AN INDIRECT METHOD IN THE CALCULUS OF VARIATIONS

F. H. CLARKE

ABSTRACT. This article presents a new approach to the issue of the existence of solutions to the basic problem in the calculus of variations. The method is indirect, and applies to certain classes of of problems with slow or no growth, in addition to those satisfying the traditional coercivity condition. The proof hinges upon showing with the help of nonsmooth analysis that a certain value function is constant. Examples are given to illustrate the applicability of the results and the necessity of the hypotheses.

1. INTRODUCTION

Consider the following basic problem in the calculus of variations: to minimize the integral functional

$$\Lambda(x) := \int_0^T L(x(t), \dot{x}(t)) \, dt$$

over a class X of functions $x(\cdot)$ mapping $[0, T]$ to $\mathbb{R}^n$, subject to prescribed boundary conditions:

$$x(0) = x_0, \qquad x(T) = x_T$$

and a unilateral constraint on the state x:

$$x(t) \in \Omega \quad \text{for all } t.$$

As with any optimization problem, a central question is that of existence: under what conditions on the data can we assert that there is a solution x to the problem? This is more than an academic point, for the principal way to solve the problem is to apply the deductive method, the chain of reasoning that goes as follows: a solution exists; any solution must satisfy such and such necessary conditions; perusal of these conditions leads to an answer. The first link in this chain is crucial, and its nonavailability for two centuries had a marked effect on the subject. For example, much of the impetus for studying differential equations qualitatively originally arose from the need to develop the theory of fields of extremals, the main ingredient in the classical sufficient conditions, which are able to assert only the presence of a local minimum.

A significant turning point in the calculus of variations is due to Leonida Tonelli who, in a seminal paper [8] entitled "Sur une méthode directe du calcul des variations", introduced what has become known as *the* direct method. Tonelli's contribution to his subject was manifold; the pertinent ones here are his introduction for X of the class of absolutely continuous functions as the natural space in which to work, and the identification of a set of conditions on the Lagrangian L of the problem assuring that the functional Λ be lower semicontinuous in an appropriate topology. A somewhat modernized version of Tonelli's theorem for our problem is the following, in which, now and forever, the class X of competing functions is taken to be the absolutely continuous functions from $[0, T]$ to $\mathbb{R}^n$, a class labeled AC, and whose elements are called *arcs*.

Theorem 1. *Suppose that* $L \colon \mathbb{R}^{2n} \to \mathbb{R}$ *is continuous, convex in the* $\dot{x}$ *variable, and coercive. Then if* Ω *is closed, the problem admits a minimizing arc.*

The word "coercive" refers to the existence of a function $\theta \colon [0, \infty) \to \mathbb{R}$ satisfying:

    (i) $L(x, v) \geq \theta(|v|)$ for all x in Ω and for all v.

    (ii) $\lim_{r \to \infty} \frac{\theta(r)}{r} = \infty$.

There is no loss of generality in assuming θ to be continuously differentiable, convex, and increasing; we shall refer to such a function θ as a *Nagumo function*. The proof of the theorem, incidentally, lives up to the word "direct": one extracts from *any* minimizing sequence a convergent subsequence and shows that the limit is a solution. Essentially what is involved is a compactness property of the level sets of the functional Λ; i.e. the sets $\{x \in \text{AC} \colon \Lambda(x) \leq \Lambda(\overline{x})\}$.

The recent book [3] of Lamberto Cesari (a student of Tonelli) gives a thorough account of existence theory for the basic problem and some of its extensions. The direct method is so robust that it has become a standard tool in variational problems of all kinds. Indeed, in most cases, subsequent results can be viewed as variations on Tonelli's theme. A notable exception is a technique pioneered by MacShane wherein one reframes a certain class of problems in order to apply the compactness theory of parametric integrals (see [3]).

In this article we present a new approach to existence. It proceeds in an indirect way which invokes necessary conditions at an intermediate point. This novel feature is shared by another approach of recent vintage due to Clarke and Loewen [6], which is otherwise different in both approach and realm of applicability.

Rather than employ hypotheses that imply the compactness of the level sets of Λ, we require that the structure of the problem imply a much weaker property. The property in question is that for some $k > 0$, every admissible arc x satisfy

$$\operatorname*{ess\,inf}_{0 \leq t \leq T} |\dot{x}(t)| < k.$$

The coercivity hypothesis certainly implies such a condition, but we are also able to treat certain classes of noncoercive problems (so-called "slow-growth" problems, and even some that have no growth whatever).

An important element in the proof is the use of the second Erdmann condition at an intermediate stage. We owe to [1] the observation that this classical necessary condition holds in a wider context than had been realized.

The main idea of the proof is to study a certain (nondifferentiable) value function via nonsmooth analysis. The key is to prove this function eventually constant. This is the first application of the fruitful technique known as "proximal analysis" (see [5]) to existence theory; we believe that others will follow.

The next two sections state and prove the existence theorem in the autonomous case, in which technicalities are minimized and the nature of the proof is clearest. Also, a number of examples are given. Section 4 carries out the extension to the nonautonomous case, while §5 is devoted to a refinement that can be made when $n = 1$ (see also [2] for an interesting approach to one-dimensional problems via relaxation). Sections 6 and 7 obtain results in nonsmooth analysis and optimization that are needed for the proof.

2. The main theorem; examples

2.1. **Basic hypotheses.** We are given a closed subset Ω of $\mathbb{R}^n$. The Lagrangian L of the problem is a lower semicontinuous function mapping $\Omega \times \mathbb{R}^n$ to $(-\infty, \infty]$. For each x in Ω, the domain of the function $v \to L(x, v)$ (i.e., the set on which L is finite-valued) is assumed to be a nonempty open convex set upon which the function is convex. We hypothesize as well that $L(x, v)$ is bounded below on subsets of (x, v) having bounded x components. These conditions assure that the integral functional $\Lambda(x)$ is well defined (possibly as $+\infty$) whenever x is an arc. The other elements of the problem are a closed subset C of $\mathbb{R}^n \times \mathbb{R}^n$ and a closed convex cone K in $\mathbb{R}^n$. The problem (P) we are studying, an extension of the one mentioned in the introduction, consists of minimizing $\Lambda(x)$ subject to the arcs satisfying the constraints

$$(x(0), x(T)) \in C, \qquad x(t) \in \Omega \quad \forall t, \quad \dot{x}(t) \in K \quad \text{a.e.}$$

Any such arc is termed *admissible*. We suppose that at least one of the following sets is compact:

$$\{x \in \mathbb{R}^n : \text{for some } y \in \mathbb{R}^n, (x, y) \in C\},$$
$$\{y \in \mathbb{R}^n : \text{for some } x \in \mathbb{R}^n, (x, y) \in C\},$$

and that at least one Lipschitz admissible arc $\overline{x}$ exists for which $\Lambda(\overline{x})$ is finite. We shall denote by $\Gamma(\overline{x})$ the following level set of admissible arcs:

$$\Gamma(\overline{x}) := \{x \in \text{AC} : (x(0), x(T)) \in C, x(t) \in \Omega \, \forall t, \dot{x}(t) \in K \text{ a.e.}, \Lambda(x) \leq \Lambda(\overline{x})\}.$$

2.2. **The theorem.** The first of our two principal hypotheses concerns the level set $\Gamma(\overline{x})$:

(H1) $\left\{ \begin{array}{l} \text{There exists } k > 0 \text{ such that any member } x \text{ of } \Gamma(\overline{x}) \text{ satisfies} \\ \qquad \operatorname*{ess\,inf}_{0 \leq t \leq T} |\dot{x}(t)| < k. \end{array} \right.$

We shall discuss presently a variety of ways in which such a condition can be derived from verifiable properties of the data of (P).

(H1) by itself does not imply the existence of a solution to (P) (as shown by Example 2.4.1 below). The supplementary condition that fills the gap is in

the form of an inequality:

$$(*) \quad \begin{aligned} &\lim_{\beta \to \infty} \sup_{\substack{x \in \Omega, \, v \in K \\ |v| > \beta}} \{L(x, v) - v \cdot \partial_v L(x, v)\} \\ &\qquad < \inf_{\substack{x \in \Omega, \, v \in K \\ |v| < k}} \{L(x, v) - v \cdot \partial_v L(x, v)\}. \end{aligned}$$

In this expression, $\partial_v L$ refers to the subgradient of L in the v variable; this is a nonempty compact set for all v in the domain of $L(x, \cdot)$, since this function is finite and convex, and hence Lipschitz, in that open set. The supremum and infimum in $(*)$ are taken over the elements in the set specified in the braces. When L is differentiable in v, the set $\partial_v L(x, v)$ reduces to the singleton set whose element is $\nabla_v L(x, v)$. Our missing hypothesis is the following:

(H2) $(*)$ holds for the k provided by (H1).

Theorem 2. *Let the basic hypotheses hold, together with* (H1) *and* (H2). *Then* (P) *admits a solution. Furthermore, any solution* x *is Lipschitz and satisfies*

$$L(x(t), \dot{x}(t)) - \dot{x}(t) \cdot p(t) = c \quad a.e.,$$

where c *is a constant, and where* p *is a measurable function such that* $p(t)$ *belongs to* $\partial_v L(x(t), \dot{x}(t))$ *a.e.*

We remark that the Lipschitz regularity of the solution is a result of independent interest, allowing as it does the writing of necessary conditions of a strong type; we refer the reader to the discussions in [5, Chapter 2].

The proof of the theorem inaugurates the use of an indirect method based upon proximal analysis; it is given in §3. In the remainder of this section, we illustrate various contexts in which the theorem applies.

2.3. **The coercive case.** We shall show that Theorem 2 subsumes Theorem 1. Suppose now that $K = \mathbb{R}^n$, and that L satisfies the hypotheses of Theorem 1. The coercivity of L certainly implies that $L(x, v)$ majorizes a function of the form $\varepsilon |v| + \gamma$, for a positive constant ε. It follows that any arc in $\Gamma(\overline{x})$ satisfies

$$\int_0^T \{\varepsilon |\dot{x}(t)| + \gamma\} \, dt \le \Lambda(\overline{x}),$$

which (since $x(0)$ or $x(T)$ is bounded a priori) gives an a priori bound on $\|x\|_\infty$. It is therefore no loss of generality to assume Ω compact, which we shall do. To see that Theorem 2 applies, we must verify the presence of hypotheses (H1) and (H2). The preceding remarks show that (H1) is satisfied for any k exceeding $[\Lambda(\overline{x}) - \gamma T]/(\varepsilon T)$. As for (H2), we have

Lemma 2.3.1. *The inequality* $(*)$ *holds for any* $k > 0$.

Proof. We shall proceed by showing first that the right side of $(*)$ is finite, and then that the left side is $-\infty$. Suppose that the right side is unbounded below. Then a sequence (x_i, v_i) in $\Omega \times k\overline{B}$ exists such that, for elements p_i in $\partial_v L(x_i, v_i)$, one has

$$L(x_i, v_i) - p_i \cdot v_i \to -\infty.$$

By passing to a subsequence if necessary, we can suppose that (x_i, v_i) converges to a limit (x, v) in $\Omega \times k\overline{B}$. Since L is finite for x in Ω, and continuous, we have

$$-\infty < L(x, v) = \lim L(x_i, v_i) < \infty.$$

It follows that $p_i \cdot v_i$ must tend to $+\infty$. Let $\varepsilon_0 > 0$ be such that $L(x, v + \varepsilon_0 v)$ is finite. Now for each i the subgradient inequality gives

$$L(x_i, v_i + \varepsilon_0 v_i) - L(x_i, v_i) \geq \varepsilon_0 p_i \cdot v_i.$$

Taking the limit, we deduce that $L(x, v + \varepsilon_0 v)$ is $+\infty$, a contradiction that establishes the first half of the proof.

For any (x, v) in $\Omega \times \mathbb{R}^n$, for any p in $\partial_v L(x, v)$, and for any ε in $(0, |v|)$, purely formal manipulation of the subgradient inequality (see [1, Lemma 4.6] for details) gives rise to

$$L(x, v) - p \cdot v \quad \leq \quad \frac{|v|}{|v| - \varepsilon} L\left(x, \frac{\varepsilon v}{|v|}\right) - \frac{\varepsilon}{|v| - \varepsilon} \theta(|v|).$$

Because L is bounded on bounded subsets of $\Omega \times \mathbb{R}^n$, the growth property of θ readily implies, in view of this inequality, that the left side of $(*)$ is $-\infty$ as claimed. $\square$

We have shown that Theorem 1 may be viewed as a special case of Theorem 2. The former's hypotheses actually imply, as is well known, that level sets such as $\Gamma(\overline{x})$ are compact in a suitable sense. In contrast to this, Theorem 2 applies in situations in which $\Gamma(\overline{x})$ even fails to be closed (see Example 2.4.3 below).

2.4. Linear growth.

We remarked in discussing the coercive case that the *linear growth condition*

$$L(x, v) \geq \varepsilon |v| + \gamma \qquad (x \in \Omega, v \in K)$$

is adequate to imply (H1) for certain values of k. In itself, however, and in the absence of (H2), the existence of a solution to (P) does not follow, as we now demonstrate:

Example 2.4.1. We set $n = 1$, $T = 1$, $C = \{(0, 1)\}$, $\Omega = \mathbb{R}$, $K = \mathbb{R}$, and we take L to be $x^2 + f(v)$, where f is the smooth convex function defined by

$$f(v) := v[1 + \min(v, 0)].$$

Then L satisfies a linear growth condition with $\varepsilon = 1$, $\gamma = -1$. Any admissible arc x satisfies

$$\Lambda(x) = \int_0^1 (x^2 + f(\dot{x})) \, dt > \int_0^1 f(\dot{x}) \, dt$$

$$\geq \int_0^1 \dot{x} \, dt \quad (\text{since } f(v) \geq v \;\; \forall v)$$

$$= 1.$$

So the infimum in (P) is no less than 1, a value which $\Lambda(x)$ can never attain. The infimum is in fact equal to 1, as evidenced by the sequence of admissible arcs x_i which equal 0 on $[0, 1 - 1/i]$, and which are linear between $1 - 1/i$ and 1. We have $\Lambda(x_i) \to 1$, showing that (P) fails to admit a solution. $\square$

Here is a simple first example not covered by previous existence theories, to which Theorem 2 applies.

Example 2.4.2. We set

$$L(x, v) := g(x)\sqrt{1 + |v|^2},$$

where g is any locally bounded lower semicontinuous function on $\mathbb{R}^n$ with $\inf\{g(x): x \in \Omega\} > 0$. We claim that for any choice of the data C, Ω, K satisfying the basic hypotheses, (P) has a solution (and all solutions are Lipschitz). To see this, note that because of the linear growth possessed by L, any element x of $\Gamma(\overline{x})$ admits an a priori bound for $\|x\|_\infty$. Consequently we may suppose that Ω is compact without affecting the issue at hand. Now, the linear growth furnishes (H1) for a suitable choice of k, so to invoke Theorem 2 it suffices to confirm that $(*)$ holds for every k. We calculate

$$L(x, v) - v \cdot L_v(x, v) = g(x)/\sqrt{1 + |v|^2},$$

so if g_1 signifies an upper bound for g on Ω, the left side of $(*)$ is bounded above by

$$\lim_{\beta \to \infty} g_1/\sqrt{1 + \beta^2} = 0.$$

The right side of $(*)$ is bounded below by $\delta/\sqrt{1 + k^2}$, where δ is a positive lower bound for g, confirming that $(*)$ holds. □

In the example above, $(*)$ is satisfied for any $k > 0$. We now examine a situation in which this is not the case; i.e., in which $(*)$ holds only for certain values of k. We remain in the presence of linear growth.

Example 2.4.3. We set $K = \Omega = \mathbb{R}^n$, $C = \{(0, 0)\}$, and $T = 1$, so that the boundary conditions are $x(0) = x(1) = 0$. The functional Λ is given by

$$\Lambda(x) := \int_0^1 \{\sqrt{1 + |\dot{x}(t)|^2} - r \sin |x(t)|\} \, dt,$$

where r is a positive constant. The inequality $(*)$ reduces to the condition

$$2r\sqrt{1 + k^2} < 1,$$

so in order to apply Theorem 2, we must show that (H1) holds for such a value of k.

The linear growth condition is satisfied with $\varepsilon = 1$, $\gamma = -r$, and we may take $\overline{x} \equiv 0$, so that (H1) certainly holds for any k exceeding $\Lambda(\overline{x}) - \gamma = 1 + r$. We deduce therefore that (P) admits a solution whenever

$$2r\sqrt{1 + (1 + r)^2} < 1,$$

i.e., when r is sufficiently small.

We claim that a solution fails to exist when r is sufficiently large. For when a solution x does exist, the Erdmann necessary condition holds (see §7):

$$-r \sin |x(t)| + \frac{1}{\sqrt{1 + |\dot{x}(t)|^2}} = c.$$

Since $x(0) = 0$, c is positive, and it follows that $-r \sin |x(t)|$ exceeds -1 for all t. This implies that $\Lambda(x)$ is positive. On the other hand, when $r > 1 + \pi$ we may define a sequence of admissible arcs x_i having $\lim \Lambda(x_i) < 0$, as follows:

$x_i(t) = y$ for $t \in (i^{-1}, 1 - i^{-1})$, where y is a constant vector of norm $\pi/2$, and the graph of x_i is linear between the pairs of points $(0, 0)$ and (i^{-1}, y), and $(1 - i^{-1}, y)$ and $(1, 0)$. Then one readily sees $\lim \Lambda(x_i) = 1 + \pi - r < 0$, showing that no solution x with $\Lambda(x) > 0$ can exist. Note incidentally that for any value of r, Λ is bounded on a sequence x_i such as the above, and that no subsequence of x_i converges to an arc. This shows that even when Theorem 2 applies, the sets $\Gamma(\overline{x})$ need not enjoy the compactness properties that they possess in the coercive case.

The question remains of whether a potentially better (i.e., lower) estimate can be made for k. Recall that k is simply an upper bound for $\inf\{|\dot{x}(t)| : 0 \leq t \leq T\}$, where x is an arbitrary element of $\Gamma(\overline{x})$, and that our choice of k resulted from the linear growth condition. We may improve upon k (and hence deduce existence for a larger set of r) by improving upon $\overline{x}$; i.e., finding another Lipschitz admissible arc giving a lower value to Λ. But the sharpest result of all would be obtained by taking k arbitrarily close to 0, and this turns out to be possible when $n = 1$, by an alternate approach peculiar to the one-dimensional case (see §5 for a continuation of the discussion). □

2.5. Examples with no growth.

In our examples up to now, linear growth of the Lagrangian has been the means of providing hypothesis (H1). We now treat some cases in which L exhibits no such growth, and in which (H1) is implied by the structure of the cone and the endpoint constraints.

Example 2.5.1. We set $n = 1$ and consider the minimization of

$$\int_0^1 \{e^{-\dot{x}(t)} + rg(x(t))\} \, dt$$

subject to $\dot{x}(t) \geq 0$ a.e., $x(0) = 0$, $x(1) = \delta > 0$. Thus $K = [0, \infty)$ and $C = \{(0, \delta)\}$. The function g is assumed lower semicontinuous and locally bounded, and r is nonnegative. Note that any admissible arc x necessarily takes values in $[0, \delta]$, so we may take $\Omega = [0, \delta]$. In addition, for any k greater than δ, an admissible arc x must have $|\dot{x}(t)| < k$ for t in a set of positive measure. This verifies (H1) for any such k. Let us examine (H2).

We calculate

$$L(x, v) - vL_v(x, v) = rg(x) + e^{-v}[1 + v].$$

The supremum of this expression for $x \in \Omega$, $v \in K$ $(= [0, \infty))$, and $|v| > \beta$ is $rg_1 + e^{-\beta}[1 + \beta]$, where g_1 is the supremum of g on $[0, \delta]$. The infimum for $x \in \Omega$, $v \in K$, $|v| < k$ is $rg_0 + e^{-k}[1 + k]$, where g_0 is the minimum of g over $[0, \delta]$. Therefore, if $(*)$ is to hold for any $k > \delta$, we must have

$$rg_1 < rg_0 + e^{-\delta}[1 + \delta],$$

which holds for r sufficiently small. (Alternatively, for certain r, this will hold for δ sufficiently small.) For such r, Theorem 2 applies and a Lipschitz solution exists. See [2] for a discussion in different terms of an example similar to this one.

It is possible to show that if the problem is modified by deleting the constraint $x(1) = \delta$, then no solution exists for any $r > 0$; of course, (H1) fails in that case. In contrast, the preceding examples are unaffected by this type of modification. □

We look next at an example in which the domain of $L(x, \cdot)$ is not all of $\mathbb{R}^n$.

Example 2.5.2. We set $n = 1$, $T = 1$, and consider minimizing

$$\int_0^1 \{|\dot{x}|^{-2} - |x|^2\}\, dt$$

subject to

$$x(0) = 0, \qquad x(1) = \delta > 0, \qquad \dot{x}(t) \geq 0 \text{ a.e.}$$

Note that the domain of $L(x, \cdot)$ is $(0, \infty)$ for each x. We may add the constraint $x(t) \in \Omega$, where $\Omega = [0, \delta]$, without changing the nature of the problem.

It is clear as in the preceding example that (H1) is satisfied for any k greater than δ. We turn now to (H2). We calculate

$$L(x, v) - v \cdot L_v(x, v) = 3v^{-2} - |x|^2.$$

The supremum over $x \in \Omega$, $v \in K$, $|v| > \beta$ is equal to $3\beta^{-2}$, and the infimum over $x \in \Omega$, $v \in K$, $|v| < k$ is $3k^{-2} - \delta^2$. In order to have (∗) for some $k > \delta$, we therefore require $\delta^2 < 3\delta^{-2}$, which is valid for δ sufficiently small. □

We look now at a no growth situation in two dimensions.

Example 2.5.3. We set $n = 2$, $T = 1$, and consider minimizing

$$\int_0^1 \{e^{-\dot{x}_1(t)} + (\dot{x}_2(t))^2 + \dot{x}_1(t)g(x(t))\}\, dt$$

subject to

$$\dot{x}_i(t) \geq 0 \text{ a.e. } (i = 1, 2), \qquad x(0) = [0, 0], \qquad x(1) = [\delta_1, \delta_2],$$

where $\delta_i > 0$ $(i = 1, 2)$ and g is continuous. Let us show that (H1) is satisfied for any k greater than $\delta_1 + \delta_2$. Let x be an admissible arc. Since $[1, 1] \cdot v \geq |v|$ for any v in the first quadrant, we have

$$\int_0^1 |\dot{x}(t)|\, dt \leq \int_0^1 [1, 1] \cdot \dot{x}(t)\, dt = [1, 1] \cdot [\delta_1, \delta_2] = \delta_1 + \delta_2,$$

which implies (H1).

We turn now to (H2). We calculate

$$L(x, v) - v \cdot L_v(x, v) = e^{-v_1}[1 + v_1] - v_2^2.$$

The supremum of this expression over $v \in K$, $|v| > \beta$ is $e^{-\beta}[1+\beta]$, while the infimum over $v \in K$, $|v| < k$ is no less than $e^{-k}[1 + k] - k^2$. It follows that (∗) holds for any k sufficiently small, so for $\delta_1 + \delta_2$ sufficiently small, Theorem 2 applies to yield the existence and Lipschitz regularity of a solution. □

2.6. *K* must be a cone.

It might be thought that the restriction in Theorem 2 to K being a cone is an artifact of the proof technique. However, the following example illustrates that even when K is a closed convex set containing 0, and all the other hypotheses of the theorem are present, existence may fail.

Example 2.6.1. We set $n = 2$, $T = 1$, and consider minimizing

$$\int_0^1 \left\{ \sqrt{1 + (\dot{x}_2(t))^2} - r\dot{x}_1(t)\sin|x_2(t)| \right\} dt$$

subject to $x(0) = (0, 0)$, $x(1) = (1, \pi/2)$, and $\dot{x}(t) \in K$ a.e., where r is positive and where $K = \{(v_1, v_2) : 0 \leq v_1 \leq 1, v_2 \geq 0\}$. When x is admissible and belongs to $\Gamma(\overline{x})$, we have

$$\int_0^1 |\dot{x}_2(t)| \, dt \quad \leq \quad \int_0^1 \{L(x(t), \dot{x}(t)) + r\} \, dt \quad \leq \quad \Lambda(\overline{x}) + r.$$

In addition, $|\dot{x}_1(t)|$ is bounded by 1. Consequently, (H1) holds for a suitable choice of k. We proceed to show that $(*)$ holds for all $k > 0$. We have

$$L(x, v) - v \cdot L_v(x, v) = \frac{1}{\sqrt{1 + v_2^2}}.$$

The supremum over K for $|v| > \beta$ tends to 0 with β, since $|v_1| \leq 1$ and $|v| \geq \beta$ implies $|v_2| \geq \sqrt{\beta^2 - 1}$. The infimum over K for $|v| < k$ is $1/\sqrt{1 + k^2}$, which confirms $(*)$. Thus all the hypotheses of Theorem 2 are present, except that K is not a cone.

We now claim that for r sufficiently large the problem admits no solution. To see this, observe that the constraints imply that any admissible arc x has $x_1(t) \equiv t$. Given this, the problem is equivalent to that of minimizing

$$\int_0^1 \left\{ \sqrt{1 + \dot{y}^2} - r\sin|y| \right\} dt$$

over the one-dimensional arcs y satisfying $y(0) = 0$, $y(1) = \pi/2$. We can show precisely as we did for Example 2.4.3 that no solution exists when r exceeds $1 + \pi/2$. □

3. Proof of Theorem 2

Let θ be a Nagumo function, and let AC_θ denote the class of arcs x in AC for which $\Lambda_\theta(x) < \infty$, where

$$\Lambda_\theta(x) := \int_0^T \theta(|\dot{x}(t)|) \, dt.$$

Consider the problem P_θ which is identical to (P) except that the arcs x are restricted to AC_θ; i.e., P_θ is the problem of minimizing $\Lambda(x)$ over the arcs x in AC_θ satisfying

$$x(t) \in \Omega \ \forall t, \qquad (x(0), x(T)) \in C, \qquad \dot{x}(t) \in K \quad \text{a.e.}$$

Most of the work lies in showing that P_θ admits a solution, as we shall see. We introduce a further parameter $\alpha \in \mathbb{R}$ by adding another constraint to P_θ; namely, $\Lambda_\theta(x) \leq \alpha$. This problem we label $P_\theta(\alpha)$, and the value of $P_\theta(\alpha)$ we denote by $V_\theta(\alpha)$. V_θ is the central figure in the proof; here are its salient properties.

Lemma 3.1. *For all α sufficiently large, $V_\theta(\alpha)$ is finite. Whenever $V_\theta(\alpha)$ is finite, the minimum defining $V_\theta(\alpha)$ is attained. $V_\theta(\cdot)$ is lower semicontinuous.*

Proof. Note that for α sufficiently large, the arc $\overline{x}$ is feasible for the problem $P_\theta(\alpha)$, so that $P_\theta(\alpha)$ admits feasible arcs. For any α having this last property, we have a minimizing sequence $\{x_i\}$ for the problem $P_\theta(\alpha)$, and for this sequence we have $\Lambda_\theta(x_i) \le \alpha$. This is known to imply [3, Theorem 10.3.i] that a subsequence of $\{\dot{x}_i\}$ (we eschew relabeling) converges weakly to an element v of $L^1([0, T]; \mathbb{R}^n)$. Since, by the basic hypotheses on C, either $\{x_i(0)\}$ or $\{x_i(T)\}$ is bounded, we may also suppose that $\{x_i(0)\}$ (to take one case) converges to a point x_0. It follows that the sequence of arcs $\{x_i\}$ converges uniformly to the arc x defined by

$$x(t) := x_0 + \int_0^t v(\tau)\, d\tau.$$

We have $\Lambda_\theta(x) \le \alpha$, for the functional Λ_θ is lower semicontinuous because θ is convex. The functional Λ being lower semicontinuous as well, we have

$$\Lambda(x) \le \lim_{i \to \infty} \Lambda(x_i) = V_\theta(\alpha).$$

The basic hypothesis on L implies that $\Lambda(x)$ is finite. Since the limit x is easily seen to be feasible for $P_\theta(\alpha)$, it follows that x solves $P_\theta(\alpha)$. Further, $V_\theta(\alpha)$ is finite. Let us now check that V_θ is lower semicontinuous. Let α_i converge to α and $V_\theta(\alpha_i)$ converge to a limit (finite or $+\infty$); we wish to show $\lim V_\theta(\alpha_i) \ge V_\theta(\alpha)$. If $V_\theta(\alpha_i) = +\infty$ for all large i, there is nothing to prove. In the contrary case, $P_\theta(\alpha_i)$ admits a solution x_i. By the same argument as above, we can suppose that x_i converges to an arc x in the sense described above. Since $\Lambda_\theta(x) \le \lim \Lambda_\theta(x_i) \le \alpha$, x is feasible for $P_\theta(\alpha)$ and we have

$$V_\theta(\alpha) \le \Lambda(x) \le \lim \Lambda(x_i) = \lim V_\theta(\alpha_i),$$

completing the proof.   □

Lemma 3.2. *P_θ admits a solution iff $V_\theta(\cdot)$ is eventually constant.*

Proof. Suppose first that V_θ is eventually constant. By Lemma 3.1, V_θ is eventually finite. Thus, for some $\alpha_0 > 0$, we have $V_\theta(\alpha_0)$ finite, and $V_\theta(\alpha) = V_\theta(\alpha_0)$ for all $\alpha \ge \alpha_0$. Let x_0 be an arc solving $P_\theta(\alpha_0)$ (see Lemma 3.1); we claim that x_0 solves P_θ. If not, there is an element x of AC_θ feasible for P_θ such that $\Lambda(x) < \Lambda(x_0)$. Choose any $\alpha > \alpha_0$ such that $\Lambda_\theta(x) < \alpha$; this is possible since x belongs to AC_θ. Then, by definition, we have

$$V_\theta(\alpha) \le \Lambda(x) < \Lambda(x_0) = V_\theta(\alpha_0),$$

a contradiction that completes the proof of one implication. The proof of the other, which we do not require, is omitted.   □

We now proceed to show that $V_\theta(\cdot)$ is eventually constant. In view of the result presented in §6, it suffices to establish that for α sufficiently large, whenever ζ is a proximal subgradient of V_θ at α, we have $\zeta = 0$. So let $\zeta \ne 0$ be a proximal subgradient of V_θ at α; we shall show that for α large a contradiction ensues. Note that V_θ is nonincreasing; this implies that its proximal subgradients are nonpositive. Thus we have $\zeta = -r$, where r is a positive quantity.

Now let z be a solution of $P_\theta(\alpha)$; thus $V_\theta(\alpha) = \Lambda(z)$ and $\Lambda_\theta(z) \leq \alpha$. We claim that equality must hold in this last inequality. For suppose not. Then $\Lambda_\theta(z) < \alpha$, and for any α' between $\Lambda_\theta(z)$ and α we have

$$V_\theta(\alpha') \leq \Lambda(z) = V_\theta(\alpha) \leq V_\theta(\alpha').$$

It follows that V_θ is constant on $(\Lambda_\theta(z), \alpha]$, which, by the definition of proximal subgradient, implies $r = 0 = \zeta$, contrary to assumption. Thus $\Lambda_\theta(z) = \alpha$.

The inequality characterizing r says that for some $\sigma > 0$, for some $\delta > 0$, and for all α' within δ of α, one has

$$V_\theta(\alpha') - V_\theta(\alpha) + \sigma|\alpha' - \alpha|^2 \geq -r(\alpha' - \alpha).$$

Let us replace $V_\theta(\alpha)$ by $\Lambda(z)$, α by $\Lambda_\theta(z)$ in this inequality, and α' by $\Lambda_\theta(x)$, where x is such that $\Lambda_\theta(x)$ is within δ of $\Lambda_\theta(z)$. Note that $V_\theta(\alpha') \leq \Lambda(x)$, which leads to

$$\Lambda(x) + r\Lambda_\theta(x) + \sigma|\Lambda_\theta(x) - \Lambda_\theta(z)|^2 \geq \Lambda(z) + r\Lambda_\theta(z).$$

This is the situation for which §7 was created. We deduce the existence of a constant c and a function p as described in Theorem 5 such that

$$(3.1) \qquad L(z(t), \dot{z}(t)) - \dot{z}(t) \cdot p(t) + r\theta(|\dot{z}(t)|) - r|\dot{z}(t)|\theta'(|\dot{z}(t)|) = c \quad \text{a.e.}$$

We have seen that $\Lambda_\theta(z) = \alpha$; thus

$$\operatorname*{ess\,sup}_{0 \leq t \leq T} |\dot{z}(t)| > \beta,$$

where β is any number less than $\theta^{-1}(\alpha/T)$. Note that we can arrange for β to be arbitrarily large by taking α sufficiently large, since $\theta^{-1}(\alpha/T)$ goes to $+\infty$ as α goes to $+\infty$. We derive from (3.1) that for some $x_1 \in \Omega$, for some $v_1 \in K$ with $|v_1| > \beta$, and for some p_1 in $\partial_v L(x_1, v_1)$, one has

$$L(x_1, v_1) - v_1 \cdot p_1 + r\theta(|v_1|) - r|v_1|\theta'(|v_1|) = c.$$

We deduce from this equality

$$c \leq \sup_{\substack{x \in \Omega, v \in K \\ |v| > \beta}} \{L(x, v) - v \cdot \partial_v L(x, v)\} + r \sup_{\substack{x \in \Omega, v \in K \\ |v| > \beta}} \{\theta(|v|) - |v|\theta'(|v|)\}.$$

Let us return again to (3.1). Since, for α sufficiently large, $\overline{x}$ is feasible for the problem $P_\theta(\alpha)$, it follows that z belongs to $\Gamma(\overline{x})$. By (H1), $|\dot{z}(t)|$ is less than k on a set of positive measure. So (3.1) gives

$$c \geq \inf_{\substack{x \in \Omega, v \in K \\ |v| < k}} \{L(x, v) - v \cdot \partial_v L(x, v)\} + r \inf_{\substack{x \in \Omega, v \in K \\ |v| < k}} \{\theta(|v|) - |v|\theta'(|v|)\}.$$

Together with the preceding inequality on c, this implies

$$r \left[\inf_{\substack{x \in \Omega, v \in K \\ |v| < k}} \{\theta(|v|) - |v|\theta'(|v|)\} - \sup_{\substack{x \in \Omega, v \in K \\ |v| > \beta}} \{\theta(|v|) - |v|\theta'(|v|)\} \right]$$

$$\leq \sup_{\substack{x \in \Omega, v \in K \\ |v| > \beta}} \{L(x, v) - v \cdot \partial_v L(x, v)\} - \inf_{\substack{x \in \Omega, v \in K \\ |v| < k}} \{L(x, v) - v \cdot \partial_v L(x, v)\}.$$

For α, and hence β, sufficiently large, the right side of this inequality is negative by hypothesis (H2), while the bracketed term on the left side is positive

(Lemma 2.3.1 applied to θ). It follows that r is negative, a contradiction that establishes that V_θ is eventually constant and hence that P_θ has a solution.

Let us label the solution of P_θ by z. Theorem 5 (for $r = 0$) shows that, for some constant c, one has

$$(3.2) \qquad c \in L(z(t), \dot{z}(t)) - \dot{z}(t) \cdot \partial_v L(z(t), \dot{z}(t)) \quad \text{a.e.}$$

This relation, in the presence of (H2), implies that $\dot{z}(t)$ is essentially bounded (the argument is the one given above, for $r = 0$).

All the above has been done for a fixed Nagumo function θ. Let ψ be any other Nagumo function. We claim that z also solves the problem P_ψ. For let the Lipschitz arc y solve P_ψ. (The solution y is provided by the proof above, for θ replaced by ψ.) Since y is feasible for P_θ, and z for P_ψ, we have

$$\Lambda(y) \leq \Lambda(z) \leq \Lambda(y),$$

which establishes the claim.

Now let x be any arc feasible for the original problem (P). There exists some Nagumo function ψ such that x belongs to AC_ψ [3, Theorem 10.3.i (c)]. Consequently, x is feasible for P_ψ. But then, since z solves P_ψ, we have $\Lambda(z) \leq \Lambda(x)$. Thus z solves (P). Finally, any other solution of (P) would solve P_ψ for some Nagumo function ψ, and hence, like z, would be Lipschitz and satisfy (3.2). This completes the proof of the theorem. $\square$

4. THE NONAUTONOMOUS PROBLEM

4.1. The modified hypotheses. We proceed to adapt the basic hypotheses to the case in which the Lagrangian L has explicit dependence on the t variable: $L(t, s, v)$. We shall assume that L is nonnegative. In all other cases, our new hypotheses reduce to what we had in the preceding sections if the t dependence is absent. The sets Ω, C, and K are as in §2. We assume that L is lower semicontinuous in (x, v), and that, for each (t, x) in $[0, T] \times \Omega$, the function $v \to L(t, x, v)$ is convex and has as domain a nonempty open convex set. All the other basic hypotheses remain in force, as does (H1) and the statement of the problem.

The dependence of L on t is required to be Lipschitz for each (x, v). More precisely, let $x \in \Omega$, $t \in [0, T]$, and $v \in \text{dom } L(t, x, \cdot)$ be given. Then the function $s \to L(s, x, v)$ is required to be Lipschitz for $s \in [0, T]$. Thus the derivative $L_s(s, x, v)$ exists for almost all s in $[0, T]$, and we further require that for nonnegative constants κ and η (not depending on (t, x, v)) the following holds:

$$(4.1) \qquad |L_s(s, x, v)| \leq \kappa L(s, x, v) + \eta.$$

Our hypotheses imply that L is lower semicontinuous in (t, x, v), and hence $\mathscr{L} \times \mathscr{B}$ measurable (see [4]), so that $\Lambda(x)$ is well defined (possibly as $+\infty$) for every admissible arc x.

The modification of $(*)$ that we require is the following:

$(*)'$

$$\lim_{\substack{\beta \to \infty}} \sup_{\substack{x \in \Omega, v \in K \\ |v| > \beta \\ t \in [0, T]}} \{L(t, x, v) - v \cdot \partial_v L(t, x, v)\} + \{\kappa \Lambda(\overline{x}) + \eta T\}$$

$$< \inf_{\substack{x \in \Omega, v \in K \\ |v| < k \\ t \in [0, T]}} \{L(t, x, v) - v \cdot \partial_v L(t, x, v)\}.$$

As in §2, we require this to hold for some k satisfying (H1); we label this hypothesis (H2)'. As in the autonomous case, (H1) and (H2)' are satisfied when L is coercive. Notice that (H2)' reduces to (H2) when L is autonomous, for then we may take $\kappa = \eta = 0$.

Theorem 3. *Let the modified basic hypotheses hold, as well as* (H1) *and* (H2)'. *Then* (P) *admits a solution. Furthermore, any solution* x *is Lipschitz and satisfies*

$$L(t, x(t), \dot{x}(t)) - \dot{x}(t) \cdot p(t) = c + \int_0^t \xi(\tau) \, d\tau \quad a.e. ,$$

where c *is a constant, and where* p *and* ξ *are measurable functions satisfying*

$$p(t) \in \partial_v L(t, x(t), \dot{x}(t)) \quad a.e.,$$

$$\xi(t) \in \partial_t L(t, x(t), \dot{x}(t)) \quad a.e.$$

Remark. Aside from the existence of a solution, the Lipschitz regularity of any solution under the stated conditions is a new result of independent interest, and adds to the known criteria for solutions to be Lipschitz (see [5, Chapter 2]).

4.2. Proof of the theorem. The proof in §3 of the autonomous case needs only slight modifications that we now indicate. The first arises in (3.1), which now becomes

(4.2)

$$L(t, z(t), \dot{z}(t)) - \dot{z}(t) \cdot p(t) + r\theta(|\dot{z}(t)|) - r|\dot{z}(t)|\theta'(|\dot{z}(t)|)$$

$$= c + \int_0^t \xi(\tau) \, d\tau \quad a.e.$$

Using (4.2), the same reasoning employed in §3 then yields the following upper bound on c:

$$\sup_{\substack{x \in \Omega, v \in K \\ |v| > \beta \\ t \in [0, T]}} \{L(t, x, v) - v \cdot \partial_v L(t, x, v)\}$$

$$+ r \sup_{\substack{x \in \Omega, v \in K \\ |v| > \beta \\ t \in [0, T]}} \{\theta(|v|) - |v|\theta'(|v|)\} + \sup_t \int_0^t \xi(\tau) \, d\tau,$$

and the following lower bound on c:

$$\inf_{\substack{x \in \Omega, v \in K \\ |v| < k \\ t \in [0, T]}} \{L(t, x, v) - v \cdot \partial_v L(t, x, v)\}$$

$$+ r \inf_{\substack{x \in \Omega, v \in K \\ |v| < k \\ t \in [0, T]}} \{\theta(|v|) - |v|\theta'(|v|)\} + \inf_t \int_0^t \xi(\tau) \, d\tau.$$

These estimates together with (H2)' will then complete the proof just as in the autonomous case, provided we can show that the term

$$\sup_t \int_0^t \xi(\tau)\,d\tau - \inf_t \int_0^t \xi(\tau)\,d\tau$$

is bounded above by $\kappa\Lambda(\overline{x})+\eta T$. To see this, note that the term we are seeking to bound can be expressed, to within an arbitrary degree of accuracy, in the form $\pm \int_{t_0}^{t_1} \xi(\tau)\,d\tau$, where t_0 and t_1 are points in the interval $[0,T]$, with $t_0 < t_1$. We have

$$\pm \int_{t_0}^{t_1} \xi(\tau)\,d\tau \le \int_{t_0}^{t_1} |\xi(\tau)|\,d\tau$$

$$\le \int_{t_0}^{t_1} \{\kappa L(\tau,z(\tau),\dot{z}(\tau))\,d\tau + \eta\}\,d\tau$$

(by (4.1))

$$\le \kappa\Lambda(z) + \eta T$$

(since L is nonnegative)

$$\le \kappa\Lambda(\overline{x}) + \eta T$$

(since, as before, $\overline{x}$ is feasible for the problem solved by z). This completes the proof. □

4.3. **An example.** We take $\eta = 1$, $T = 1$, and

$$L(t,x,v) = \varphi(t)\sqrt{1+v^2},$$

where φ is a continuously differentiable function whose minimum on $[0,1]$ is $m > 0$. We consider the problem (P) of minimizing $\Lambda(x)$ subject to $x(0) = 0$, $x(1) = c > 0$. Thus, for this example, $K = \mathbb{R}$. Futhermore, since L satisfies the linear growth condition $L(t,x,v) \ge m|v|$, we may add a constraint $x(t) \in \Omega$ for some compact set Ω without altering the nature of the problem. For $\overline{x}$ we take the linear arc joining $(0,0)$ to $(1,c)$; we have

$$\Lambda(\overline{x}) = \overline{\varphi}\sqrt{1+c^2},$$

where

$$\overline{\varphi} := \int_0^1 \varphi(t)\,dt.$$

It follows as in § 2.3 that (H1) holds for any k exceeding to $\overline{\varphi}(1+c^2)^{1/2}/m$. Let us now examine (4.1). Set

$$\mu := \max_{0 \le t \le T} |\varphi'(t)|.$$

Then $|\varphi'(t)| \le (\mu/m)\varphi(t)$ for all t, and it follows that L satisfies (4.1) with $\kappa = \mu/m$, $\eta = 0$.

There remains to satisfy (H2)' for some k arbitrarily near $\overline{\varphi}(1+c^2)^{1/2}/m$; this condition is readily calculated to reduce to

$$(2\mu\overline{\varphi}/m)\sqrt{1+c^2} < \frac{m}{\sqrt{1+k^2}} = \frac{m^2}{\sqrt{m^2+\overline{\varphi}^2(1+c^2)}}.$$

Since $\overline{\varphi}$ is bounded above by $m + \mu$, a sufficient condition to insure this inequality is the following:

$$2\mu\,(m + \mu)\sqrt{1 + c^2}\;\sqrt{m^2 + (m + \mu)^2(1 + c^2)}\; < \; m^3.$$

We therefore deduce:

Proposition 4.1. *For given* m *and* c, (P) *admits a solution if* φ *is sufficiently close to being constant; i.e., if* μ *is sufficiently small.*

In the analysis above, we have of course been speaking of conditions which are merely *sufficient* to guarantee the existence of a solution. It turns out, however, that a solution actually fails to exist unless some such restriction is placed upon φ. Because the functional Λ is convex, an exact characterization of the conditions on φ under which existence holds is possible; this is a result of Kaiser (see [3, p. 440]).

5. THE CASE $n = 1$

When $n = 1$ a somewhat different tack can be taken in the proof to allow us to exploit the mean value theorem in producing the k invoked in (H1) and (H2). The following additions are made to the basic hypotheses of §4 (which remain in force):

(5.1)   every arc x in $\Gamma(\overline{x})$ is such that $x(t) \in \operatorname{int}\Omega\;\forall t$;

(5.2) L is locally Lipschitz in (t, x, v) and satisfies, for constants k_0 and c_0,

$$|\partial_x L(t, x, v)| \le k_0 |L(t, x, v)| + c_0, \quad \forall (t, x, v) \in [0, T] \times \Omega \times \mathbb{R}^n.$$

We no longer assume (H1), but rather that $(*)'$ (see §4) holds for a value of k greater than k_0, where

$$k_0 := \sup\left\{ \frac{|x_1 - x_0|}{T} : (x_0, x_1) \in C \right\}.$$

Theorem 4. *When* $n = 1$ *and the above hypotheses hold, then the conclusions of Theorem 3 are valid.*

Proof. The proof remains essentially that of Theorem 3 (see §4), the only change being how the lower bound on c is obtained. When Theorem 5 is invoked, the second part yields that z is continuously differentiable. Since z is admissible for $P_\theta(\alpha)$, we have $(z(0), z(T)) \in C$. The mean value theorem implies that for some t^* in $(0, T)$, one has $\dot{z}(t^*)$ equal to $(z(T) - z(0))/T$. We deduce

$$\operatorname*{ess\,inf}_{0 \le t \le T} |\dot{z}(t)| \le k_0,$$

where k_0 is defined above. The proof then proceeds without change.

Example. Let us return to Example 2.4.3, in the special case $n = 1$. We have $k_0 = 0$, so $(*)'$ holds for some k exceeding k_0 provided $2r < 1$. Since all the other hypotheses of Theorem 4 are satisfied, we obtain existence for all such r. This is an improvement over what was obtained by applying Theorem 2. □

6. A RESULT IN NONSMOOTH ANALYSIS

Let U be an open subset of $\mathbb{R}^n$, and let f be a lower semicontinuous function mapping U to $(-\infty, \infty]$.

Proposition 6.1. *Suppose that whenever ζ is a proximal subgradient to f at a point x in U, we have $\zeta = 0$. Then f is locally constant in U.*

Proof. Let x_0 be a point in U at which f is finite (there is nothing to prove if no such point exists), and choose $\varepsilon > 0$ so that the ε-ball about x_0 is contained in U, and so that f is bounded below on the set $x_0 + \varepsilon B$. Let x be any other point in $x_0 + \varepsilon B$; let us proceed to show that $f(x) \geq f(x_0)$. This is evident if $f(x) = +\infty$, so we can assume that $f(x)$ is finite. Now suppose that we had $f(x) < f(x_0)$, and let us derive a contradiction. There is a continuously differentiable function g on $x_0 + \varepsilon B$ which has a unique global minimum at x_0, whose derivative vanishes only at x_0, and such that both $g(w)$ and $|g'(w)|$ approach $+\infty$ as w approaches the boundary of $x_0 + \varepsilon B$. Further, we can scale g so that

$$g(x) - g(x_0) < f(x_0) - f(x).$$

The function $f + g$ has a global minimum at some point z in $x_0 + \varepsilon B$, so $\zeta = 0$ is a proximal subgradient of $f + g$ at z. By the approximate sum formula [5, Proposition 1.4], there exist for each integer i points u_i and v_i within $1/i$ of z such that

$$|\zeta_i + \xi_i| < 1/i \ ,$$

where ζ_i is a proximal subgradient of f at u_i and ξ_i is a proximal subgradient of g at v_i. Now we have $\zeta_i = 0$ by hypothesis, and also $g'(v_i) = \xi_i$. Since v_i converges to z and $g'(v_i)$ converges to 0, we must have $g'(z) = 0$, whence the minimizer z equals x_0. Thus

$$f(x) + g(x) \geq f(x_0) + g(x_0),$$

which contradicts an earlier statement about g.

It follows from the above that f has a constant value f_0 at all points in $x_0 + \varepsilon B$ at which f is finite. To complete the proof, it suffices to show that f is finite everywhere on $x_0 + (\varepsilon/4)B$. Suppose not; let z in $x_0 + (\varepsilon/4)B$ have $f(z) = +\infty$. Pick $\delta > 0$ with $\delta < \varepsilon/4$, and consider the point $(z, f_0 - \delta)$ in $\mathbb{R}^n \times \mathbb{R}$. This point is outside the closed set epi f. Let (w, r) be a point in epi f closest to $(z, f_0 - \delta)$. Then, since (x_0, f_0) belongs to epi f,

$$|(z, f_0 - \delta) - (w, r)| \leq |(z, f_0 - \delta) - (x_0, f_0)| \leq |z - x_0| + \delta < \varepsilon/2.$$

It follows that $|z - w|$ is less than $\varepsilon/2$, so that w belongs to the set $x_0 + \varepsilon B$. We deduce that $r = f(w) = f_0$. We have now shown that $(w, f(w)) = (w, f_0)$ is the closest point in epi f to the point $(z, f_0, -\delta)$. In geometric terms, this says that $(z - w, -\delta)$ is a proximal normal [5] to epi f at $(w, f(w))$, which is equivalent to the fact that the nonzero vector $(z - w)/\delta$ is a proximal subgradient of f at w. This contradiction of the hypothesis completes the proof. □

Remark. A simpler, self-contained proof of this result is given by F. H. Clarke and R. M. Redheffer in an article to appear in the Canadian Mathematical Bulletin.

7. A RESULT IN NONSMOOTH OPTIMIZATION

We now derive the necessary conditions needed in the proof of Theorem 2. The hard work has already been done in [4]; it is a matter of doing the bookkeeping that makes this evident. The hypotheses on L are those of §4 except that we do not require L to be nonnegative (the hypotheses therefore reduce to those of §2 when L is autonomous), and θ is a Nagumo function. We suppose that z is an arc satisfying the constraints of (P):

$$(x(0),\, x(T)) \in C, \qquad x(t) \in \Omega \quad (0 \le t \le T), \qquad \dot{x}(t) \in K \quad \text{a.e.},$$

and for positive constants r and σ we consider the functional f on AC_θ given by

$$f(x) := \Lambda(x) + r\Lambda_\theta(x) + \sigma|\Lambda_\theta(x) - \Lambda_\theta(z)|^2.$$

Theorem 5. *Suppose that for every arc x admissible for (P), whenever $\Lambda_\theta(x)$ is sufficiently near $\Lambda_\theta(z)$, one has $f(x) \ge f(z)$. Then there exists an integrable function $\xi(t)$ such that*

$$\xi(t) \in \partial_t L(t,\, z(t),\, \dot{z}(t)) \quad \text{a.e.}$$

and, for some constant c,

(7.1)
$$\begin{aligned} & L(t,\, z(t),\, \dot{z}(t)) - \langle \dot{z}(t),\, p(t) \rangle + r\theta(|\dot{z}(t)|) - r|\dot{z}(t)|\theta'(|\dot{z}(t)|) \\ & \qquad = c + \int_0^t \xi(\tau)\, d\tau \quad \text{a.e.}, \end{aligned}$$

where, for t almost everywhere, $p(t)$ belongs to $\partial_v L(t,\, z(t),\, \dot{z}(t))$. If in addition we have $n = 1$, if $(*)'$ holds for some k, and if the hypotheses (5.1), (5.2) hold, then z is continuously differentiable.

Proof. Let u be any measurable function on $[0,\, T]$ with values in the interval $[-\varepsilon_0,\, \varepsilon_0]$, where ε_0 is a positive parameter less than 1, and where $\int_0^T u(t)\, dt = 0$. We use u to define an invertible transformation of $[0,\, T]$ into itself:

$$t = \tau + \int_0^\tau u(s)\, ds.$$

This induces via reparametrization a new arc x:

$$x(t) = z(\tau), \quad \text{where } \tau \text{ corresponds to } t \text{ as above.}$$

This is known as the Erdmann transformation.

We find $\dot{x}(t) = \dot{z}(\tau)/(1 + u(\tau))$, so that $\dot{x}(t)$ belongs to K a.e. Further, x satisfies the other constraints of (P). Let us calculate $\Lambda_\theta(x)$:

$$\int_0^T \theta(|\dot{x}(t)|)\, dt = \int_0^T \theta\!\left(\frac{|\dot{z}(\tau)|}{1 + u(\tau)}\right)(1 + u(\tau))\, d\tau.$$

For any given $\varepsilon > 0$, there exists for almost every τ a positive number $\gamma(\tau)$ such that if $|u(\tau)| \le \gamma(\tau)$ then the last integrand above is within ε of $\theta(|\dot{z}(\tau)|)$. We shall limit the choice of u in this way. The function γ can be taken measurable, and also such that for t a.e., $\dot{x}(t)$ lies in $\operatorname{int}(\operatorname{dom} L(t,\, x(t),\, \cdot))$.

For ε sufficiently small, we have by hypothesis $f(x) \ge f(z)$. We shall translate this into terms of an optimal control problem whose state is $(s,\, y)$ in

$\mathbb{R}^2$ and whose dynamics are

$$\frac{ds}{dt} = 1 + u(t), \qquad \frac{dy}{dt} = \theta\left(\left|\frac{\dot{z}(t)}{1 + u(t)}\right|\right)(1 + u(t)).$$

The control u is subject to $|u(t)| \leq \gamma(t)$, and the boundary conditions are $s(0) = 0$, $s(T) = T$, $y(0) = 0$. The objective functional integrand $\widetilde{L}(t, s, u)$ is given by

$$\widetilde{L}(t, s, u) := L\left(s, z(t), \frac{\dot{z}(t)}{1 + u}\right)(1 + u) + r\theta\left(\left|\frac{\dot{z}(t)}{1 + u}\right|\right)(1 + u),$$

and we know that the control $u \equiv 0$, and corresponding state

$$s(t) = t, \qquad y(t) = \int_0^t \theta(|\dot{z}(\tau)|)\, d\tau,$$

solve the problem of minimizing

$$\sigma|y(T) - \Lambda_\theta(z)|^2 + \int_0^T \widetilde{L}(t, s(t), u(t))\, dt.$$

Our goal is to apply the necessary conditions of [4, Theorem 5.2.1], whose hypotheses we need to verify. It follows easily that $\widetilde{L}$ is $\mathscr{L} \times \mathscr{B}$ measurable, and that the control set depends measurably on t. The Lipschitz requirement on $\widetilde{L}$ requires that for some integrable function $k_0(t)$, for some $\delta > 0$, for all s_1 and s_2 in the ball $B(t, \delta)$, and for t a.e., one has

$$|L(s_1, z(t), \dot{z}(t)) - L(s_2, z(t), \dot{z}(t))| \leq k_0(t)|s_1 - s_2|.$$

The growth hypothesis (4.1) on the t-dependence of L leads precisely to this conclusion, as demonstrated in [4, p. 181].

The necessary conditions yield the existence of arcs λ_1, λ_2 and a scalar λ_0 equal to 0 or 1, such that

$$\dot{\lambda}_1(t) \in \lambda_0 \partial_t L(t, z(t), \dot{z}(t)) \quad \text{a.e.}, \qquad \dot{\lambda}_2(t) = 0,$$

$$\max_{-\gamma(t) \leq u \leq \gamma(t)} \left\{\lambda_1(t)(1 + u) + \lambda_2(t)\theta\left(\left|\frac{\dot{z}(t)}{1 + u}\right|\right)(1 + u) - \lambda_0 \widetilde{L}(t, 0, u)\right\}$$

occurs at $u = 0$ a.e. and such that if $\lambda_0 = 0$, then (λ_1, λ_2) is nonvanishing. In addition, we have the transversality condition $\lambda_2(T) = 0$.

Analysis of these conditions shows that λ_2 is identically zero. If $\lambda_0 = 0$, then the maximum condition gives $\lambda_1 \equiv 0$ as well, which is proscribed; hence $\lambda_0 = 1$. It follows that for some constant c and for some ξ as described in the statement of the theorem, one has

$$\lambda_1(t) = c + \int_0^t \xi(\tau)\, d\tau.$$

Taking the subdifferential at 0 of the quantity being maximized a.e., and expressing the fact that 0 belongs to it, gives precisely (7.1).

There remains to prove the final assertion of the theorem. In the context described there, one seeks now to apply the necessary conditions of [4, Theorem 5.2.1] directly to the problem without reducing it to a one-dimensional problem via the Erdmann transformation. By reframing it much as we did above, and

invoking the fact that (in the added presence of $(*)'$) we know $\dot{z}$ to be essentially bounded, we can invoke the necessary conditions to deduce the existence of an arc p such that for t a.e., the maximum

$$\max_{v \in K} \{p(t) \cdot v - \lambda_0 L(t, z(t), v) - \lambda_0 r\theta(|v|)\}$$

occurs at $v = \dot{z}(t)$, where λ_0 is either 0 or 1, and where p is nonvanishing if $\lambda_0 = 0$. The arc p also satisfies the adjoint equation

$$\dot{p}(t) \in \lambda_0 \partial_x L(t, z(t), \dot{z}(t)) \quad \text{a.e.}$$

Suppose first that $\lambda_0 = 0$. Then it follows from the above that $p(t) \equiv p_0 \neq 0$. From the maximum condition, it now ensues that $\dot{z}(t) = 0$ a.e., since the only point in a cone K in $\mathbb{R}$ admitting a nonzero normal vector is the origin. Thus z is constant, and hence C^1.

Let us now examine the case $\lambda_0 = 1$. We then have, from the maximum condition, $p(t) \in \partial_v g(t, \dot{z}(t))$ a.e., where g is given by

$$g(t, v) := L(t, z(t), v) + r\theta(|v|) + \psi_K(v) ,$$

ψ_K being the indicator of K. Because $g(t, \cdot)$ is strictly convex, this implies that $\dot{z}$ is continuous; the argument is given in [7, p. 86].

References

1. L. Ambrosio, O. Ascenzi, and G. Buttazzo, *Lipschitz regularity for minimizers of integral functionals with highly discontinuous integrands*, J. Math. Anal. Appl. **142** (1989), 301–316.

2. B. Botterton and P. Marcellini, *A general approach to the existence of minimizers of one-dimensional non-coercive integrals of the calculus of variations*, preprint.

3. L. Cesari, *Optimization—Theory and applications*, Springer-Verlag, New York, 1983.

4. F. H. Clarke, *Optimization and nonsmooth analysis*, 2nd ed., Classics in Applied Math. V, SIAM, Philadelphia, Pa., 1990.

5. _____, *Methods of dynamic and nonsmooth optimization*, CBMS-NSF Regional Conf. Ser. in Appl. Math., vol. 57, SIAM, Philadelphia, Pa., 1989.

6. F. H. Clarke and P. D. Loewen, *An intermediate existence theory in the calculus of variations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **16** (1989), 487–526.

7. F. H. Clarke and R. B. Vinter, *Regularity properties of solutions to the basic problem in the calculus of variations*, Trans. Amer. Math. Soc. **289** (1985), 73–98.

8. L. Tonelli, *Sur une méthode directe du calcul des variations*, Rend. Circ. Mat. Palermo **39** (1915), 233–264; also in Opere Scelte, Vol. II, Edizioni Cremonese, Rome, 1961, pp. 289–333.

CENTRE DE RECHERCHES MATHÉMATIQUES (CRM), UNIVERSITÉ DE MONTRÉAL, C.P. 6128–A, MONTRÉAL, QC H3C 3J7, CANADA