly in L^1 for some $p \in W^{1,1}$. We can ensure by again extracting a subsequence that $\lambda_i \rightarrow \lambda'$ for some $\lambda' \geq 0$, $\mu_i \rightarrow \mu$ weakly* and $\gamma_i d\mu_i \rightarrow \gamma d\mu$ weakly* for some constant $\lambda' \geq 0$, nonnegative measure μ and μ -integrable function γ . The convergence analysis in the proof of Theorem 4 may now be reproduced to justify passing to the limit in the relationships (A)-(D). We thereby obtain

$$\begin{aligned} & \|p\| + \lambda' + \|\mu\| = 1, \\ \dot{p}(t) & \in \text{co}\{w : (w, p(t) + \int_{[0,t)} \gamma(s)\mu(ds)) \in \lambda' \partial \rho_F(t, \bar{x}(t), \dot{\bar{x}}(t))\} \quad \text{a.e.}, \\ & (p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)]) \in \lambda' \partial d_C(\bar{x}(0), \bar{x}(1)), \\ & \gamma(t) \in \partial_x^> h(t, \bar{x}(t)) \text{ and } \text{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}. \end{aligned}$$

Suppose first that $0 \leq \lambda' < 1$. In this case $\|p\| + \|\mu\| > 0$. Scale the multipliers so that $\|p\| + \|\mu\| = 1$. Recalling that $\partial \rho_F(t, \bar{x}(t), \dot{\bar{x}}(t)) \subset N_{\text{Gr}F(t, \cdot)}(\bar{x}(t), \dot{\bar{x}}(t))$ and $\partial d_C(\bar{x}(0), \bar{x}(1)) \subset N_C(\bar{x}(0), \bar{x}(1))$, we see that p, μ and γ satisfy the conditions (i)–(iv) of Theorem 5 with $\lambda = 0$.

The remaining case to consider is when $\lambda' = 1$. We complete the proof by showing that this cannot arise.

The following lemma will be required:

Lemma 9. *Take any $t \in [0, 1]$ and any index value $i \geq 4$. Suppose that $(x, v) \in (\bar{x}(t) + \epsilon(12k(t))^{-1}B) \times (\dot{\bar{x}}(t) + \epsilon/3B)$ and $(w, q) \in R^n \times R^n$ are such that*

$$(w, q) \in \lambda' \partial \rho_F(t, x, v) + \{0\} \times (\lambda' i^{-1}B) + \{0\} \times N_{R(t)}(v).$$

Then $|w| \leq k(t)$. Furthermore, if $v \notin F(t, x)$ then

$$|q| \geq \lambda' \sqrt{5/16}.$$

Proof. (w, q) can be expressed as

$$(30) \quad (w, q) = \lambda'(a, b) + \lambda'(0, c) + (0, \alpha(v - \dot{\bar{x}}(t))/|v - \dot{\bar{x}}(t)|)$$

(the last term is interpreted as $(0,0)$ if $v = \dot{\bar{x}}(t)$) with $(a, b) \in \partial \rho_F(t, x, v)$ and $|c| \leq 1/4$. α is a non-negative member which is zero if $v \notin \text{boundary}\{R(t)\}$.

We see that $w = \lambda'a$. Lemma 6 tells us that $|w| \leq k(t)$. Now suppose that $v \notin F(t, x)$. According to Lemma 6,

$$b = (v - u)/|v - u|,$$

where u is the closest point to v in the convex set $F(t, x)$. If $v \notin \text{boundary}\{R(t)\}$ then the parameter α in (30) takes value zero, and we deduce that $q = \lambda'(b + c)$. In this case $|q| \geq \lambda'(|b| - |c|) = \lambda'3/4 \geq \lambda'\sqrt{5/16}$. So we may assume that $v \in \text{boundary}\{R(t)\}$, i.e.

$$|v - \dot{x}(t)| = \epsilon/3.$$

By hypothesis (G2)' (or (G2)'') there exists $v' \in F(t, x)$ such that

$$|\dot{x}(t) - v'| \leq \epsilon/12.$$

$v - u$ is a normal vector to the convex set $F(t, x)$ at u . Since $v' \in F(t, x)$, we have

$$(u - v') \cdot (v - u) \geq 0.$$

We deduce (writing $\bar{v} = \dot{x}(t)$) that

$$\begin{aligned} 0 \leq (u - v') \cdot (v - u) &= (u - \bar{v}) \cdot (v - u) + (\bar{v} - v') \cdot (v - u) \\ &= (v - \bar{v}) \cdot (v - u) + (\bar{v} - v') \cdot (v - u) - |v - u|^2 \\ &\leq (v - \bar{v}) \cdot (v - u) + \epsilon|v - u|/12 - |v - u|^2 \end{aligned}$$

Since $v \notin F(t, x)$ and $u \in F(t, x)$, it follows that $u \neq v$, and we conclude that

$$\frac{(v - u)}{|v - u|} \cdot (v - \bar{v}) \geq |v - u| - \epsilon/12 \geq -\epsilon/12.$$

However, for any $\alpha \geq 0$

$$\begin{aligned} & \left| \frac{(v - u)}{|v - u|} + c + \alpha(v - \bar{v}) \right|^2 \\ & \geq \left| \frac{(v - u)}{|v - u|} + c \right|^2 + 2\alpha \left(\frac{(v - u)}{|v - u|} \cdot (v - \bar{v}) + c \cdot (v - \bar{v}) \right) + \alpha^2 |v - \bar{v}|^2 \\ & \geq |1 - 1/4|^2 + 2\alpha(-\epsilon/12 - \epsilon/12) + (\alpha\epsilon/3)^2 \\ & = (3/4)^2 + (\alpha\epsilon/3 - 1/2)^2 - 1/4 \\ & \geq 5/16. \end{aligned}$$

It follows that

$$|q| = \lambda' \left| \frac{(v - u)}{|v - u|} + c + \alpha(v - \bar{v}) \right| \geq \lambda' \sqrt{5/16}$$

The lemma is proved. □

Recall that, by assumption, $\lambda' = 1$. It follows that $p = 0, \mu = 0$. Now in view of (29) we can arrange by subsequence extraction that either

- (α) $(x_i(0), x_i(1)) \notin C$ for all i ; or
- (β) $\dot{x}_i(t) \notin F(t, x_i(t))$ for all t 's in a set of positive measure.

Suppose first that (α) is true. We deduce from (C) and the properties of the distance function that

$$|(p_i(0), -[p_i(1) + \int_{[0,1]} \gamma_i(s)\mu_i(ds)])| \geq \lambda_i(1 - i^{-1}).$$

Passing to the limit, we obtain $|(p(0), -[p(1) + \int_{[0,1]} \gamma(s)\mu(ds)])| \geq 1$. This contradicts $p = 0$ and $\mu = 0$.

Suppose finally that (β) is true. According to Lemma 9,

$$|p_i(t) + \int_{[0,t)} \gamma_i(s)\mu_i(ds)| \geq \lambda_i \sqrt{5/16}$$

on a set of positive measure. Now in view of the fact that $|\gamma_i(t)| \leq k_h$ μ -a.e., we deduce that

$$\|p_i\|_\infty + k_h \|\mu_i\| \geq \lambda_i \sqrt{5/16}.$$

In the limit we obtain

$$\|p\|_\infty + k_h \|\mu\| \geq \sqrt{5/16}.$$

Once again we have arrived at a contradiction to the fact that $p = 0$ and $\mu = 0$. The theorem is proved. Q.E.D.

REFERENCES

1. F. H. Clarke, *Optimal Solutions to Differential Inclusions*, J. Optim. Theory Appl. **19** (1976), 469–478. MR **54**:13367
2. F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983. MR **85m**:49002
3. F. H. Clarke, “*Methods of Dynamic and Nonsmooth Optimization*”, (CBMS-NSF Regional conference series in applied mathematics, vol 57) SIAM Publication, Philadelphia, 1989. MR **91j**:49001
4. A. D. Ioffe, *Euler-Lagrange and Hamiltonian Formalisms in Dynamic Optimization*, Trans. Amer. Math. Soc. **349** (1997), 2871–2900. MR **97i**:49028
5. A. D. Ioffe and R. T. Rockafellar, *The Euler and Weierstrass Conditions for Nonsmooth Variational Problems*, Calc. Var. Partial Differential Equations **4** (1996), 59–87. MR **96k**:49031
6. P. D. Loewen, “*Optimal Control Via Nonsmooth Analysis*”, CRM Proceedings and Lecture Notes, American Mathematical Society, Providence, 1993. MR **94h**:49003
7. P. D. Loewen and R. T. Rockafellar, *Optimal Control of Unbounded Differential Inclusions*, SIAM J. Control Optim. **32** (1994), 442–470. MR **95h**:49043
8. P. D. Loewen and R. T. Rockafellar, *New Necessary Conditions for the Generalized Problem of Bolza*, SIAM J. Control. Optim. **34** (1996), 1496–1511. MR **97d**:49021
9. B. S. Mordukhovich, *Maximum Principle in Problems of Time Optimal Control with Nonsmooth Constraints*, J. Appl. Math. Mech. **40** (1976), 960–969. MR **58**:7284
10. B. S. Mordukhovich, *Generalized Differential Calculus for Nonsmooth and Set-Valued Mappings*, J. Math. Anal. Appl. **183** (1994), 250–288. MR **95i**:49029
11. B. S. Mordukhovich, *Discrete Approximations and Refined Euler-Lagrange Conditions for Nonconvex Differential Inclusions*, SIAM J. Control Optim. **33** (1995), 882–915. MR **96d**:49028
12. R. T. Rockafellar, *Equivalent Subgradient Versions of Hamiltonian and Euler Lagrange Equations in Variational Analysis*, SIAM J. Control. Optim. **34** (1996), 1300–1314. CMP 96:14
13. R. B. Vinter and G. Pappas, *A Maximum Principle for Nonsmooth Optimal Control Problems with State Constraints*, J. Math. Anal. Appl. **89** (1982), 212–232. MR **84b**:49026
14. R. B. Vinter and H. Zheng, *The Extended Euler-Lagrange Condition for Nonconvex Variational Problems*, SIAM J. Control Optim. **35** (1997), 56–77. CMP 97:07

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING AND CENTRE FOR PROCESS SYSTEMS ENGINEERING, IMPERIAL COLLEGE, EXHIBITION ROAD, LONDON SW7 2BT, UK
E-mail address: r.vinter@ic.ac.uk

DEPARTMENT OF BUSINESS STUDIES, UNIVERSITY OF EDINBURGH, 50 GEORGE SQUARE, EDINBURGH EH9 9JY, SCOTLAND, UK
E-mail address: zheng@maths.ed.ac.uk