

## REGULARITY PROPERTIES OF SOLUTIONS TO THE BASIC PROBLEM IN THE CALCULUS OF VARIATIONS

BY

F. H. CLARKE<sup>1</sup> AND R. B. VINTER

ABSTRACT. This paper concerns the basic problem in the calculus of variations: minimize a functional  $J$  defined by

$$J(x) = \int_a^b L(t, x(t), \dot{x}(t)) dt$$

over a class of arcs  $x$  whose values at  $a$  and  $b$  have been specified.

Existence theory provides rather weak conditions under which the problem has a solution in the class of absolutely continuous arcs, conditions which must be strengthened in order that the standard necessary conditions apply. The question arises: What necessary conditions hold merely under hypotheses of existence theory, say the classical Tonelli conditions? It is shown that, given a solution  $x$ , there exists a relatively open subset  $\Omega$  of  $[a, b]$ , of full measure, on which  $x$  is locally Lipschitz and satisfies a form of the Euler-Lagrange equation.

The main theorem, of which this is a corollary, can also be used in conjunction with various classes of additional hypotheses to deduce the global smoothness of solutions. Three such classes are identified, and results of Bernstein, Tonelli, and Morrey are extended. One of these classes is of a novel nature, and its study implies the new result that when  $L$  is independent of  $t$ , the solution has essentially bounded derivative.

**1. Introduction.** The basic problem in the calculus of variations, which we denote by (P), is that of minimizing the functional  $J$  defined by

$$J(x) := \int_a^b L(t, x(t), \dot{x}(t)) dt$$

over a given class of functions  $x$ , assuming given values at  $a$  and  $b$ :  $x(a) = A$ ,  $x(b) = B$ . (Here  $L$  is a function from  $[a, b] \times R^n \times R^n$  to  $R$ , and  $\dot{x}$  signifies the derivative of  $x$ .) It has been studied now for almost three hundred years.

One of the fundamental issues that was broached relatively late is that of *existence*: under what hypotheses on  $L$ , and within what class of functions  $x$ , can one be assured that a solution exists? It was Tonelli, in work that also had great significance in several areas of functional analysis, who was able to develop a satisfactory existence theory. Let us suppose that  $L$  is  $C^2$ , the function  $v \rightarrow L(t, x, v)$  satisfies globally  $L_{vv} \geq 0$  for all  $(t, x)$ , and the following coercivity condition holds:

$$L(t, x, v) \geq \beta|v|^2 + \lambda \quad \text{for all } (t, x, v),$$

---

Received by the editors January 29, 1984 and, in revised form, July 5, 1984.

1980 *Mathematics Subject Classification*. Primary 49B05.

<sup>1</sup>The support of the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged.

©1985 American Mathematical Society  
0002-9947/85 \$1.00 + \$.25 per page

where  $\beta$  is a positive constant. (We refer to these as the classical Tonelli hypotheses, although we have strengthened his slightly for ease of presentation.) Tonelli's celebrated existence theorem asserts that under these hypotheses, and with the class of competing functions taken to be the absolutely continuous functions mapping  $[a, b]$  to  $R^n$  (we call such functions *arcs* for short), a solution to (P) exists. The gamut of existence results for (P), including refinements and extensions of Tonelli's is presented in [3, Chapters 11–16].

Much of the attention in the calculus of variations has been devoted to necessary conditions for optimality, the best known of which is the Euler-Lagrange equation (strong form):

$$(1.1) \quad L_v(t, x(t), \dot{x}(t)) = c + \int_a^t L_x(s, x(s), \dot{x}(s)) ds \quad \text{a.e., } a \leq t \leq b.$$

One finds however that the customary derivations of necessary conditions such as this have imposed additional hypotheses, notably:

(a) an a priori assumption that  $x$  belongs to some special subclass of arcs (e.g., the piecewise-smooth functions of the classical calculus of variations, or arcs having essentially bounded derivative; see for example [2–4]), or

(b) a priori growth conditions on  $L$  of the form

$$(1.2) \quad |L_x| + |L_v| \leq c|L| + k$$

(also invoked in optimal control theory; see for example [7, 10, 12]).

The fact is that hypotheses of both these types can fail to be satisfied for very reasonable-looking problems satisfying the Tonelli conditions (as in the example of [9] in which  $L$  is a polynomial).

A natural question then, and the first that we address in this article, is to ask what necessary conditions hold under strictly those assumptions guaranteeing the existence of a solution (for example, the classical Tonelli hypotheses).

Tonelli was not insensible to this issue. Under the classical hypotheses he proved [14] that when  $n = 1$  (i.e., in the case of a single unknown function) and  $L_{vv}$  is everywhere strictly positive, then there is an open set  $\Omega$  in  $[a, b]$  of full measure such that in  $\Omega$  the solution  $x$  is  $C^2$  and satisfies

$$(1.3) \quad \frac{d}{dt} L_v(t, x(t), \dot{x}(t)) = L_x(t, x(t), \dot{x}(t)).$$

(We refer to this as *Tonelli's regularity theorem*.) Note that (1.3) is strictly weaker than (1.1) (since, in general, we cannot integrate (1.3)).

The first object of this article is to obtain regularity properties and necessary conditions for solutions to (P) in the vector case ( $n > 1$ ), and under hypotheses considerably weaker than the classical ones of Tonelli. This is done in Theorem 2.1, where we find that progressively more can be said about  $x$  as the hypotheses on  $L$  accumulate until we find as an extreme case the one we have labelled Tonelli's regularity theorem (see Corollary 3).

As implied above, the question of whether the stronger condition (1.1) holds is closely related to that of regularity of solutions: when can we assert that the solution  $x$  to (P) lies in certain desirable subclasses of arcs and, in so doing, obtain stronger necessary conditions? If, for example,  $\dot{x}$  is essentially bounded (i.e.,  $x$  is Lipschitz continuous) then (as is well known) under the classical Tonelli hypotheses  $x$  is  $C^2$

and satisfies (1.1). This is proven, for example, in [3, §2.6], together with other results in this vein, and some counterexamples are provided in different contexts in which the solution turns out to be neither Lipschitz nor smooth.

The complexity of these issues is illustrated by the Lavrentiev phenomenon (see [3, §18.5]) in which the minimum in (P) (over all feasible arcs) can be strictly less than the infimum taken over feasible arcs having essentially bounded derivative. In [3, §18.4] conditions due to T. S. Angell are given which guarantee sufficiently strong approximation properties to exclude the Lavrentiev phenomenon. These results, which subsume earlier ones of Tonelli and Mania, hinge upon a hypothesis called Condition (D) by Cesari and Suryanarayana, which also plays a key role in a large class of existence theorems (see [3, Chapter 13]).

The second object of this article is to explore the nature of additional hypotheses upon the problem which would serve to exclude (or limit) bad behaviour of the solution  $x$ . Of special interest are hypotheses which exclude points of bad behaviour of  $x$  (i.e., points at which  $x$  is not locally Lipschitz) altogether and which therefore assure validity of the strong form of the Euler–Lagrange equation (1.1). Several sets of hypotheses of this nature are given. In each case the regularity of  $x$  on an open set of full measure, established in Theorem 2.1, plays a decisive part in improving known results and in generating new ones.

First we show that points of bad behaviour can occur only at one endpoint of  $[a, b]$  or the other, depending on whether  $t \rightarrow L_t(t, x(t), \dot{x}(t))$  is bounded above or below by a summable function. We deduce as a consequence of these properties the striking, new result that in a nonsmooth vector setting, merely under hypotheses H1–H3, autonomous problems cannot give rise to points of bad behaviour, and, hence, solutions are globally Lipschitz and satisfy the Euler–Lagrange equation in strong form (1.1).

We then consider how Theorem 2.1 can be used to weaken growth conditions of the type (1.2), under which points of bad behaviour cannot occur. Tonelli, in the smooth, scalar case, has shown that the bound on  $L_v$  can be discarded. We generalize Tonelli's result to apply in a nonsmooth, vector setting and weaken, somewhat, the restrictions on  $L_x$ .

We next turn to hypotheses expressed in terms of growth assumptions on  $F = L_{vv}^{-1}[L_x - L_{vt} - L_{vx}]$  as a function of  $v$ . Tonelli established in the scalar case that points of bad behaviour cannot occur under hypotheses considered by Bernstein in which the growth of  $F$  is restricted to be at most quadratic. This, and a little bit more, is shown to be true also in a vector setting.

Finally, on this topic we present a novel technique which leads to information concerning regularity of  $x$  when  $F$  is merely polynomially bounded. Such a condition alone cannot be expected to exclude points of bad behaviour; indeed the example of [9] exhibiting such a point satisfies the condition. However, we can exclude them first when the Lagrangian is sufficiently close to a certain class of well-behaved Lagrangians, and second when the interval of integration is suitably small.

Note that each set of hypotheses in §3 which serves to exclude points of bad behaviour ensures that any absolutely continuous arc which is minimizing is actually Lipschitz and therefore excludes also the Lavrentiev phenomenon.

We conclude this introduction with a word regarding hypotheses and methodology. We impose the following conditions:

- (H1) The function  $(t, x, v) \rightarrow L(t, x, v)$  is locally bounded on  $[a, b] \times R^n \times R^n$ , measurable as a function  $t$ , and convex as a function of  $v$ .
- (H2)  $L$  is locally Lipschitz in  $(x, v)$  (uniformly in  $t$ ); i.e., for each bounded subset  $C$  of  $R^n \times R^n$ , there exists a constant  $K$  such that, for all  $(x_1, v_1), (x_2, v_2)$  in  $C$ , for all  $t$  in  $[a, b]$ , one has

$$|L(t, x_1, v_1) - L(t, x_2, v_2)| \leq K|(x_1 - x_2, v_1 - v_2)|.$$

- (H3) There is a constant  $\alpha$  and a convex function  $\theta: [0, \infty) \rightarrow R$  such that  $L(t, x, v) \geq -\alpha|x| + \theta(|v|)$  for all  $(t, x, v)$  in  $[a, b] \times R^n \times R^n$ , where  $\theta(r)/r \rightarrow \infty$  as  $r \rightarrow \infty$ .

It is easy to see that (H1)–(H3) extend the classical Tonelli hypotheses; in particular, it is not necessary that  $L$  be differentiable. This is not done merely in quest of the greatest generality, but rather because the methods and techniques of nonsmooth analysis and optimization [7] are central to our approach. Thus, nonsmooth analysis would remain an intrinsic part of the proof of Theorem 1, for example, even if  $L$  were presumed to satisfy the classical Tonelli hypotheses. We remark that the construction of an auxiliary Lagrangian with the requisite properties in that proof, if one were constrained to smooth functions, is quite problematic and may have posed the chief obstacle to treating the vector case ( $n > 1$ ).

The reader whose interest lies in the results as they pertain to smooth contexts need of course not be concerned with the methodology. To follow all the details, however, it is necessary to have some familiarity with the generalized gradient  $\partial f$  of a (not necessarily differentiable) function  $f$ . We refer to [7] (or [5, 6]) for an introduction to this topic.

## 2. The regularity theorem.

2.1. *Statement of the result.* We deal throughout this section with the basic variational problem (P) under hypotheses (H1)–(H3).

**THEOREM 2.1.** *A solution  $x(\cdot)$  to (P) exists. Let  $\tau$  be any point in the interval  $[a, b]$  for which*

$$(2.1) \quad \liminf_{\substack{s, t \rightarrow \tau \\ a \leq s \leq \tau \leq t \leq b \\ s \neq t}} \frac{|x(s) - x(t)|}{|s - t|} < \infty.$$

Then:

(i) *There is an interval  $I$  which is a neighborhood of  $\tau$  in  $[a, b]$  in which the arc  $x$  is Lipschitz and satisfies the Euler–Lagrange inclusion; i.e., there is an arc  $p$  defined in  $I$  such that for almost all  $t$  in  $I$  one has*

$$(2.2) \quad (\dot{p}(t), p(t)) \in \partial L(t, x(t), \dot{x}(t)).$$

(ii) If, in addition, for each  $t$  in  $[a, b]$  and  $w$  in  $R^n$  the function  $v \rightarrow L(t, x(t), v)$  is strictly convex and the function  $s \rightarrow L(s, x(t), w)$  is continuous at  $t$ , then  $x$  is  $C^1$  in  $I$ .

(iii) If in addition to the hypotheses of (ii), for each  $t$  in  $[a, b]$  the function  $L$  is  $C^r$  (in all its arguments,  $r \geq 2$ ) near  $(t, x(t), \dot{x}(t))$ , and  $L_{vv}(t, x(t), \dot{x}(t)) > 0$  (positive definite), then  $x$  is  $C^r$  in  $I$ .

(When  $\tau$  is either  $a$  or  $b$ , the term  $C^1$  may require explanation. We say  $x$  is  $C^1$  on  $[a, a + \varepsilon)$ , for example, when  $x$  is  $C^1$  on  $(a, a + \varepsilon)$  and  $\dot{x}(t)$  tends to a finite limit as  $t$  decreases to  $a$ .)

We pause to make a few remarks about the theorem before turning to its proof. Since  $x(\cdot)$  is absolutely continuous, and so differentiable a.e., it follows that those points  $\tau$  (if any) at which the mild condition (2.1) fails, form a set of zero measure. In consequence:

**COROLLARY 1.** *There is a set  $\Omega$  open in  $[a, b]$  and of full measure in which  $\dot{x}$  is locally bounded.*

Of course, a great deal of interest resides in the question of how large or small  $\Omega$  must be. The example of [9] confirms that under hypotheses (H1)–(H3) it is possible for the solution arc  $x$  to have unbounded derivative (so that necessarily  $\Omega$  cannot include all of  $[a, b]$ ). In §3 we develop, among other things, various additional hypotheses in the presence of which we can assert  $\Omega = [a, b]$ .

A word about the differential inclusion (2.2) is in order. The generalized gradient  $\partial L$  is taken with respect to the  $(x, v)$  variables (for each  $t$ ). If  $L$  happens to be  $C^1$  in these variables, then (2.2) implies that almost everywhere in  $I$  one has

$$\frac{d}{dt} \nabla_v L(t, x(t), \dot{x}(t)) = \nabla_x L(t, x(t), \dot{x}(t)),$$

the familiar form of the Euler–Lagrange equation. Necessary conditions for problems with nonsmooth data play an essential role in the proof of the theorem, regardless of whether the Lagrangian  $L$  is smooth or not. A discussion of such necessary conditions and the underlying nonsmooth calculus appears in Chapter 1 of [7].

In the presence of strict convexity, one of the conclusions of the theorem may be recast as follows:

**COROLLARY 2.** *Let  $L$  satisfy the hypotheses of (ii). Then for any  $\tau$  in  $[a, b]$  one has*

$$(2.3) \quad \liminf \frac{|x(s) - x(t)|}{|s - t|} = \limsup \frac{|x(s) - x(t)|}{|s - t|}$$

(where the limits are taken as in (2.1)).

This is easy to see, for if the left side of (2.3) is finite, then the theorem asserts that  $x(\cdot)$  is  $C^1$  near  $\tau$ , which implies equality (2.3), while if the left side of (2.3) is  $+\infty$ , then equality holds automatically.

When there is a single dependent variable (i.e.,  $n = 1$ ), the case considered by Tonelli, (2.3) lends itself to the interpretation that  $x(\cdot)$  is everywhere differentiable in an extended sense. The following is Tonelli's result captured under greatly reduced hypotheses:

**COROLLARY 3.** *If  $n = 1$  and  $L$  satisfies the hypotheses of (ii), then  $x(\cdot)$  is differentiable everywhere in  $[a, b]$ , in the sense that the following limit exists (finite or infinite) for each  $\tau$  in  $[a, b]$ :*

$$(2.4) \quad \lim_{\substack{t \rightarrow \tau \\ a \leq t \leq b}} \frac{x(t) - x(\tau)}{t - \tau}.$$

**PROOF.** If  $\tau$  is such that the left side of (2.1) is finite, then by the theorem  $x$  is  $C^1$  near  $\tau$ , whence the limit above exists (and is finite). When the left side of (2.1) is  $+\infty$  and  $\tau = a$ , the limit (2.4) could only fail to exist if one had

$$(2.5) \quad \limsup_{t \downarrow a} \frac{x(t) - x(a)}{t - a} = +\infty, \quad \liminf_{t \downarrow a} \frac{x(t) - x(a)}{t - a} = -\infty.$$

But if (2.5) holds, there are points  $t$  distinct from, but arbitrarily near,  $a$  for which  $x(t) = x(a)$ , whence the left side of (2.1) is finite, a contradiction. The case  $\tau = b$  is similar.

We are left then with the case in which  $\tau$  lies in  $(a, b)$  and the left side of (2.1) is  $+\infty$ . It follows (from reasoning as above) that the phenomenon to be ruled out is that

$$(2.6) \quad \lim_{t \downarrow \tau} \frac{x(t) - x(\tau)}{t - \tau} = +\infty, \quad \lim_{s \uparrow \tau} \frac{x(s) - x(\tau)}{s - \tau} = -\infty$$

(or the opposite, which is handled similarly). (The situation (2.6) is like that of the cusp at zero of the function  $r \rightarrow |r|^{1/2}$ .) Fix any  $\varepsilon > 0$ . We assert the existence of  $\delta > 0$  with the following property: for any value  $r$  in  $(0, \delta)$ , there is a point  $s$  in  $(\tau - \varepsilon, \tau)$  such that  $x(s) = x(\tau) + r$  (for if not, the intermediate value property applied to  $x$  yields  $x(s) \leq x(\tau)$  for  $s$  in  $(\tau - \varepsilon, \tau)$ , contradicting the second condition of (2.6)). By the same reasoning there is a point  $t$  lying in  $(\tau, \tau + \varepsilon)$  and some  $r$  in  $(0, \delta)$  for which  $x(t) = x(\tau) + r$  (for otherwise we would have  $\limsup_{t \downarrow \tau} (x(t) - x(\tau))/(t - \tau) \leq 0$ , contradicting the first condition of (2.6)). Pick  $s, t$ , and  $r$  as above. Then  $x(s) = x(t)$  and  $|s - t| < 2\varepsilon$ . Since  $\varepsilon$  is arbitrary, it follows that the left side of (2.1) is zero, a contradiction.  $\square$

The proof of the theorem will also show:

**COROLLARY 4.** *If  $x$  is merely assumed to be a strong local solution to (P) (i.e., relative to arcs  $y$  satisfying  $\|x - y\| < \varepsilon$  for some  $\varepsilon > 0$ ), then conclusions (i)–(iii) of the theorem remain valid for any  $\tau$  satisfying (2.1).*

**2.2. Some reductions to simpler cases.** We begin by asserting that there is no loss of generality in assuming that the function  $\theta$  in (H3) is nondecreasing on  $[0, \infty)$ . For suppose this is not initially the case. From the growth and convexity of  $\theta$  there exists  $r_0 \geq 0$  such that  $\theta$  is nondecreasing on  $[r_0, \infty)$ , and such that  $\max\{\theta(r) : 0 \leq r \leq r_0\} = \theta(r_0)$ . Define a new function  $\tilde{\theta}$  via

$$\tilde{\theta}(r) := \max\{\theta(r), \theta(r_0)\} - \theta(r_0) + \min\{\theta(s) : 0 \leq s \leq r_0\}.$$

Because pointwise maxima of convex functions are convex,  $\tilde{\theta}$  continues to satisfy the conditions of (H3) (note  $\tilde{\theta} \leq \theta$ ) and, in addition, is nondecreasing. We are therefore content to prove the theorem under the additional hypothesis:

$$(H4) \quad \theta \text{ is nondecreasing on } [0, \infty).$$

The existence of a solution  $x$  to (P) follows from existence results that are now standard (see for example [7, Theorem 4.1.3]); note that (H1) and (H2) play the role of guaranteeing that there is at least one admissible arc  $y$  for which the functional

$$J(y) := \int_a^b L(t, y(t), \dot{y}(t)) dt$$

is finite (e.g., take  $y(t) = A + (t - a)(B - A)/(b - a)$ ).

Let  $\tau$  be the point in the statement of Theorem 2.1. Then there are sequences  $s_i$  and  $t_i$  converging to  $\tau$ , with  $a \leq s_i \leq \tau \leq t_i \leq b$ ,  $s_i \neq t_i$ , such that

$$(2.7) \quad \lim_{i \rightarrow \infty} \frac{x(t_i) - x(s_i)}{t_i - s_i} =: z,$$

where  $z$  is a point in  $R^n$ . Note that if (for example) we have  $\tau > a$ , then we can arrange to have  $s_i < \tau$  while preserving (2.7) (since  $x$  is continuous); if  $\tau = a$  then of course every  $s_i$  equals  $\tau$ . A similar argument at  $b$  shows that we can assume that for each  $i$ ,  $[s_i, t_i]$  is a neighbourhood of  $\tau$  relative to  $[a, b]$ .

Our next step is to show that we need only consider the case  $z = 0$ . For suppose  $z$  is not zero. Define a new version ( $\bar{P}$ ) of (P) which has data

$$\bar{A} := A - az, \quad \bar{B} := B - bz, \quad \bar{L}(t, y, v) := L(t, y + tz, v + z).$$

Then a solution to ( $\bar{P}$ ) is provided by  $\tilde{x}(t) := x(t) - tz$ . Note that  $\bar{L}$  satisfies (H1) and (H2). We wish to show that for some suitable  $\tilde{\theta}$ , (H3) and (H4) are satisfied. We have

$$\bar{L}(t, y, v) \geq -\alpha|y + tz| + \theta(|v + z|) \geq -\alpha|y| - \alpha C|z| + \theta(|v + z|),$$

where  $C = \max\{|a|, |b|\}$ . We set

$$m := \min\{\theta(r) : 0 \leq r \leq 2|z|\},$$

and we define  $\tilde{\theta}$  by

$$(2.8) \quad \tilde{\theta}(r) := \theta(\max[r - |z|, 0]) - \max[\theta(0), 0] + \min[m, 0] - \alpha C|z|.$$

If  $|v| \geq |z|$ , then by (H4) and (2.8) one has

$$\theta(|v + z|) \geq \theta(|v| - |z|) \geq \tilde{\theta}(|v|) + \alpha C|z|.$$

If  $|v| \leq |z|$ , then  $|v + z| \leq 2|z|$ , whence

$$\theta(|v + z|) \geq m \geq \tilde{\theta}(|v|) + \alpha C|z|.$$

Combining these observations with the preceding we arrive at

$$\bar{L}(t, y, v) \geq -\alpha|y| + \tilde{\theta}(|v|),$$

and it follows that  $\bar{L}$  and  $\tilde{\theta}$  satisfy (H3) and (H4).

Note that the limit in (2.7), when  $x$  is replaced by  $\tilde{x}$ , is zero, and the conclusions of the theorem for  $\bar{L}, \tilde{\theta}$  imply those for  $L, x$ . This establishes that without loss of generality the theorem may be proven under the additional hypothesis:

- (H5) There are sequences  $s_i, t_i, i \geq 1$ , converging to  $\tau$  such that  $[s_i, t_i]$  is a neighbourhood of  $\tau$  in  $[a, b]$  for each  $i$  and

$$\lim_{i \rightarrow \infty} \frac{x(t_i) - x(s_i)}{t_i - s_i} = 0.$$

Let  $M$  be such that  $\|x\| < M$  (where  $\|\cdot\|$  signifies the supremum norm on  $[a, b]$ ), and define a new Lagrangian  $L_1$  by

$$(2.9) \quad L_1(t, y, v) := \max[L(t, y, v), -\alpha M + \theta(|v|)].$$

Observe that (H1)–(H4) continue to be satisfied (with  $\tilde{\theta}(r) := \theta(r) - \alpha M$  and  $\tilde{\alpha} = 0$ ), and (H5) is unaffected. We have  $L_1 \geq L$ , and it follows from (H3) that  $L_1(t, y, v)$  and  $L(t, y, v)$  coincide for  $y$  in a neighbourhood of  $x(t)$ . Thus,  $x$  remains a solution to the new problem, and it is easy to see that the conclusions of Theorem 1 for  $L_1$  imply the same for  $L$ . Because  $L_1$  satisfies (H3) with  $\alpha = 0$ , we deduce that no loss of generality results in proving the theorem under the additional hypothesis:

$$(H6) \quad \text{The } \alpha \text{ in (H3) is equal to zero.}$$

No change in hypotheses or conclusions results if a constant is added to  $L_1$  either. We can then add the same constant to  $\theta$ , so that no generality is lost in supposing that  $\theta$  is nonnegative. We can then replace  $\theta$  by  $\tilde{\theta}$ , where  $\tilde{\theta}$  is the convex hull of the function  $\theta(r)$  and the function  $r^2$  (i.e.,  $\tilde{\theta}$  is the greatest convex function majorized by each of the given functions, see [13]). This preserves all previous properties of  $\theta$ . Consequently, we can suppose  $0 \leq \theta(r) \leq r^2$  without loss of generality. One last property of  $\theta$  will prove of value: we wish  $\theta$  to be strictly convex for  $r$  sufficiently large. To effect this, construct a new  $\tilde{\theta}$  (majorized by the old) as follows:

$$(2.10) \quad \tilde{\theta}(r) := \theta(0) + \int_0^r \left[ \frac{1+s}{2+s} \right] \theta'(s) ds$$

(note that  $\theta$  is locally Lipschitz, so  $\theta'$  is defined, by [6, Proposition 2.2.6]). Then  $\tilde{\theta}'(r) = (1+r)\theta'(r)/(2+r)$  is strictly increasing as soon as  $\theta'$  is strictly positive, so that  $\tilde{\theta}$  is eventually strictly convex. It is simple to verify via (2.10) that  $\tilde{\theta}$  continues to satisfy all the previous properties listed for  $\theta$ . In summary, we have shown that there is no loss of generality in assuming:

$$(H7) \quad \text{The function } \theta \text{ in (H3) satisfies } 0 \leq \theta(r) \leq r^2 \text{ and is strictly convex for } r \text{ sufficiently large.}$$

We now turn to the proof of the theorem under the additional hypotheses (H4)–(H7), which have been shown to entail no loss of generality. We can summarize our extra baggage as follows:  $\theta(r)$  is nonnegative, nondecreasing, bounded above by  $r^2$ , and eventually strictly convex, the  $\alpha$  in (H3) is 0, and condition (H5) holds. We shall be glad to have these items for the next step.

2.3. *Construction of an auxiliary Lagrangian.* Let

$$S = \{(t, y) : a \leq t \leq b, |y| \leq M\},$$

where  $M$ , as earlier, satisfies  $\|x\| < M$ , and define

$$c_0 := \sup\{|L(t, y, 0)| : (t, y) \in S\},$$

a finite quantity by (H1). Now pick  $R_0 > 0$  such that

$$(2.11) \quad \theta(R_0) > 2[1 + c_0]$$



and  $\theta$  is strictly convex on  $[R_0, \infty)$ . (This is possible by (H7).) Let

$$\sigma_0 := \sup\{|\zeta| : \zeta \in \partial_v L(t, y, v), (t, y) \in S, |v| \leq R_0\}$$

( $\sigma_0$  is finite by (H2)), and choose  $R_1 > R_0$  such that

$$(2.12) \quad \theta(r) > 2r[1 + \sigma_0 + c_0/R_1] \quad \text{if } r \geq R_1.$$

Now let

$$c_1 := \sup\{|L(t, y, v)| : (t, y) \in S, |v| \leq R_1\},$$

$$\sigma_1 := \sup\{|\zeta| : \zeta \in \partial_v L(t, y, v), (t, y) \in S, |v| \leq R_1\},$$

and choose  $R_2 > R_1$  such that

$$(2.13) \quad \theta(R_2) > 2[c_1 + 2R_2\sigma_1].$$

Let

$$\phi(w) := \frac{1}{2} \max[\theta(|w|), \theta(R_2)].$$

We now define the function  $\tilde{L}(t, y, v)$ , for each  $(t, y)$ , as the convex hull of the functions  $u \rightarrow L(t, y, u)$  (restricted to  $|u| \leq R_2$ ) and  $w \rightarrow \phi(w)$ . Formally (see [13]),  $\tilde{L}(t, y, v)$  is defined as

$$\inf\{\lambda L(t, y, u) + (1 - \lambda)\phi(w) : 0 \leq \lambda \leq 1, |u| \leq R_2, \lambda u + (1 - \lambda)w = v\}.$$

PROPOSITION 2.1.  $\tilde{L}$  possesses the following properties:

- (a)  $\tilde{L}(t, y, v)$  is (globally) measurable in  $t$ , convex in  $v$ , continuous in  $(y, v)$ .
- (b) There exist constants  $k_1$  and  $k_2$  such that for all  $(t, y)$  in  $S$ , for all  $v$ , for all  $\zeta$  in  $\partial \tilde{L}(t, y, v)$ , one has  $|\zeta| \leq k_1 + k_2|v|$  (where  $\partial$  denotes generalized gradient in  $(y, v)$ ); in particular then,  $\tilde{L}$  is locally Lipschitz in  $(y, v)$ .
- (c)  $\tilde{L}(t, y, v) \geq \theta(|v|)/2$  for all  $(t, y, v)$ .
- (d) If  $|v| \leq R_1$  and  $(t, y) \in S$ , then in the inf defining  $\tilde{L}(t, y, v)$  we can limit attention to  $|w| \leq R_2$ .
- (e) For  $(t, y)$  in  $S$ , we have  $\tilde{L}(t, y, v) = L(t, y, v)$  if  $|v| \leq R_1$ ,  $\tilde{L}(t, y, v) \leq L(t, y, v)$  for  $|v| \leq R_2$ , and  $\tilde{L}(t, y, v) < L(t, y, v)$  for  $|v| > R_2$ .
- (f) For  $(t, y)$  in  $S$ , we have  $\tilde{L}(t, y, v) = \theta(|v|)/2$  if  $|v| > R_2$ .
- (g) If  $(t, y) \in S$  and  $|v| \geq R_1$ , then any element  $\zeta$  of  $\partial_v \tilde{L}(t, y, v)$  satisfies  $|\zeta| > 1 + \sigma_0$ .

PROOF.  $\tilde{L}$  is convex in  $v$ , by construction (see [13]), and measurable in  $t$ , since, for fixed  $(y, v)$ , the inf defining  $\tilde{L}(t, y, v)$  is equivalent to a pointwise inf of countably many measurable functions of  $t$  (obtained by taking  $(\lambda, u, w)$  in a suitable countable dense set). Thus (a) follows. (c) follows from the fact that both  $L$  and  $\phi$  satisfy the required inequality, and, hence, so does  $\tilde{L}$ , the convex hull.

Let us turn to (d). Let  $u, w, \lambda$  be such that  $\lambda u + (1 - \lambda)w = v$ , where  $|v| \leq R_1$ ,  $|u| \leq R_2$ , and  $|w| > R_2$ . It follows that  $\lambda > 0$  necessarily. There exists  $r > 0$  such that the point  $\tilde{w} := w - r(w - u)$  has norm  $R_2$ . Note that we can take  $0 < r < \lambda$ , since  $|v| \leq R_1 < R_2$ . Set  $\tilde{\lambda} := (\lambda - r)/(1 - r)$ , and note the relation  $\tilde{\lambda}u + (1 - \tilde{\lambda})\tilde{w} = v$ . We now claim that

$$(2.14) \quad \tilde{\lambda}L(t, y, u) + (1 - \tilde{\lambda})\phi(\tilde{w}) \leq \lambda L(t, y, u) + (1 - \lambda)\phi(w),$$

which will show that in the inf defining  $\tilde{L}$ , we gain nothing from taking  $|w| > R_2$ , as claimed. Now (2.14) can be rewritten as

$$\begin{aligned}\phi(\tilde{w}) &\leq (\lambda - \tilde{\lambda})/(1 - \tilde{\lambda})L(t, y, u) + (1 - \lambda)/(1 - \tilde{\lambda})\phi(w) \\ &= rL(t, y, u) + (1 - r)\theta(|w|)/2\end{aligned}$$

(since  $\phi(w) = \theta(|w|)/2$  by (H4)), so it would suffice to establish

$$\theta(|\tilde{w}|)/2 \leq r\theta(|u|)/2 + (1 - r)\theta(|w|)/2,$$

since  $\phi(\tilde{w}) = \theta(|\tilde{w}|)/2$  and  $L(t, y, u) \geq \theta(|\mu|) \geq \theta(|u|)/2$  (the last inequality in view of (H7)). But this last inequality is an immediate consequence of the convexity of the function  $y \rightarrow \theta(|y|)$ , since  $\tilde{w} = ru + (1 - r)w$  (we have used the fact that  $f \circ g$  is convex when  $f: R \rightarrow R$  is convex and nondecreasing, and  $g: R^n \rightarrow R$  is convex). Thus, (2.14) and, hence, (d) are proven.

Let us examine (e). Note first that  $\tilde{L}(t, y, v)$  is bounded above by  $L(t, y, v)$  if  $|v| \leq R_2$ , since we can take  $\lambda = 1$ ,  $u = v$  in the inf defining  $\tilde{L}$ . If  $|v| > R_2$  take  $\lambda = 0$ ,  $w = v$  to deduce

$$\tilde{L}(t, y, v) \leq \phi(|v|) = \theta(|v|)/2 < \theta(|v|) \leq L(t, y, v).$$

So in all cases,  $\tilde{L} \leq L$ , and the inequality is strict if  $|v| > R_2$ . Now let  $|v| \leq R_1$  and let  $\delta > 0$ . Then there exist  $\lambda, u, w$  as in the definition of  $\tilde{L}(t, y, v)$  and with  $|w| \leq R_2$ , in view of (d), such that

$$\begin{aligned}\tilde{L}(t, y, v) + \delta &\geq \lambda L(t, y, u) + (1 - \lambda)\phi(w) \\ &\geq \lambda\{L(t, y, v) + (u - v) \cdot \zeta\} + (1 - \lambda)\phi(w),\end{aligned}$$

where  $\zeta$  is any element of  $\partial_v L(t, y, v)$  (since  $\zeta$  is a subgradient as in convex analysis because  $L$  is convex in  $v$  (see [7, Proposition 2.27])),

$$\begin{aligned}&\geq \lambda L(t, y, v) - |u - v|\sigma_1 + (1 - \lambda)\theta(R_2)/2 \\ &\hspace{15em}(\text{by definition of } \sigma_1, \text{ and since } \phi \geq \theta(R_2)/2) \\ &= L(t, y, v) + (1 - \lambda)\{\theta(R_2)/2 - L(t, y, v)\} - |u - v|\sigma_1 \\ &\geq L(t, y, v) + (1 - \lambda)\{\theta(R_2)/2 - c_1\} - |u - v|\sigma_1 \quad (\text{by definition of } c_1) \\ &\geq L(t, y, v) + (1 - \lambda)\{\theta(R_2)/2 - c_1 - 2R_2\sigma_1\} \\ &\hspace{15em}(\text{since } |u - v| = (1 - \lambda)|u - w| \leq (1 - \lambda)2R_2) \\ &\geq L(t, y, v),\end{aligned}$$

in light of (2.13). Since  $\delta$  is arbitrary we deduce  $\tilde{L}(t, y, v) = L(t, y, v)$ , and (e) is proven.

We showed above that we have  $\tilde{L}(t, y, v) \leq \theta(|v|)/2$  whenever  $|v| > R_2$ . Equality must then hold for such  $v$  in view of (c), which is assertion (f).

Let us turn now to (b). Let  $K$  be a Lipschitz constant for  $y \rightarrow L(t, y, u)$ , uniformly valid for  $t$  in  $[a, b]$ ,  $|u| \leq R_2$ , and for  $y$  in the region  $|y| \leq M$ . Let  $y_1, y_2$  be points in that region, and for any  $\delta > 0$  and any  $v$ , let  $u, w$ , and  $\lambda$ , as in the definition of  $\tilde{L}(t, y, v)$ , be such that

$$\tilde{L}(t, y_1, v) + \delta \geq \lambda L(t, y_1, u) + (1 - \lambda)\phi(w).$$

Then

$$\begin{aligned} \tilde{L}(t, y_2, v) &\leq \lambda L(t, y_2, u) + (1 - \lambda)\phi(w) \\ &\leq \lambda[L(t, y_1, u) + K|y_1 - y_2|] + (1 - \lambda)\phi(w) \\ &\leq \tilde{L}(t, y_1, v) + K|y_1 - y_2| + \delta. \end{aligned}$$

Since  $y_1, y_2$  are interchangeable and  $\delta$  is arbitrary, it follows that  $\tilde{L}(t, \cdot, v)$  is Lipschitz on  $|y| \leq M$  with rank  $K$ . A simple argument also verifies that  $\tilde{L}(t, y, \cdot)$  is Lipschitz on  $|v| \leq R_2$  of rank  $K_1$  for some constant  $K_1$ . Thus  $\partial\tilde{L}(t, y, v)$  is uniformly bounded for  $(t, y) \in S, |v| \leq R_2$ . If  $|v| > R_2$ , then by (f),  $\tilde{L}(t, y, w) = \theta(|w|)/2$  for  $w$  near  $v$ , so the truth of (b) is evident if we are armed with the following:

LEMMA 2.1. *For every  $r > 0$  and every  $\zeta$  in  $\partial\theta(r)$ , one has  $|\zeta| \leq 4r$ .*

To see this, express the subdifferential inequality:  $\theta(s) \geq \theta(r) + \zeta(s - r)$  for all  $s$ . Put  $s = 2r$  and use (H7) to obtain  $\zeta r \leq 4r^2$ , whence the result.

Only (g) now remains to be proved. Let  $t, y, v$  and  $\zeta$  be as described there. Then

$$\tilde{L}(t, y, w) - \tilde{L}(t, y, v) \geq \langle w - v, \zeta \rangle \quad \text{for all } w,$$

whence, using (c) and (e),

$$\langle w - v, \zeta \rangle \leq L(t, y, w) - \theta(|v|)/2.$$

Put  $w = 0$  to derive

$$|\zeta| \geq [\theta(|v|)/2 - L(t, y, 0)]/|v| > 1 + \sigma_0 \quad (\text{by (2.12)}). \quad \square$$

2.4. *End of the proof.* Let us now consider the following variational problem  $(P_i)$ , where  $i$  is a positive integer: minimize

$$\tilde{J}_i := \int_{s_i}^{t_i} \tilde{L}(t, y, \dot{y}) dt$$

over the arcs  $y$  on  $[s_i, t_i]$  satisfying  $y(s_i) = A_i, y(t_i) = B_i$ , where we define  $A_i := x(s_i), B_i := x(t_i)$ . In view of the properties of  $\tilde{L}$  enumerated in Proposition 2.1, it is an easy consequence of standard existence theorems (e.g. [7, Theorem 4.1.3]) that  $(P_i)$  admits a solution, which we label  $x_i$ . The linear arc  $\hat{y}_i$ , defined by

$$\hat{y}_i(t) := A_i + (t - s_i)(B_i - A_i)/(t_i - s_i),$$

is feasible for  $(P_i)$ , whence  $\tilde{J}_i(x_i) \leq \tilde{J}_i(\hat{y}_i) \leq J_i(\hat{y}_i)$ . In view of Proposition 2.1(c), we derive

$$\frac{1}{2} \int_{s_i}^{t_i} \theta(|\dot{x}_i|) dt \leq J_i(\hat{y}_i).$$

Since the terms on the right go to zero as  $i$  goes to  $\infty$ , we conclude that, given any  $\varepsilon > 0$ , for  $i$  sufficiently large one has  $|x_i(t) - A_i| < \varepsilon$  on  $[s_i, t_i]$ . It follows that for some  $I_0$  and  $i \geq I_0$ , one has  $|x_i(t)| < M$ . The point of this observation is that near  $x_i$ , for  $i \geq I_0$ , the additional properties of  $\tilde{L}$  which hold for  $(t, y)$  in  $S$  become available to us.

We now wish to set the stage for applying necessary conditions to the solution  $x_i$  to  $(P_i)$ , namely those of [7, Theorem 4.2.2]. The two hypotheses to be verified

are “calmness” and the “strong Lipschitz condition” on the Hamiltonian  $H$  of the problem, i.e., the function defined by

$$H(t, y, p) := \sup\{\langle p, v \rangle - \tilde{L}(t, y, v) : v \in R^n\}.$$

Note that  $H$  is finite everywhere by Proposition 2.1. Let  $(t, y_1)$  and  $(t, y_2)$  in  $S$  be given. Then

$$\begin{aligned} H(t, y_1, p) &= \sup_v \{\langle p, v \rangle - \tilde{L}(t, y_1, v)\} \\ &\leq \sup_v \{\langle p, v \rangle - \tilde{L}(t, y_2, v)\} + \sup_v |\tilde{L}(t, y_2, v) - \tilde{L}(t, y_1, v)| \\ &= H(t, y_2, p) + \sup\{|\tilde{L}(t, y_2, v) - \tilde{L}(t, y_1, v)| : |v| \leq R_2\} \end{aligned}$$

(since  $\tilde{L}(t, y_2, v)$  and  $\tilde{L}(t, y_1, v)$  agree for  $|v| > R_2$  by Proposition 2.1)

$$\leq H(t, y_2, p) + k|y_1 - y_2|,$$

where  $k$  is a Lipschitz constant valid on the relevant (bounded) set. This verifies the strong Lipschitz condition. The argument proving calmness proceeds by showing that the problem  $(P_i)$  is equivalent to a certain penalized free endpoint problem; it is detailed in [7, Step 1, p. 183] (it is here that property (b) of Proposition 2.1 is used).

We derive then (for  $i \geq I_0$ ) the existence of an arc  $p_i$  on  $[s_i, t_i]$  such that

$$(2.15) \quad (-\dot{p}_i, \dot{x}_i) \in \partial H(t, x_i, p_i) \quad \text{a.e. on } [s_i, t_i],$$

where  $\partial H$  signifies the generalized gradient in the  $(y, p)$  variables of  $H(t, y, p)$ . Relation (2.15) implies

$$(2.16) \quad |\dot{p}_i(t)| \leq k,$$

where  $k$  is the Lipschitz constant just introduced above, as well as (see [7, Proposition 2.5.3])

$$(2.17) \quad \dot{x}_i \in \partial_p H(t, x_i, p_i) \quad \text{a.e.}$$

and (by the conjugacy of convex analysis)

$$(2.18) \quad p_i \in \partial_v \tilde{L}(t, x_i, \dot{x}_i) \quad \text{a.e.}$$

**LEMMA 2.2.** *There exist  $I_1 \geq I_0$  and  $R_3 > 0$  such that for all  $i \geq I_1$ , one has  $|\dot{x}_i(t)| \leq R_3$  a.e. in  $[s_i, t_i]$ .*

To see this, begin by choosing  $R_3 > R_2$  such that for all  $r \geq R_3$  and all  $s$  in  $\partial\theta(r)$ , one has  $s \geq 2[\sigma_2 + 2]$ , where  $\sigma_2$  is defined by

$$\sigma_2 := \sup\{|\zeta| : \zeta \in \partial_v \tilde{L}(t, y, v) : (t, y) \in S, |v| \leq R_2 + 1\}.$$

This is possible since  $\theta(r)/r \rightarrow \infty$  as  $r \rightarrow \infty$ . Now let us choose  $I_1 \geq I_0$  such that  $k(t_i - s_i) < 1$  for all  $i \geq I_1$ , and such that  $|\Delta_i| < R_2$  for  $i \geq I_1$ , where  $\Delta_i := (B_i - A_i)/(t_i - s_i)$  (recall (H5)).

Suppose that for such an  $i$  we fail to have  $|\dot{x}_i| \leq R_3$  a.e. Then  $p_i(t) \in \partial\theta(r)/2$  for at least one  $t$  (by (2.18) and Proposition 2.1(f)) and for  $r > R_3$ , whence  $|p_i(t)| \geq \sigma_2 + 2$ . In view of (2.16) and the choice of  $I_1$ , we have

$$|p_i| \geq \sigma_2 + 1 \quad \text{for all } t \text{ in } [s_i, t_i],$$

so that  $|\dot{x}_i| \geq R_2 + 1$  for all  $t$ , by (2.18) and by definition of  $\sigma_2$ . But then for each  $t$  there is a neighbourhood of  $(x_i(t), p_i(t))$  in which  $H(t, \cdot, \cdot)$  is given by

$$H(t, y, p) = \sup_v \{ \langle p, v \rangle - \theta(|v|)/2 \},$$

in view of Proposition 2.1(f). From (2.15) we conclude that  $\dot{p}_i = 0$  a.e., so that  $p_i$  is constant. Since  $\theta(r)$  is strictly convex for  $r > R_0$  (see (H7)), (2.18) implies that  $\dot{x}_i$  is constant, in fact, by the above, a constant of norm exceeding  $R_2$ . This implies

$$|\Delta_i| := |(x_i(t_i) - x_i(s_i))/(t_i - s_i)| > R_2,$$

a contradiction which establishes the lemma.

The proof also established

$$(2.19) \quad |p_i(t)| \leq \sigma_2 + 2 \quad \text{for } t \text{ in } [s_i, t_i].$$

LEMMA 2.3. *There exists  $I_2$  such that, for  $i \geq I_2$ , there is a subset of  $[s_i, t_i]$  of positive measure in which one has  $|\dot{x}_i| < R_0$ .*

Choose  $I_2 \geq I_1$  such that, for all  $i \geq I_2$ ,

$$(2.20) \quad kR_3(t_i - s_i) + (\sigma_2 + 2)|\Delta_i| < 1$$

(recall that  $\Delta_i \rightarrow 0$ ). Suppose the assertion of the lemma is false for such an  $i$ . Then  $|\dot{x}_i| \geq R_0$  a.e. on  $[s_i, t_i]$ . Because  $p_i(t)$  is a subgradient a.e. of the convex function  $v \rightarrow \tilde{L}(t, x_i(t), v)$  at  $v = \dot{x}_i(t)$ , one has, for almost all  $t$ ,

$$\tilde{L}(t, x_i(t), 0) - \tilde{L}(t, x_i(t), \dot{x}_i(t)) \geq -\langle \dot{x}_i(t), p_i(t) \rangle.$$

Now the left side is bounded above by  $L(t, x_i(t), 0) - \theta(|\dot{x}_i|)/2$ , which is in turn majorized by  $c_0 - \theta(R_0)/2$ , whence

$$(2.21) \quad \langle \dot{x}_i(t), p_i(t) \rangle \geq \theta(R_0)/2 - c_0 > 1 \quad \text{a.e.}$$

by (2.11). Now we have

$$|p_i(t) - p_i(s_i)| \leq k|t - s_i| \leq k(t_i - s_i)$$

by (2.16), so

$$k(t_i - s_i)|\dot{x}_i(t)| + \langle \dot{x}_i(t), p_i(s_i) \rangle > 1 \quad \text{a.e.}$$

Recalling Lemma 2.2, integrating this over  $[s_i, t_i]$  and invoking (2.19) yield

$$kR_3(t_i - s_i)^2 + (\sigma_2 + 2)|\Delta_i|(t_i - s_i) > (t_i - s_i),$$

contradicting (2.20) and proving the lemma.

LEMMA 2.4. *There exists  $I_3$  such that, for  $i \geq I_3$ , one has  $|\dot{x}_i(t)| \leq R_1$  a.e. in  $[s_i, t_i]$ .*

Pick  $I_3 \geq I_2$  such that  $k(t_i - s_i) < 1$  for  $i \geq I_3$ . By Lemma 2.3 and (2.18), there is at least one  $t$  in  $[s_i, t_i]$  for which  $|p_i(t)| \leq \sigma_0$ , so that  $|p_i| < \sigma_0 + 1$  throughout  $[s_i, t_i]$ . It is now a direct consequence of (2.18) and Proposition 2.1(g) that necessarily  $|\dot{x}_i| \leq R_1$  a.e., establishing the lemma.

We now observe that for  $i \geq I_3$  one has

$$\begin{aligned} \tilde{J}_i(x_i) &\leq \tilde{J}_i(x) && \text{(since } x_i \text{ solves } (P_i)) \\ &\leq J_i(x) && \text{(since } \tilde{L} \leq L) \\ &\leq J_i(x_i) && \text{(since } x \text{ solves } (P)) \\ &= \tilde{J}_i(x_i) && \text{(by Lemma 2.4 and Proposition 2.1(e)).} \end{aligned}$$

It follows that  $\tilde{J}_i(x)$  and  $J_i(x)$  agree, whence  $|\dot{x}| \leq R_2$  on  $[s_i, t_i]$  by Proposition 2.1(e). When  $\dot{x}$  is bounded, the Euler–Lagrange inclusion is a known necessary condition [4] in this context. This proves the first assertions of the theorem (take  $I = [s_i, t_i]$ ).

We now posit the additional hypotheses of (ii). Pick any  $R > R_2$  and consider the following problem for any integer  $i > I_3$  as above: minimize  $\int_{s_i}^{t_i} L(t, y, \dot{y}) dt$  over the arcs  $y$  satisfying  $|\dot{y}| \leq R$  a.e.,  $y(s_i) = A_i$ ,  $y(t_i) = B_i$ . We know that  $x$  solves this problem and (because of the compact set to which the velocity is constrained) the necessary conditions [7, Theorem 4.2.2] for this optimal control problem apply (for  $R > |\Delta_i|$ ; the arguments that verify the requisite hypotheses are precisely those of [7, Proposition 4.2.4]). The Hamiltonian  $H$  for this problem is given by

$$H(t, y, p) = \sup\{\langle p, v \rangle - L(t, y, v) : |v| \leq R\}.$$

The necessary conditions assert the existence of an arc  $p$  on  $[s_i, t_i]$  such that  $(-\dot{p}, \dot{x}) \in \partial H(t, x, p)$  a.e., which implies [7, Proposition 2.5.3]

$$(2.22) \quad \dot{x}(t) \in \partial_p H(t, x(t), p(t)) \quad \text{a.e.}$$

The elements of  $\partial_p H(t, x(t), p(t))$  are those  $v$  maximizing  $v \rightarrow \langle p(t), v \rangle - L(t, x(t), v)$  over  $|v| \leq R$ . Thus the maximum is attained a.e. at  $\dot{x}(t)$ .

If the bounded function  $\dot{x}$  has other than removable discontinuities, then we can find two sequences  $c_j, d_j$  in  $[s_i, t_i]$  at each point of which (2.22) holds, converging to a point  $r$ , and such that

$$\lim_{j \rightarrow \infty} \dot{x}(c_j) =: \alpha \neq \beta := \lim_{j \rightarrow \infty} \dot{x}(d_j).$$

Let  $v$  be any point such that  $|v| \leq R$ . Then by the preceding remarks we have

$$(2.23) \quad \langle p(c_j), v \rangle - L(c_j, x(c_j), v) \leq \langle p(c_j), \dot{x}(c_j) \rangle - L(c_j, x(c_j), \dot{x}(c_j)).$$

Because  $L(t, y, v)$  is Lipschitz in  $(y, v)$ , the sequence  $L(c_j, x(c_j), \dot{x}(c_j))$  has the same limit as the sequence  $L(c_j, x(r), \alpha)$ , namely  $L(r, x(r), \alpha)$ . Similarly,  $L(c_j, x(c_j), v)$  converges to  $L(r, x(r), v)$ . Taking limits in (2.23) gives

$$\langle p(r), v \rangle - L(r, x(r), v) \leq \langle p(r), \alpha \rangle - L(r, x(r), \alpha).$$

Since  $v$  is arbitrary, the point  $\alpha$  maximizes the strictly concave function in question over  $|v| \leq R$ . But  $|\alpha| < R$  (recall  $|\dot{x}| \leq R_2 < R$ ), so  $\alpha$  globally maximizes the strictly concave function  $v \rightarrow \langle p(r), v \rangle - L(r, x(r), v)$  (local and global maxima of concave functions coincide). The same conclusion obtains for  $\beta$  (distinct from  $\alpha$ ), the required contradiction. Hence,  $\dot{x}$  is essentially continuous on  $[s_i, t_i]$ .

And now the third and final part of the theorem. We merely sketch an argument to derive this classical result, which goes back to Weierstrass (see for example [3,

Theorem 2.6.iii)]. Write the Euler–Lagrange inclusion (which now reduces to an equation) as

$$(2.24) \quad L_v(t, x, \dot{x}) = p_0 + \int_{s_i}^t L_x(s, x, \dot{x}) ds.$$

The right side is  $C^1$  and so is the function  $(t, v) \rightarrow L_v(t, x(t), v)$ . It follows from the Implicit Function Theorem that  $\dot{x}$  is  $C^1$  on  $I$ ; i.e., that  $x$  is  $C^2$  on  $I$ . This allows us to differentiate through in (2.24) and solve to get

$$(2.25) \quad \ddot{x} = [L_{vv}]^{-1}\{L_x - L_{vt} - L_{vx}\dot{x}\},$$

where the derivatives of  $L$  are evaluated at  $(t, x(t), \dot{x}(t))$ . (It follows that  $\ddot{x}$  admits finite limiting values at  $s_i$  and  $t_i$ .) Now suppose  $x$  is known to be  $C^{r-1}$  and  $L$  is  $C^r$  ( $r \geq 2$ ). The right side of (2.25) is readily seen to be  $C^{r-2}$ , hence so is  $\ddot{x}$ ; i.e.,  $x$  is  $C^r$  on  $I$ .  $\square$

**3. Hypotheses which restrict the set of points of bad behaviour.**

Throughout this section (H1)–(H3) are in force. We take  $x(\cdot)$  to be a solution to problem (P) and  $\Omega$  to be the subset of  $[a, b]$  in which  $x$  is locally Lipschitz.

We have already shown (§2) that the set of points of bad behaviour, namely points at which  $x$  is not locally Lipschitz, is a closed set of zero measure. Various additional hypotheses are now supplied under which it is possible to infer further properties about the set; in certain cases, points of bad behaviour cannot exist.

In view of the results of §2, hypotheses which assure that  $\Omega = [a, b]$ , when supplemented by conditions under which the stronger conclusions of Theorem 2.1 apply, immediately translate into hypotheses under which  $x(\cdot)$  is globally  $C^1$  or  $C^2$ , etc.

The proof of a number of our results requires the following lemma:

LEMMA 3.1. *Let  $\bar{t} \in \Omega \cap [a, b]$  be given and define*

$$(3.1) \quad t_{\max} = \sup \left\{ t \in (\bar{t}, b] : \text{ess sup}_{s \in [\bar{t}, t]} |\dot{x}(s)| < \infty \right\}.$$

*Then under either hypothesis (i) or (ii) below the restriction of  $x(\cdot)$  to  $[\bar{t}, b]$  is Lipschitz.*

(i)  *$|\dot{x}(\cdot)|$  is essentially bounded on  $[\bar{t}, t_{\max}]$ .*

(ii) *Corresponding to any increasing sequence of points  $\{t_i\}$  in  $[\bar{t}, t_{\max})$  which converges to  $t_{\max}$ , there exists an absolutely continuous function  $p_i(\cdot) : [\bar{t}, t_i] \rightarrow R^n$  and a constant  $k$  independent of  $i$ , such that*

$$(3.2) \quad p_i(t) \in \partial_v L(t, x(t), \dot{x}(t)) \quad \text{a.e. } t \in [\bar{t}, t_i],$$

and

$$(3.3) \quad |p_i(t)| \leq k, \quad \text{all } t \in [\bar{t}, t_i].$$

PROOF. Notice first that  $t_{\max}$  is a well-defined number in  $(\bar{t}, b]$  since  $\bar{t} \in \Omega \cap [a, b]$ , and, therefore,  $\dot{x}(\cdot)$  is essentially bounded on a neighbourhood of  $\bar{t}$  relative to  $[a, b]$ .

Suppose first (i). By Theorem 2.1  $t_{\max} \in \Omega$ , but then we must have  $t_{\max} = b$ , i.e.  $x(\cdot)$  is Lipschitz on  $[\bar{t}, b]$ , for if  $t_{\max}$  were interior to  $[\bar{t}, b]$  then, again by Theorem

2.1,  $\dot{x}(\cdot)$  would be essentially bounded on a neighbourhood of  $t_{\max}$ , in contradiction of the definition of  $t_{\max}$ .

Suppose next (ii). Since  $L$  is convex in the  $v$  variable, (3.2) implies

$$L(t, x(t), v) - L(t, x(t), \dot{x}(t)) \geq p_i(t) \cdot (v - \dot{x}(t)) \quad \text{a.e. } t \in [\bar{t}, t_i].$$

Fix  $v \in R^n$ . By (3.3)

$$(3.4) \quad L(t, x(t), \dot{x}(t)) \leq k|\dot{x}(t)| + k_1 \quad \text{a.e. } t \in [\bar{t}, t_i]$$

for  $k$  as in (3.3) and  $k_1$  a constant which does not depend on  $i$ . By hypothesis (H3), however, we can choose a constant  $K > k_1$  such that

$$|w|^{-1}L(t, x(t), w) \geq k + 1 \quad \text{for all } t \in [a, b]$$

whenever  $|w| > K$ . Now consider the set

$$S = \{t \in [\bar{t}, t_i]: \dot{x}(t) \text{ exists and } |\dot{x}(t)| > K\}.$$

For any  $t \in S$  we have

$$L(t, x(t), \dot{x}(t)) > k|\dot{x}(t)| + |\dot{x}(t)| \geq k|\dot{x}(t)| + k_1.$$

Comparing this inequality with (3.4), we see that  $S$  is a null-set. It follows that

$$|\dot{x}(t)| \leq K \quad \text{a.e. } t \in [\bar{t}, t_i].$$

Bearing in mind that  $K$  does not depend on  $i$ , we see that (i) is confirmed. We have shown that this yields  $[\bar{t}, b] \subset \Omega$ .

3.1. *The autonomous case and extensions.* Perhaps the most noteworthy special case of (P) when points of bad behaviour cannot occur is when  $L$  does not depend on  $t$ . This is an immediate corollary of a more general result; here conditions are imposed on the time derivative of  $L$ , appropriately defined, in order that points of bad behaviour be confined to one endpoint or the other of  $[a, b]$ . These conditions are automatically satisfied in the autonomous case.

At this stage some new notation is required:  $\pi_t \partial_{t,x} L(t, x, v)$  denotes the projection onto the  $t$ -coordinate of the generalized gradient of  $(t, x) \rightarrow L(t, x, v)$  for fixed  $v$ . This reduces to the partial derivative in the  $t$ -variable  $L_t(t, x, v)$  when  $L(t, x, v)$  is continuously differentiable in  $(t, x)$ . Also, we point out in connection with inclusion (3.7) that, given a subset  $Q \subset R$ , we write  $|Q|$  for the set  $\{|q|: q \in Q\}$ .

Recall that here, and elsewhere in this section,  $x(\cdot)$  is a solution to (P).

PROPOSITION 3.1. *Suppose that  $L(t, x, v)$  is Lipschitz in  $(t, x, v)$  on bounded sets. Suppose further that there exists a summable function  $\gamma(\cdot)$  such that*

$$(3.5) \quad \gamma(s) \notin \pi_t \partial_{t,x} L(s, x(s), \dot{x}(s)) + [0, \infty) \quad \text{a.e. } s \in [a, b];$$

then  $(a, b] \subset \Omega$ .

If (3.5) is replaced by

$$(3.6) \quad \gamma(s) \notin \pi_t \partial_{t,x} L(s, x(s), \dot{x}(s)) - [0, \infty) \quad \text{a.e. } s \in [a, b],$$

then  $[a, b) \subset \Omega$ .

If (3.5) is replaced by

$$(3.7) \quad \gamma(s) \notin |\pi_t \partial_{t,x} L(s, x(s), \dot{x}(s))| - [0, \infty) \quad \text{a.e. } s \in [a, b],$$

then  $\Omega = [a, b]$ .



Note that, when  $L(t, x, v)$  is continuously differentiable in  $(t, x)$ , (3.5)–(3.7) reduce to

$$\gamma(s) \leq L_t(s, x(s), \dot{x}(s)), \quad \gamma(s) \geq L_t(s, x(s), \dot{x}(s)), \quad \gamma(s) \geq |L_t(s, x(s), \dot{x}(s))|,$$

respectively.

**PROOF OF PROPOSITION 3.1.** Suppose that (3.5) is true. We show that  $(a, b) \subset \Omega$ . Choose an arbitrary point  $\bar{t}$  from  $\Omega \cap [a, b)$ . Define  $t_{\max}$  by (3.1) and let  $\{t_i\}$  be an increasing sequence of points in  $(\bar{t}, t_{\max})$  which converges to  $t_{\max}$ .

Fix  $i$ . Note that the restriction of  $x(\cdot)$  to  $[\bar{t}, t_i]$  solves (P) when  $(x(\bar{t}), \bar{t})$ ,  $(x(t_i), t_i)$  replace  $(A, a)$ ,  $(B, b)$  as boundary data. By definition of  $t_{\max}$ ,  $|\dot{x}(\cdot)|$  is essentially bounded on  $[\bar{t}, t_i]$ . In view of these observations it is not difficult to deduce from [7, Theorem 5.2.3] the following facts concerning  $x(\cdot)$ :

There exist an absolutely continuous function  $p_i(\cdot): [\bar{t}, t_i] \rightarrow R^n$  and a constant  $c_i$  such that

$$(3.8) \quad p_i(t) \in \partial_v L(t, x(t), \dot{x}(t)) \quad \text{a.e. } t \in [\bar{t}, t_i]$$

and

$$(3.9) \quad L(t, x(t), \dot{x}(t)) - p_i(t) \cdot \dot{x}(t) = c_i + \int_{\bar{t}}^t \xi_i(s) ds \quad \text{a.e. } t \in [\bar{t}, t_i]$$

for some summable function  $\xi_i(\cdot)$  such that

$$\xi_i(t) \in \pi_t \partial_{t,x} L(t, x(t), \dot{x}(t)) \quad \text{a.e. } t \in [\bar{t}, t_i]$$

(to be more precise, these conclusions are obtained by applying [7, Theorem 5.2.3] to the autonomous optimal control problem in which time is treated as a dependent variable governed by the differential equation  $\dot{t} = 1$  on the interval  $[\bar{t}, t_i]$  and subject to the constraints  $t(\bar{t}) = 0$ ,  $t(t_i) \in R$ ).

Now since  $\bar{t} \in \Omega$ ,  $|\dot{x}(\cdot)|$  is bounded (off a null set) on some neighbourhood of  $\bar{t}$ . It follows from (3.8), continuity of  $p_i(\cdot)$  and the local Lipschitz continuity of  $L$  that  $|p_i(\bar{t})|$  is bounded by some constant independent of  $i$ . But then  $c_i$ , which, by (3.9), can be expressed

$$c_i = \lim_{t \downarrow \bar{t}} \left\{ \text{ess sup}_{s \in [\bar{t}, t]} [L(s, x(s), \dot{x}(s)) - p_i(s) \cdot \dot{x}(s)] \right\},$$

satisfies

$$(3.10) \quad |c_i| \leq k_1,$$

where  $k_1$  is a constant which does not depend on  $i$ .

Since  $L$  is convex in the velocity variable, (3.8) implies

$$L(t, x(t), v) - L(t, x(t), \dot{x}(t)) \geq p_i(t) \cdot (v - \dot{x}(t)), \quad \text{a.e. } t \in [\bar{t}, t_i],$$

for arbitrary  $v \in R^n$ . Combination of this inequality with (3.9), for arbitrary  $v$ , yields

$$\begin{aligned} L(t, x(t), v) - p_i(t) \cdot v &\geq c_i + \int_{\bar{t}}^t \xi_i(s) ds, & t \in [\bar{t}, t_i], \\ &\geq c_i + \int_{\bar{t}}^t \gamma(s) ds, & t \in [\bar{t}, t_i], \end{aligned}$$

where  $\gamma(\cdot)$  is the summable function of hypothesis (3.5). In view of (3.10)

$$L(t, x(t), v) - p_i(t) \cdot v \geq k_2, \quad t \in [\bar{t}, t_i], \quad v \in R^n,$$

for some constant  $k_2$  which does not depend on  $i$ . It follows that

$$|p_i(t)| \leq k_3, \quad \text{all } t \in [\bar{t}, t_i],$$

for  $i = 1, 2, \dots$ , where  $k_3$  is the constant

$$k_3 = \sup_{\substack{|v| \leq 1 \\ t \in [a, b]}} |L(t, x(t), v)| + |k_2|.$$

In view of (3.8), we have verified the hypotheses of Lemma 3.1. We conclude that the restriction of  $x(\cdot)$  to  $[\bar{t}, b]$  is Lipschitz. By Theorem 2.1,  $[\bar{t}, b] \subset \Omega$ .

The above argument applies for arbitrary  $\bar{t} \in \Omega \cap [a, b)$ . However,  $\bar{t}$  can be chosen arbitrarily close to the point  $a$  since  $\Omega$  has full measure. It follows that  $(a, b] \subset \Omega$ .

A similar argument yields  $[a, b) \subset \Omega$  under assumption (3.6); the difference is that we must consider a maximal interval  $[t_{\min}, \bar{t}]$  such that  $|\dot{x}(\cdot)|$  is essentially bounded on  $[t_{\min} + \varepsilon, \bar{t}]$  for arbitrary  $\varepsilon > 0$  to the left of an arbitrary point  $\bar{t} \in \Omega \cap (a, b]$ .

Observe finally that if (3.7) is true then certainly (3.5) and (3.6) are true, so  $\Omega = [a, b]$ .

**COROLLARY 3.1.** *Suppose  $L$  is independent of  $t$ . Then  $\Omega = [a, b]$ .*

Even in the smooth, scalar case Corollary 3.1 is, apparently, a new result. In this context Tonelli has shown [15, p. 366] that the hypothesis “ $L$  is independent of  $t$ ”, supplemented by the hypothesis: “there exist positive constants  $\alpha, M, M_1$  such that

$$|L - vL_v| \geq M|v|^\alpha - M_1 \quad \text{for all } (t, x, v) \in [a, b] \times R \times R^n,$$

assures that  $\Omega = [a, b]$ . We have shown that (in the presence of (H1)–(H3)) the supplementary hypothesis is superfluous.

The hypotheses of Proposition 3.1 can be replaced by others of less generality but of a more directly verifiable nature. As an illustration we note:

**COROLLARY 3.2.** *Suppose  $L(t, x, v)$  is Lipschitz in  $(t, x, v)$  on bounded sets, and, corresponding to any bounded open set  $S \subset R^n$ , there exist a constant  $c$  and a summable function  $\gamma(\cdot)$  such that*

$$(3.11) \quad L_t(t, x, v) \leq c|L(t, x, v)| + \gamma(t)$$

for all points  $(t, x, v)$  in  $(a, b) \times S \times R^n$  at which  $(\sigma, y) \rightarrow L(\sigma, y, v)$  is differentiable. Then  $[a, b] \subset \Omega$ .

If inequality (3.11) is replaced by  $-L_t(t, x, v) \leq c|L(t, x, v)| + \gamma(t)$ , then  $(a, b) \subset \Omega$ ; if it is replaced by  $|L_t(t, x, v)| \leq c_1|L(t, x, v)| + \gamma(t)$ , then  $\Omega = [a, b]$ .

**PROOF.** Choose  $S$  sufficiently large to contain the values of  $x(\cdot)$ . In view of the characterization of the generalized gradient provided in [7, Theorem 2.5.1], (3.11) implies

$$c|L(t, x, v)| + \gamma(t) \notin \pi_t \partial_{t,x} L(t, x, v) - [0, \infty)$$

for all points in  $(a, b) \times S \times R^n$ . But  $t \rightarrow L(t, x(t), \dot{x}(t))$  is summable by the properties of the arc  $x(\cdot)$ . Replacing  $(t, x, v)$  by  $(t, x(t), \dot{x}(t))$  wherever  $\dot{x}(t)$  is defined

and bearing in mind  $x(t) \in S$ , we see that (3.6) is satisfied. By Proposition 3.1,  $[a, b] \subset \Omega$ . The other cases are treated in similar fashion.

**3.2. Conditions of Morrey type.** Among hypotheses which assure existence of a solution to (P) which is smooth and satisfies the Euler–Lagrange equation, those which involve some kind of summability assumptions on the derivatives of  $L$ , both in  $x$  and  $v$  along the minimizing arc, have received the most attention. A simple form of such hypotheses, applicable to the smooth, vector case, is given in Morrey’s book [12, p. 28]. See also [10, p. 379 and 7, Theorem 4.4.1]. In the smooth, scalar case Tonelli [15, p. 368] showed that summability requirements on  $L_v$  can be dispensed with.

In a nonsmooth, vector setting we now give hypotheses in which  $L_v$  (strictly speaking the partial generalized gradient) is unrestricted, and in which the summability requirement on  $L_x$  is weakened by inclusion of a term which involves  $L_v$  in the bound.

**PROPOSITION 3.2.** *Suppose there exist a summable function  $\gamma(\cdot)$  and a non-negative constant  $c$  such that*

$$(3.12) \quad \gamma(s) \notin |\partial_x L(s, x(s), \dot{x}(s))| - c|\partial_v L(s, x(s), \dot{x}(s))| - [0, \infty) \quad \text{a.e. } s \in [a, b].$$

Then  $\Omega = [a, b]$ .

In inclusion (3.12),  $|\partial_x L(s, x(s), \dot{x}(s))|$  denotes the set  $\{|d|: d \in R^n, d \in \partial_x L(s, x(s), \dot{x}(s))\}$ , and  $|\partial_v L(s, x(s), \dot{x}(s))|$  has a similar interpretation. Condition (3.12) reduces to

$$|L_x(s, x(s), \dot{x}(s))| \leq c|L_v(s, x(s), \dot{x}(s))| + \gamma(s) \quad \text{a.e. } s \in [a, b]$$

when  $L(t, x, v)$  is continuously differentiable in  $x$  and  $v$ .

**PROOF.** Choose an arbitrary point  $\bar{t}$  from  $\Omega \cap [a, b]$ , define  $t_{\max}$  by (3.1), and let  $t_i$  be an increasing sequence of points in  $(\bar{t}, t_{\max})$  which converges to  $t_{\max}$ .

Fix  $i$ . The subarc  $x(s), \bar{t} \leq s \leq t_i$ , solves (P) for appropriately modified boundary data, and  $\dot{x}(t)$  is essentially bounded on  $[\bar{t}, t_i]$ . From [7, Theorem 5.2.1] we conclude:

There exists an absolute continuous function  $p_i(\cdot): [\bar{t}, t_i] \rightarrow R^n$  such that

$$(3.13) \quad \dot{p}_i(t) \in \partial_x L(t, x(t), \dot{x}(t)) \quad \text{a.e. } t \in [\bar{t}, t_i]$$

and

$$(3.14) \quad p_i(t) \in \partial_v L(t, x(t), \dot{x}(t)) \quad \text{a.e. } t \in [\bar{t}, t_i].$$

Since  $\dot{x}(t)$  is essentially bounded on a neighbourhood of  $\bar{t}$ , we deduce from (3.13) and the continuity of  $p_i(\cdot)$  that  $|p_i(\bar{t})| \leq k_1$  for some constant  $k_1$  which does not depend on  $i$ .

In view of hypothesis (3.12)

$$|\dot{p}_i(t)| - c|p(t)| \leq \gamma(t) \quad \text{a.e. } t \in [\bar{t}, t_i]$$

for a summable function  $\gamma(\cdot)$ . In view of these inequalities, we deduce from Gronwall’s inequality that

$$|p_i(t)| \leq k_2, \quad t \in [\bar{t}, t_i],$$

for some constant  $k_2$  which does not depend on  $i$ .

Noting (3.14), we see that the hypotheses of Lemma 3.1 are satisfied. In consequence,  $[\bar{t}, b] \subset \Omega$ . However,  $\bar{t}$  can be chosen arbitrarily close to the point  $a$ , whence  $(a, b) \subset \Omega$ .

A similar argument applied to a "maximal" interval to the left of an arbitrary point  $\bar{t} \in \Omega \cap (a, b]$  yields the conclusion  $[a, b) \subset \Omega$ . It follows that  $[a, b] \subset \Omega$ .

Essentially the same arguments which permitted us to deduce Corollary 3.2 from Proposition 3.1 lead to:

**COROLLARY 3.3.** *Suppose that, corresponding to any bounded, open set  $S \in R^n$ , there exist constants  $c_1, c_2$  and a summable function  $\gamma(\cdot)$  such that*

$$|L_x(t, x, v)| \leq c_1 |L(t, x, v)| + c_2 |L_v(t, x, v)| + \gamma(t)$$

for all points  $(t, x, v)$  in  $(a, b) \times S \times R^n$  at which  $(y, v) \rightarrow L(t, y, v)$  is differentiable. Then  $\Omega = [a, b]$ .

**3.3. Conditions of Bernstein type.** In [1, p. 433] Bernstein gave conditions under which the Euler-Lagrange equation admits a global solution on  $[a, b]$ . The main feature of Bernstein's conditions is the requirement that  $F$  has growth in  $v$  of degree at most quadratic, uniformly over bounded sets in the  $(t, x)$  variables. Here  $F$  signifies the "Bernstein function":

$$(3.15) \quad F = L_{vv}^{-1}(L_x - L_{vt} - L_{vx}v),$$

which is the function resulting from expressing the Euler-Lagrange equation in the form  $\ddot{x}(t) = F(t, x(t), \dot{x}(t))$ .

Tonelli [15, p. 363] showed in the scalar case that Bernstein's condition is sufficient to exclude points of bad behaviour. It is evident from the corollary to the next result that the same is true in a vector setting, and, further, that we can permit growth in  $v$  more than quadratic when additional hypotheses are placed upon  $L$ .

**PROPOSITION 3.3.** *Suppose  $L(t, x, v)$  is twice continuously differentiable in  $(t, x, v)$ ,  $L_{vv}(t, x(t), \dot{x}(t)) > 0$  a.e.  $t \in [a, b]$ , and there exists a summable function  $\gamma(\cdot)$  such that*

$$|F(t, x(t), \dot{x}(t))| \leq \gamma(t)(|\dot{x}(t)| + 1) \quad \text{a.e. } t \in [a, b],$$

where  $F$  was defined in (3.15). Then  $\Omega = [a, b]$ .

**PROOF.** Choose any  $\bar{t} \in \Omega \cap (a, b)$ . Define  $t_{\max}$  as in (3.1) and consider a sequence  $\{t_i\}$  such that  $t_i \uparrow t_{\max}$  as before. Fix  $i$ . According to Theorem 2.1,  $x(\cdot)$  is twice continuously differentiable at all points in  $[\bar{t}, t_i]$ . The restriction of  $x(\cdot)$  to  $[\bar{t}, t_i]$  solves (P) for suitably modified boundary conditions and therefore satisfies the Euler-Lagrange equation there:

$$\frac{d}{dt} L_v(t, x(t), \dot{x}(t)) = L_x(t, x(t), \dot{x}(t)), \quad t \in [\bar{t}, t_i].$$

This can be expressed

$$\ddot{x}(t) = F(t, x(t), \dot{x}(t)), \quad t \in [\bar{t}, t_i],$$

whence, for any  $t, s \in [\bar{t}, t_{\max}]$ ,

$$\dot{x}(t) - \dot{x}(s) = \int_s^t F(\sigma, x(\sigma), \dot{x}(\sigma)) d\sigma.$$

Under the hypotheses, then,

$$\begin{aligned} ||\dot{x}(t) - \dot{x}(s)|| &\leq \int_s^t |F(\sigma, x(\sigma), \dot{x}(\sigma))| d\sigma \\ &\leq c_1 \int_s^t \gamma(\sigma)(|\dot{x}(\sigma)| + 1) d\sigma. \end{aligned}$$

We conclude that the absolutely continuous function  $|\dot{x}(\cdot)|$  satisfies

$$\frac{d}{dt}|\dot{x}(t)| \leq c_1\gamma(t)[|\dot{x}(t)| + 1] \quad \text{a.e. } t \in [\bar{t}, t_i].$$

By Gronwall's inequality

$$|\dot{x}(t)| \leq k_1, \quad t \in [\bar{t}, t_i],$$

for some constant  $k_1$  which does not depend on  $i$ .

By Lemma 3.1 then,  $x(\cdot)$  is Lipschitz on  $[\bar{t}, b]$ . Likewise, we show that  $x(\cdot)$  is Lipschitz on  $[a, \bar{t}]$ . It follows that  $x(\cdot)$  is Lipschitz on all of  $[a, b]$  and, therefore,  $\Omega = [a, b]$ .

Given  $\alpha \geq 0$  we now define the Lagrangian  $L$  to be  $\alpha$ -coercive when there exist a function  $g: R^k \rightarrow R$ , bounded on bounded sets, and a positive constant  $k$  such that

$$(3.16) \quad L(t, x, v) \geq g(x) + k|v|^{1+\alpha} \quad \text{for all } (t, x, v) \in [a, b] \times R^n \times R^n.$$

Note that, in view of hypothesis (H3),  $L$  is always 0-coercive.

**COROLLARY 3.4.** *Suppose  $L$  is  $\alpha$ -coercive for some  $\alpha \geq 0$ ,  $L(t, x, v)$  is twice continuously differentiable in  $(t, x, v)$ , and  $L_{vv}(t, x, v) > 0$ , all  $(t, x, v) \in [a, b] \times R^n \times R^n$ . Suppose further that, corresponding to every bounded set  $S \subset R^n$ , there exists a constant  $c$  such that*

$$(3.17) \quad |F(t, x, v)| \leq c(1 + |v|^{2+\alpha}) \quad \text{for all } (t, x, v) \in [a, b] \times S \times R^n.$$

Then  $\Omega = [a, b]$ .

**PROOF.** Take  $S$  to be a set which contains the values of  $x(\cdot)$ . Then, by (3.17),

$$\begin{aligned} |F(t, x(t), \dot{x}(t))| &\leq c(1 + |\dot{x}(t)|^{2+\alpha}) \\ &\leq \gamma(t)(1 + |\dot{x}(t)|) \quad \text{a.e. } t \in [a, b], \end{aligned}$$

where  $\gamma(\cdot)$  is the function  $t \rightarrow c(1 + |\dot{x}(t)|^{1+\alpha})$ . However,  $t \rightarrow L(t, x(t), \dot{x}(t)) - g(x(t))$  is summable by the properties of the minimizing arc  $x(\cdot)$ . But then, by (3.16),  $\gamma(\cdot)$  also is summable. The hypotheses of Proposition 3.3 have been verified, and we are justified in concluding that  $\Omega = [a, b]$ .

**3.4. The case in which the Bernstein function is polynomially bounded.** We have established that if  $L$  is  $\alpha$ -coercive and  $F$  (given by (3.15)) has polynomial growth of degree at most  $2 + \alpha$ , then points of bad behaviour cannot occur. What can be said when  $F$  has polynomial growth of arbitrary degree? In such circumstances it is possible at least to calculate a lower bound for  $\int_a^b |\dot{x}(t)| dt$  in the event that points of bad behaviour exist. This information can be used to exclude points of bad behaviour when it permits us to deduce that, if they occurred, they would give rise to an excessively large value of  $\int_a^b |\dot{x}(t)| dt$ . We prove the lower bound first and then proceed to illustrate the technique with two examples.

PROPOSITION 3.4. Suppose  $L(t, x, v)$  is twice continuously differentiable in  $(t, x, v)$ ,  $L_{vv}(t, x(t), \dot{x}(t)) > 0$  a.e.  $t \in [a, b]$ , and there exist constants  $c_1$  and  $c_2$  and  $m > 1$  such that

$$(3.18) \quad |F(t, x(t), \dot{x}(t))| \leq c_1(|\dot{x}(t)| + c_2)^{m+1} \quad \text{for a.e. } t \in [a, b],$$

where  $F(t, x, v)$  is defined by (3.15).

Suppose also that  $\Omega$  is not the whole interval  $[a, b]$ . Then

$$\int_a^b |\dot{x}(t)| dt \geq \frac{1}{(mc_1)^{1/m}} \frac{m}{(m-1)} |b-a|^{1-1/m} - c_2 |b-a|.$$

Before proving the proposition we note

LEMMA 3.2.  $\Omega$  can be expressed as a countable union

$$(3.19) \quad \Omega = \bigcup_i \Omega_i$$

of nonempty, disjoint, relatively open subintervals of  $[a, b]$  in which, for each  $i$ ,  $\Omega_i$  is of the form:

- (i)  $[a, b_i)$  for some  $b_i \in (a, b)$ , or
- (ii)  $(a_i, b]$  for some  $a_i \in [a, b)$ , or
- (iii)  $(a_i, b_i)$  for some  $a_i, b_i \in [a, b]$ .

In (i)  $\limsup_{t \uparrow b_i} |\dot{x}(t)| = \infty$ ; in (ii)  $\limsup_{t \downarrow a_i} |\dot{x}(t)| = \infty$ ; and in (iii)

$$\limsup_{t \uparrow b_i} |\dot{x}(t)| = \infty \quad \text{and} \quad \limsup_{t \downarrow a_i} |\dot{x}(t)| = \infty.$$

PROOF. Since  $\Omega$  is a relatively open subset of  $[a, b]$ , it is known that  $\Omega$  can be expressed as a countable union (3.19) of nonempty, disjoint, relatively open subintervals. Such subintervals are of the form (i), (ii), or (iii).

Suppose  $\Omega_i = [a, b_i)$  for some  $b_i \in (a, b)$ . We claim  $b_i \notin \Omega$ . If, on the contrary,  $b_i \in \Omega$  then, since  $b_i \notin \Omega_i$ , we must have  $b_i \in \Omega_j$ , for some  $j \neq i$ . But  $\Omega_j$  is relatively open so there exists some  $b' \in [a, b_i)$  such that  $b' \in \Omega_j$ . But  $b' \in \Omega_i$ . This is not possible since  $\Omega_i, \Omega_j$  are disjoint.

Now  $b_i \notin \Omega$  and  $b_i \in (a, b]$  imply  $\limsup_{t \uparrow b_i} |\dot{x}(t)| = \infty$ , for otherwise the restriction of  $x(\cdot)$  to  $[a, b_i]$  would be a Lipschitz function. By Theorem 2.1, however, this would imply  $b_i \in \Omega$ , a contradiction.

The other cases are treated in similar fashion.

PROOF OF PROPOSITION 3.4. We decompose  $\Omega$  into nonempty, disjoint, relatively open subintervals  $\Omega = \bigcup_i \Omega_i$  as in Lemma 3.2. For each  $i$ , let  $a_i$  and  $b_i$  be the left and right endpoints of  $\Omega_i$ , respectively. Consider  $\Omega_1$ . By Lemma 3.2 either

$$(3.20) \quad \limsup_{t \uparrow b_1} |x(t)| = \infty$$

or

$$(3.21) \quad \limsup_{t \downarrow a_1} |x(t)| = \infty.$$

Let us assume (3.20). We shall obtain a lower bound on  $\int_{\Omega_1} |\dot{x}| dt$ . Similar arguments will give the same bound if, alternatively, (3.21) is assumed.

By the results of §2,  $x(\cdot)$  is  $C^2$  on  $(a_1, b_1)$  and  $x(\cdot)$  satisfies the Euler equation, which we can write as

$$\dot{x}(s) - \dot{x}(t) = \int_t^s F(\sigma, x(\sigma), \dot{x}(\sigma)) d\sigma, \quad t, s \in (a_1, b_1).$$

Arguing as in the proof of Proposition 3.3, we deduce from the hypothesis (3.18) that the absolutely continuous function  $s \rightarrow |\dot{x}(s)|$  satisfies

$$\frac{d}{ds} |\dot{x}(s)| \leq c_1(|\dot{x}(s)| + c_2)^{m+1} \quad \text{a.e. } t \in (a_1, b_1).$$

Now take  $t \in (a_1, b_1)$  and consider the function  $y(\cdot)$  which satisfies the differential equation (with boundary condition)

$$(3.22) \quad \begin{cases} \frac{d}{ds} y(s) = c_1(y(s) + c_2)^{m+1}, & s \geq t, \\ y(t) = |\dot{x}(t)|. \end{cases}$$

The differential equation well defines  $y(\cdot)$  on some maximal half-open interval  $[t, t + \varepsilon)$  with  $\varepsilon \in (0, \infty]$ . By a comparison theorem [11, Theorem 1.4.1]

$$(3.23) \quad |\dot{x}(s)| \leq y(s) \quad \text{for all } s \in [t, t + \varepsilon).$$

Now

$$(3.24) \quad b_1 \geq t + \varepsilon,$$

for if this inequality did not apply, it would follow from (3.23) and the monotonicity of  $y(\cdot)$  that

$$\infty = \limsup_{s \uparrow b_1} |\dot{x}(s)| \leq \lim_{s \uparrow b_1} y(s) < \infty,$$

a contradiction.

The differential equation (3.22) can be solved by the separation of variables technique.  $\varepsilon$  is finite and, in particular, (3.24) can be satisfied only if  $|\dot{x}(t)| + c_2 > 0$ . In this case we have

$$y(s) = [(|\dot{x}(t)| + c_2)^{-m} - mc_1(s - t)]^{-1/m} - c_2 \quad \text{for all } s \in [t, t + \varepsilon),$$

where

$$(3.25) \quad \varepsilon = (mc_1)^{-1} (|\dot{x}(t)| + c_2)^{-m}.$$

We deduce from (3.24) and (3.25) that

$$b_1 - t \geq (mc_1)^{-1} (|\dot{x}(t)| + c_2)^{-m}.$$

This inequality can be rearranged to give

$$|\dot{x}(t)| \geq (mc_1)^{-1/m} (b_1 - t)^{-1/m} - c_2.$$

But then

$$\begin{aligned} \int_{\Omega_1} |\dot{x}(t)| dt &\geq (mc_1)^{-1/m} \int_{a_1}^{b_1} (b_1 - t)^{-1/m} dt - c_2 |b_1 - a_1| \\ &= (mc_1)^{-1/m} |\Omega_1|^{1-1/m} \cdot \frac{m}{m-1} - c_2 |\Omega_1|, \end{aligned}$$

in which  $|\Omega_i|$  denotes  $|b_i - a_i|$ , the length of the interval  $\Omega_i$ . Identical reasoning applied when  $i = 2, 3, \dots$  gives the same inequality when  $\Omega_1$  is replaced by  $\Omega_i$ ,  $i =$

2, 3, \dots Since the  $\Omega_i$ 's are disjoint we deduce that, for any integer  $N$  in the index set for the  $\Omega_i$ 's,

$$\begin{aligned} \int_a^b |\dot{x}(t)| dt &\geq (mc_1)^{-1/m} \cdot \frac{m}{(m-1)} \cdot \sum_{i=1}^N |\Omega_i|^{1-1/m} - c_2 \sum_{i=1}^N |\Omega_i| \\ &\geq (mc_1)^{-1/m} \cdot \frac{m}{(m-1)} \cdot \left( \sum_{i=1}^N |\Omega_i| \right)^{1-1/m} - c_2 \sum_{i=1}^N |\Omega_i|. \end{aligned}$$

However,  $\bigcup_i \Omega_i$  has full measure, whence  $\sum_i |\Omega_i| = |b-a|$ . Taking the limit as  $N \rightarrow \infty$ , if necessary then, we deduce that

$$\int_a^b |\dot{x}(t)| dt \geq (mc_1)^{-1/m} \cdot \frac{m}{(m-1)} |b-a|^{1-1/m} - c_2 |b-a|,$$

as required.  $\square$

We now consider Lagrangians  $L = N + \varepsilon P$  expressed in terms of a nominal Lagrangian  $N$  and a perturbation term  $\varepsilon P$ . Attention is limited to a certain class of nominal Lagrangians  $N$  with the property that they cannot give rise to points of bad behaviour. We show that, under very mild hypotheses on  $P$ , this property is preserved under sufficiently small perturbations.

**COROLLARY 3.5.** *Take  $L(t, x, v) = N(v) + \varepsilon P(t, x, v)$  in which  $\varepsilon$  is a nonnegative parameter. We assume that  $N$  satisfies:*

*$N(v)$  is twice continuously differentiable in  $v$ , and there exists  $\delta > 0$  such that  $N_{vv}(v) > \delta I$  for all  $v \in R^n$ .*

*We assume further that  $P$  satisfies*

*$P(t, x, v)$  is twice continuously differentiable in  $(t, x)$ ,  $P(t, x, v) \geq 0$  for all  $(t, x, v) \in [a, b] \times R^n \times R^n$ ,  $P_{vv}(t, x, v) \geq 0$  for all  $(t, x, v) \in [a, b] \times R^n \times R^n$ , and there exists some constant  $c$  and some  $m > 1$  such that*

$$|P_x - P_{vt} - P_{vx}|(t, x(t), \dot{x}(t)) \leq c(|\dot{x}(t)| + 1)^{m+1}, \quad t \in [a, b].$$

*Then  $\Omega = [a, b]$ , provided  $\varepsilon$  is sufficiently small.*

**PROOF.** It is easy to check that conditions are satisfied under which a solution to (P) exists, whatever  $\varepsilon$ . In particular, there exists a solution  $z(\cdot)$  to the problem which results from setting  $\varepsilon$  equal to zero. Since  $N$  does not depend on  $x$ , we deduce from Proposition 3.2 that  $\dot{z}(\cdot)$  is essentially bounded. It follows that  $t \rightarrow P(t, z(t), \dot{z}(t))$  is summable.

Now our solution  $x(\cdot)$  to (P) must satisfy

$$(3.26) \quad \int_a^b N(\dot{x}(t)) dt \leq \int_a^b N(\dot{z}(t)) dt + \varepsilon \int_a^b P(t, z(t), \dot{z}(t)) dt,$$

since  $P$  is nonnegative. In view of the hypothesis  $N_{vv} > \delta I$ , consideration of a Taylor expansion of  $N$  about the origin with second order remainder term yields

$$(3.27) \quad \lim \frac{\int_a^b N(y(t)) dt}{\int_a^b |y(t)| dt} = +\infty,$$



where the limit is taken as  $\int_a^b |y(t)| dt$  tends to  $+\infty$ . Now (3.26) and (3.27) imply

$$(3.28) \quad \int_a^b |\dot{x}(t)| dt \leq k_1$$

for some constant  $k_1$  which does not depend on  $\varepsilon$ , as  $\varepsilon$  ranges over the interval  $[0, 1]$ , say.

Now examine  $F = L_{vv}^{-1}[L_x - L_{vt} - L_{vx}]$ . We deduce from the hypotheses that

$$|F(t, x, \dot{x})| \leq (\varepsilon c/\delta)(|\dot{x}| + 1)^{m+1} \quad \text{for almost all } t \text{ in } [a, b].$$

Suppose that  $\Omega \neq [a, b]$ . The last inequality implies, via Proposition 3.4, that

$$(3.29) \quad \int_a^b |\dot{x}(t)| dt \geq \left(\frac{m\varepsilon c}{\delta}\right)^{-1/m} \frac{m}{m-1} |b-a|^{1-1/m} - |b-a|.$$

Now suppose  $\varepsilon \in (0, 1]$  is chosen so that the right side is greater than  $k_1$ . For such an  $\varepsilon$ , (3.29) contradicts (3.28). We conclude that  $\Omega = [a, b]$ .  $\square$

A different approach to some of the issues treated in this article is possible—one that leads to results that are not global, but which hold “in the small”. A thorough discussion along these lines will appear in [8]. We now derive one such result as a second consequence of Proposition 3.4. It is distinguished by the precision with which the degree of “smallness” can be specified.

**COROLLARY 3.6.** *Suppose that for a given interval  $[a, T]$  the Lagrangian  $L$  and the Bernstein function  $F$  satisfy the following hypotheses on  $[a, T] \times R^n \times R^n$ :*

(a)  *$L$  is twice continuously differentiable,  $L_{vv} > 0$ ,  $L(t, x, v) \geq \beta|v|^p + \lambda$ , where  $\beta, p, \lambda$  are constants with  $\beta > 0$ ,  $p > 1$ .*

(b)  *$|F(t, x, v)| \leq c_1(|v| + c_2)^{m+1}$ , where  $c_1, c_2$  and  $m$  are constants with  $c_2 \geq 0$ ,  $m > 1$ .*

*Let any number  $M$  and point  $A$  in  $R^n$  be given, and set*

$$\Sigma := \max\{L(t, y, v) : a \leq t \leq T, |y - A| \leq (T - a)M, |v| \leq M\},$$

$$\Delta := [mc_1(m - 1)^m]^{-1}[\beta/(\Sigma - \lambda)]^{m/p}.$$

*Then for any  $b$  in  $(a, T]$  satisfying  $b < a + \Delta$  and for any  $B$  in  $R^n$  satisfying  $|B - A| \leq M(b - a)$ , the problem of minimizing  $\int_a^b L(t, y, \dot{y}) dt$  subject to  $y(a) = A$ ,  $y(b) = B$  has a solution, and any solution is twice continuously differentiable.*

**PROOF.** Let  $q$  satisfy  $1/p + 1/q = 1$ , and apply Holder’s inequality to get, for any solution  $x$ ,

$$\begin{aligned} \left\{ \int_a^b |\dot{x}| dt \right\}^p &\leq (b - a)^{p/q} \int_a^b |\dot{x}|^p dt \\ &\leq (b - a)^{p/q} \left\{ \int_a^b [L(t, x, \dot{x}) - \lambda] dt \right\} / \beta. \end{aligned}$$

If  $y$  denotes the linear arc joining  $(a, A)$  to  $(b, B)$ , the observation

$$\int_a^b L(t, x, \dot{x}) dt \leq \int_a^b L(t, y, \dot{y}) dt \leq (b - a)\Sigma$$

together with the preceding inequality lead to an upper bound for  $\int_a^b |\dot{x}| dt$ . If  $x$  fails to be  $C^2$ , then  $\Omega$  must fail to be all of  $[a, b]$ , and Proposition 3.4 provides a lower bound for  $\int_a^b |\dot{x}| dt$ . A contradiction results, as long as  $b < a + \Delta$ .  $\square$

## REFERENCES

1. S. Bernstein, *Sur les équations du calcul des variations*, Ann. Sci. Ecole Norm. Sup. (3) **29** (1912), 431–485.
2. G. A. Bliss, *Lectures on the calculus of variations*, Univ. of Chicago Press, Chicago, Ill., 1946.
3. L. Cesari, *Optimization—theory and applications*, Springer-Verlag, New York, 1983.
4. F. H. Clarke, *The Euler-Lagrange differential inclusion*, J. Differential Equations **19** (1975), 80–90.
5. ———, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.
6. ———, *Generalized gradients of Lipschitz functionals*, Adv. in Math. **40** (1981), 52–67.
7. ———, *Optimization and nonsmooth analysis*, Wiley Interscience, New York, 1983.
8. F. H. Clarke and R. B. Vinter, *Existence and regularity in the small in the calculus of variations*, J. Differential Equations (to appear).
9. ———, *On the conditions under which the Euler equation or the maximum principle hold*, Appl. Math. Optim. **12** (1984), 73–79.
10. A. D. Ioffe and V. M. Tihomirov, *Theory of extremal problems*, North-Holland, Amsterdam, 1979.
11. V. Lakshmikantham and S. Leela, *Differential and integral inequalities*, Academic Press, New York, 1969.
12. C. B. Morrey, Jr., *Multiple integrals in the calculus of variations*, Springer-Verlag, Berlin, 1966.
13. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.
14. L. Tonelli, *Sur une méthode directe du calcul des variations*, Rend. Circ. Mat. Palermo **39** (1915), 233–264, also appears in *Opere scelte*, Vol. 2, Cremonese, Rome, 1961, pp. 289–333.
15. ———, *Fondamenti di calcolo delle variazioni*, Vols. 1, 2, Zanichelli, Bologna, 1921, 1923.

CENTRE DE RECHERCHE DE MATHÉMATIQUES APPLIQUÉES, UNIVERSITÉ DE MONTRÉAL, C.P. 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, CANADA H3C 3J7

DEPARTMENT OF ELECTRICAL ENGINEERING, IMPERIAL COLLEGE, LONDON SW7 2BT, ENGLAND