ANNALI DELLA
SCUOLA NORMALE SUPERIORE DI PISA
Classe di Scienze

FRANK H. CLARKE
PHILIP D. LOEWEN

An intermediate existence theory in the calculus of variations

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4e série, tome 16, nº 4 (1989), p. 487-526
<http://www.numdam.org/item?id=ASNSP_1989_4_16_4_487_0>

An Intermediate Existence Theory in the Calculus of Variations

FRANK H. CLARKE - PHILIP D. LOEWEN

1. - Introduction

This article develops an existence theory for solutions to the following basic problem (P) of the calculus of variations:

$$(P) \qquad \text{minimize} \left\{ \Lambda(x) := \int_a^b L\left(t, x(t), \dot{x}(t)\right) \mathrm{d}t : x(a) = x_a, x(b) = x_b \right\}.$$

There is a vast literature on this subject for which the definitive reference is the monograph of Cesari [1]. In this introduction we shall merely make some general remarks to help situate the results of the article and contrast them to existing work. It was Tonelli's seminal work that first produced a general existence theorem for (P), by identifying: (i) appropriate conditions on the Lagrangian L, and (ii) a suitable space of functions from which the competing x's are drawn. The space in question (AC) is the class of absolutely continuous functions x mapping $[a, b]$ to $\mathbb{R}^n$ (we call such functions *arcs*), and the conditions on L include notably:

1) convexity in the velocity: for each (t, x), the function $v \to L(t, x, v)$ is convex;

2) coercivity: for some $\varepsilon > 0$ and constant c, L satisfies

$$L(t, x, v) \geq \varepsilon |v|^2 + c, \qquad \text{for every } (t, x, v).$$

Tonelli's work introduced the topological approach, still the standard one, whereby existence is deduced from compactness properties of level sets of the (lower semicontinuous) functional Λ.

One of the numerous extensions of Tonelli's existence theory is that of Rockafellar [12]; we shall give it here (informally) to facilitate comparison later,

Pervenuto alla Redazione l'1 Agosto 1988.

and because it makes use of the Hamiltonian which plays an important rôle in this article too. The Hamiltonian function H is defined as follows:

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

The *basic growth condition* [12] on H requires that, for each fixed $p \in \mathbb{R}^n$ and bounded subset S of $\mathbb{R}^n$, there exists a summable function φ on $[a,b]$ such that $H(t,x,p) \leq \varphi(t)$ for all t in $[a,b]$ and x in S. In the following, a form of coercivity is provided by this condition.

THEOREM [12]. *If L is convex in the velocity and H satisfies the basic growth condition, then the following level set is compact for every λ and R:*

$$\{x \in AC : \Lambda(x) \leq \lambda,\ \|x\|_\infty \leq R\}.$$

We shall not specify the sense of the word "compact", but remark only that under the hypotheses of the theorem, it follows immediately that the problem (P) admits a solution in AC under the additional state constraint $|x(t)| \leq R$.

The use of existence theory is often a prelude to invoking necessary conditions, for example the Euler equation. Indeed, the necessary conditions are sometimes themselves the goal, as is the case in §8. There are in this regard two difficulties with standard existence results such as the one given above. The first has to do with the state constraint $|x(t)| \leq R$: if the solution x has $|x(t)| = R$ at some point t (i.e., it is not "interior"), then the Euler equation is no longer a necessary condition. The second difficulty is that even if the solution x is interior, it may not satisfy the standard necessary conditions unless it belongs to one of certain subclasses of AC (for example, the class AC^∞ consisting of Lipschitz arcs, i.e., those having essentially bounded derivative). This latter fact has only recently been clarified, and we refer to [8] for a complete discussion.

The theory developed in this article applies to the following problem:

$$(P_R) \qquad \text{minimize}\{\Lambda(x) : \|x\|_\infty < R,\ x(a) = x_a,\ x(b) = x_b\}.$$

Note the strict inequality in the state constraint, which assures that any solution is "interior". Here is a sample result (§3) which invokes two hypotheses defined only in later sections; we also omit here the provenance of the quantities m and $\Lambda(\overline{x})$.

THEOREM. *Let L satisfy the extremal growth condition, and let L be strictly convex at infinity in the velocity. Suppose that for some $\overline{\alpha} > 0$, for all t in $[a,b]$ and for all x in $\mathbb{R}^n$ satisfying $|x| < R$, we have:*

$$\text{if } |p| \leq \overline{\alpha} \text{ then } H(t,x,p) \leq \frac{(R-m)\overline{\alpha} - \Lambda(\overline{x})}{b-a}.$$

Then the set of solutions to (P_R) is a nonempty subset of AC^∞.

Let us proceed to note some advantages of this type of result. Foremost, not only the existence of an interior solution is asserted, but also that all solutions are Lipschitz. Thus existence and regularity are combined, and the solutions obtained satisfy the standard necessary conditions.

Observe that the growth of the Hamiltonian is restricted only for "intermediate" values of p, and that one could imagine attaining the desired bound by adjusting the value of R (and $b - a$) appropriately. Another sense in which the word "intermediate" may be invoked is related to the growth properties of L. As examples will show, the theorem applies to non-coercive ("slow-growth") Lagrangians as well as coercive ("fast growth") ones. Perhaps the easiest way to delineate the distinction between our approach and the standard ones is to point out that it applies even to cases in which the following level sets fail to be compact or closed:

$$\{x : \Lambda(x) \leq \lambda, \ \|x\|_\infty < R\}, \ \{x : \Lambda(x) \leq \lambda, \ \|x\|_\infty \leq R\}.$$

An example to illustrate this is given in §3.

There is a price to pay for these gains, and it is reflected in the more restrictive conditions imposed on L. Because the approach used here involves necessary conditions at some point, we must impose on L hypotheses appropriate to such conditions, in the present case that L be locally Lipschitz. (In contrast, the result of Rockafellar cited above allows L to be extended-valued). Further, the extremal growth condition used in the illustrative theorem above is not always readily verifiable. We will give however several classes of more verifiable conditions which imply it.

Interestingly, two of these are among those used by Clarke and Vinter [8] in deriving regularity theorems, which may point to the extremal growth condition as a useful unifying concept for regularity and existence.

The principal techniques used in the proof are those of nonsmooth analysis, together with Tonelli's method of auxiliary Lagrangians as extended by Clarke and Vinter [8], [9]. The next section is devoted to some technical preliminaries, concerning in particular the hypothesis of "strict convexity at infinity". The main result is presented and discussed in §3, and proved in §§ 4 and 5. In §6 we specialize to existence "in the small" and in the following section to existence "in the large" ($R = \infty$). The last section makes an application of the theory to obtain a new result on the existence of periodic Hamiltonian trajectories.

2. - Technical background

Before stating our main result precisely in Section 3, we pause to fix some notation and introduce the notion of strict convexity at infinity. We also review the idea of essential values and indicate its relevance to subsequent sections. At the end we define the extremal growth condition (ECG) and prove several technical results for later use.

The positive integer n, fixed throughout this paper, denotes the dimension of our problem's state space. For any real numbers $a < b$, we write either $AC([a,b];\mathbb{R}^n)$ or simply $AC[a,b]$ for the space of absolutely continuous functions ("arcs") carrying $[a,b]$ into $\mathbb{R}^n$. A similar abbreviation is used for the subspace of Lipschitzian arcs, defined by

$$AC^\infty([a,b];\mathbb{R}^n) = \{X \in AC[a,b] : \dot{x} \in L^\infty[a,b]\}.$$

The letter B stands for the open unit ball of $\mathbb{R}^n$.

2.1 *Strict convexity at infinity*

Recall that a convex function $\ell : \mathbb{R}^n \to \mathbb{R}$ is called strictly convex if the graph of ℓ contains no line segments. To produce a local version of this definition, one might call the function ℓ strictly convex at x if the graph of ℓ contains no line segments through the point $(x, \ell(x))$. Our notion of strict convexity at infinity extends this natural idea to the point at infinity, and proves to be considerably weaker than global strict convexity. Thus, define ℓ's conjugate $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$h(p) = \ell^*(p) = \sup\{\langle p, v\rangle - \ell(v) : v \in \mathbb{R}^n\}.$$

Then h is a lower semicontinuous convex function which may take the value $+\infty$, but which must be finite somewhere. Fenchel's duality theorem states that for any $p,\ v \in \mathbb{R}^n$,

$$p \in \partial\ell(v) \Leftrightarrow v \in \partial h(p) \Leftrightarrow h(p) + \ell(v) = \langle p, v\rangle. \tag{2.1}$$

We define the multifunction $w(p) = \{v : p \in \partial\ell(v)\}$. By (2.1),

$$w(p) = \partial h(p) = \{v : h(p) = \langle p, v\rangle - \ell(v)\}. \tag{2.2}$$

DEFINITION 2.1. A convex function $\ell : \mathbb{R}^n \to \mathbb{R}$ is strictly convex at infinity if any one of the following equivalent conditions is satisfied:

(a) The graph of ℓ contains no rays;

(b) $w(p)$ is either empty or bounded for every $p \in \mathbb{R}^n$;

(c) $\partial\ell(v) \subseteq \operatorname{int}(\operatorname{dom} h)$ for every $v \in \mathbb{R}^n$;

(d) For any $(v,p) \in Gr\ \partial\ell$ and any $(w,r) \in 0^+ \operatorname{epi} \ell$ with $w \neq 0$, one has $\langle p, w\rangle < r$;

(e) For any $r > 0$ there exists $M > r$ such that the following quantity is positive:

$$\inf\{\ell(w) - \ell(v) - \langle p, w - v\rangle : |v| \leq r,\ p \in \partial\ell(v),\ |w| \geq M\};$$

(f) For any $r > 0$ there exists $M > r$ such that for all $p \in \mathbb{R}^n$, $w(p) \cap r\overline{B} \neq \emptyset$ implies $w(p) \subseteq MB$.

The equivalence of (a)-(f) is a painless exercise in convex analysis, for which [11] is the standard reference. We list all six conditions because different viewpoints are always useful in both analysis and interpretation, but for practical reasons we favour condition (f). This is because (f) makes explicit a compactness property of the set

$$\cup\{w(p) : p \text{ obeys } w(p) \cap r\overline{B} \neq \emptyset\}$$

which is not so evident in conditions like (a) and (b). We remark that a sufficient (but not necessary) condition for strict convexity at infinity is that h be finite everywhere, for in that case it is known that h is Lipschitz and that $\partial h(p)$ is everywhere compact so that criterion (b) applies.

An example of a function which is strictly convex at infinity but whose graph contains line segments arbitrarily far from the origin is easy to imagine. When $n = 1$, for example, one may start with $f(v) = (1 + v^2)^{1/2}$ and sketch the graph of a suitable ℓ by drawing straight lines between $(i, f(i))$ and $(i + 1, f(i + 1))$, $i \in \mathbb{Z}$. Note that for this example, $h(p) = +\infty$ when $|p| > 1$.

In the body of the paper we will consider finite-valued Lagrangians $L(t, x, v)$ defined on $\Omega \times \mathbb{R}^n$, where $\Omega := [a, b] \times R\overline{B}$ is compact and convex. Our assumptions will imply that L is jointly continuous in all its arguments and convex in v for each fixed $(t, x) \in S$. For such Lagrangians the Hamiltonian $H : \Omega \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is given by

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

and we write

$$\begin{aligned} W(t, x, p) &= \{v : p \in \partial_v L(t, x, v)\} \\ &= \{v : H(t, x, p) = \langle p, v\rangle - L(t, x, v)\} = \partial_p H(t, x, p). \end{aligned}$$

PROPOSITION 2.2. *Let $L : \Omega \times \mathbb{R}^n \to \mathbb{R}$ be continuous. For each $\omega \in \Omega$, suppose that $L(\omega, \cdot)$ is a convex function which is strictly convex at infinity. Then for every $r > 0$ there exists $M > 0$ such that*

(2.3) *for all $(\omega, p) \in \Omega \times \mathbb{R}^n$, $W(\omega, p) \cap r\overline{B} \neq \emptyset$ implies $W(\omega, p) \subseteq M\overline{B}$.*

PROOF. Fix $r > 0$. According to Def. 2.1(e), for every $\omega \in \Omega$ the number $M(r, \omega) = \inf\{\mu : I(r, \omega, \mu) > 0\}$ is finite, where

$$I(r, \omega, \mu) = \inf\{L(\omega, w) - L(\omega, v) - \langle p, w - v\rangle : |v| \leq r,\ p \in \partial_v L(\omega, v), |w| \geq \mu\}.$$

We now claim that $\sup_{\omega \in \Omega} M(r, \omega) < +\infty$. To prove this, suppose the contrary – i.e. suppose there is a sequence $\{\omega_i\}$ in Ω along which $M(r, \omega_i) \to +\infty$. Along a suitable subsequence (which we do not relabel) we may assume $M(r,\omega_i) > i$ for every i. This means that $I(r,\omega_i,i) = 0$ for every i, so there exist points $v_i \in r\overline{B}$, $p_i \in \partial_v L(\omega_i, v_i)$, and w_i with $|w_i| = i$, such that

$$(*) \qquad L(\omega_i, w_i) - L(\omega_i, v_i) - \langle p_i, w_i - v_i\rangle = 0 \text{ for every } i.$$

(Here we have used the fact that the constraint $|w| \geq \mu$ defining $I(r,\omega,\mu)$ may be replaced by the constraint $|w| = \mu$. This reduction displays $I(r,\omega,\mu)$ as the infimum of a continuous function over a compact set, which we know is attained). We may pass to a subsequence along which $v_i \to v \in r\overline{B}$, $\omega_i \to \omega \in \Omega$, and $p_i \to p \in \partial_v L(\omega, v)$. It is easy to deduce from (*) that

$$L\left(\omega_i, v_i + \lambda(w_i - v_i)\right) = L(\omega_i, v_i) + \lambda\langle p_i, w_i - v_i\rangle, \text{ for all } \lambda \in [0,1], \text{ for all } i.$$

In particular, if we fix any $k > r$, then for every $i > k$ there is a point u_i of norm k having the form $u_i = v_i + \lambda(w_i - v_i)$ for some $\lambda \in (0,1)$ such that this point obeys

$$L(\omega_i, u_i) = L(\omega_i, v_i) + \langle p_i, u_i - v_i\rangle, \text{ for every } i.$$

Along a further subsequence, we may assume that u_i converges to a point u_k of norm k, and taking limits in the previous line then gives

$$L(\omega, u_k) - L(\omega, v) - \langle p, u_k - v\rangle = 0, \text{ for every } k.$$

This forces $M(r,\omega) = +\infty$, a contradiction.

We may therefore define $M(r) = 1 + \sup_{\omega\in\Omega} M(r,\omega)$, a finite number for which $I(r,\omega,M(r)) > 0$ for every $\omega \in \Omega$. Let us show that $M(r)$ will serve in (2.3). Indeed, if this assertion were false then there would be some pair $(\omega, p) \in \Omega \times \mathbb{R}^n$ for which $W(\omega,p)$ contains both a point $v \in r\overline{B}$ and a point $w \notin M(r)\overline{B}$. In particular,

$$H(\omega,p) = \langle p, v\rangle - L(\omega,v) = \langle p, w\rangle - L(\omega, w)$$

and this implies $I(r,\omega,M(r)) = 0$, a contradiction. ∎

Let us now record a useful property of the multifunction W.

PROPOSITION 2.3. *Let $L : \Omega\times\mathbb{R}^n \to \mathbb{R}$ be continuous and suppose $L(t,x,\cdot)$ is convex for each $(t,x) \in \Omega$. Then the multifunction W has closed graph.*

PROOF. We must show that if sequences v_i, t_i, x_i, p_i are given converging to points v, t, x, p respectively, then the relationships

$$(2.4) \qquad v_i \in W(t_i, x_i, p_i) \quad \text{for all } i$$

imply the limiting relationship $v \in W(t,x,p)$. To prove this, note that (2.4) is equivalent to

$$\langle p_i, w - v_i \rangle \le L(t_i, x_i, w) - L(t_i, x_i, v_i) \quad \forall i,\ \forall w \in \mathbb{R}^n.$$

The continuity assumed for L allows us to take the limit as $i \to \infty$ and obtain

$$\langle p, w - v \rangle \le L(t,x,w) - L(t,x,v) \quad \text{for every } w \in \mathbb{R}^n.$$

This is equivalent to $p \in \partial_v L(t,x,v)$, i.e. $v \in W(t,x,p)$. ∎

We now define a function ρ_0 which is more or less the inverse of the map $r \to M$ expressed in Proposition 2.2. For any $s > 0$, we set $\rho_0(s)$ equal to

$$\inf\{r \ge 0 : \text{ for some } (\omega,p) \in \Omega \times \mathbb{R}^n, \text{ one has both } W(\omega,p) \cap r\overline{B} \ne \emptyset \text{ and } W(\omega,p) \cap (\mathbb{R}^n \backslash sB) \ne \emptyset\}.$$

We proceed to define $\rho(s) := \min[s, \rho_0(s)]$.

PROPOSITION 2.4. *Let $L : \Omega \times \mathbb{R}^n \to \mathbb{R}$ be convex in v and continuous. Then the function ρ has the following properties*:

(i) *ρ is nondecreasing and obeys $0 \le \rho(s) \le s$, for all $s > 0$*

(ii) *$\forall(\omega,p) \in \Omega \times \mathbb{R}^n$, $\forall s > 0$, if $W(\omega,p) \cap \rho(s)B \ne \emptyset$ then $W(\omega,p) \subseteq sB$*

(iii) *if $L(\omega,\cdot)$ is strictly convex for each $\omega \in \Omega$, then $\rho(s) = s$, for all $s > 0$*

(iv) *if $L(\omega,\cdot)$ is strictly convex at infinity for each $\omega \in \Omega$, then $\rho(s) \to \infty$ as $s \to \infty$.*

PROOF. The first property is immediate. Property (ii) is satisfied by ρ_0 simply by its definition; since $\rho \le \rho_0$, (ii) is all the more true for ρ. As for (iii), when $L(t,x,v)$ is strictly convex in v then $W(t,x,p)$ is either empty or a singleton for each (t,x,p). Consequently the r occurring in the definition of $\rho_0(s)$ may clearly be limited to $[s,\infty)$, whence $\rho_0(s) \ge s$. It follows that we have $\rho(s) = s$. Finally, to prove (iv) we need only to show that $\rho_0(s) \to \infty$, as $s \to \infty$. Were this assertion false, there would then exist a sequence $s_i \to \infty$, corresponding points $\omega_i \in \Omega$ and $p_i \in \mathbb{R}^n$, and a number r_0 such that

$$W(\omega_i,p_i) \cap r_0\overline{B} \ne \emptyset \text{ and } W(\omega_i,p_i) \cap (\mathbb{R}^n \backslash s_iB) \ne \emptyset, \text{ for every } i.$$

This clearly contradicts Proposition 2.2. ∎

2.2 *Essential values*

For any measurable function $f : \mathbb{R} \to \mathbb{R}^n$, the set of essential values of f at time t is defined by

$$\text{Ess } f(t) = \{y \in \mathbb{R}^n : \forall \varepsilon > 0,\ m\{s \in (t - \varepsilon, t + \varepsilon) : |y - f(s)| < \varepsilon\} > 0\}.$$

It is quite possible for this set to be empty, as the example Ess $f(0)$ for $f(t) = t^{-1}$ illustrates.

Nonetheless it is easy to see that Ess $f(t)$ is a closed set for each t. One can even say more: the multifunction $t \to$ Ess $f(t)$ has closed graph. This is the content of the following.

PROPOSITION 2.5. *Let* $f : \mathbb{R} \to \mathbb{R}^n$ *be measurable. For any sequences* t_i, y_i *such that* $t_i \to t$, $y_i \to y$, *and* $y_i \in$ Ess $f(t_i)$ *for every* i, *one has the limiting relationship* $y \in$ Ess $f(t)$.

PROOF. Given any $\varepsilon > 0$, choose I so large that $i \geq I$ implies both

$$|t_i - t| < \frac{\varepsilon}{2} \text{ and } |y_i - y| < \frac{\varepsilon}{2}. \tag{2.5}$$

Now, for any fixed $i \geq I$, the relation $y_i \in$ Ess $f(t_i)$ implies that the following set has positive measure:

$$S_i\left(\frac{\varepsilon}{2}\right) = \left\{s \in \left(t_i - \frac{\varepsilon}{2}, t_i + \frac{\varepsilon}{2}\right) : |y_i - f(s)| < \frac{\varepsilon}{2}\right\}.$$

But (2.5) implies that

$$\begin{aligned} S_i\left(\frac{\varepsilon}{2}\right) &\subseteq \left\{s \in (t - \varepsilon, t + \varepsilon) : |y_i - f(s)| < \frac{\varepsilon}{2}\right\} \\ &\subseteq \{s \in (t - \varepsilon, t + \varepsilon) : |y - f(s)| < \varepsilon\}. \end{aligned}$$

Thus, the set on the right side must have positive measure; and since $\varepsilon > 0$ is arbitrary, it follows that $y \in$ Ess $f(t)$. ∎

A multifunction $\Gamma : \mathbb{R} \to \mathbb{R}^n$ is called upper semicontinuous at t_0 if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|t - t_0| < \delta \text{ implies } \Gamma(t) \subseteq \Gamma(t_0) + \varepsilon B.$$

(This tacitly assumes $\Gamma(t_0) \neq \emptyset$). Even when $\Gamma(t) =$ Ess $f(t)$ for a measurable function f, this concept is strictly stronger than the closed graph property proved in Prop. 2.5. For example, the function $f(t)$ which is 0 on $(-\infty, 0]$ and t^{-1} on $(0, +\infty)$ obeys Ess $f(t) = \{f(t)\}$: although this multifunction has closed graph and nonempty values, it fails to be upper semicontinuous at $t_0 = 0$. ∎

Our next result shows how this cannot happen if f is essentially bounded.

COROLLARY 2.6. *Let* $f : \mathbb{R} \to \mathbb{R}^n$ *be measurable and essentially bounded on a neighbourhood of some point* t_0. *Then* Ess $f(t_0)$ *is a nonempty compact set, and the multifunction* Ess f *is upper semicontinuous at* t_0.

PROOF. By assumption, there exist $\varepsilon > 0$ and $M > 0$ such that $|f(t)| \leq M$ a.e. on $(t_0 - \varepsilon, t_0 + \varepsilon)$. It follows immediately that Ess $f(t) \subseteq M\overline{B}$ for every $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$; being a closed subset of the compact set $M\overline{B}$ makes Ess $f(t_0)$ compact. The compactness of $M\overline{B}$ is also central to the verification that Ess $f(t_0)$ must contain at least one point. (Hint: Suppose not...)

To check that Ess f is upper semicontinuous at t_0, suppose the contrary. That is, suppose there is some $\varepsilon > 0$ and some sequence $t_i \to t_0$ for which the corresponding sets Ess $f(t_i)$ each contain a point y_i outside Ess $f(t_0) + \varepsilon B$. Since $|y_i| \leq M$ for all i sufficiently large, we may pass to a subsequence along which y_i converges to some y_0 also outside Ess $f(t_0) + \varepsilon B$. The existence of such a subsequence, however, contradicts Proposition 2.5. So Ess f must be upper semicontinuous at t_0. ∎

Our next result describes a useful construction based on the upper semicontinuity of Ess f.

Anticipating the notation of future sections, we now consider $f = \dot{x}$ for a Lipschitz arc x defined on a fixed interval $[a, b]$.

PROPOSITION 2.7. *Let* $x \in AC^\infty[a, b]$ *be given. Suppose that for some* $\tau \in [a, b]$ *and* $\sigma > 0$, *one has* Ess $\dot{x}(\tau) \subseteq \sigma B$. *Define*

$$t_0 = \sup\left(\{t \in [a, \tau) : \text{ Ess } \dot{x}(t) \cap (\mathbb{R}^n \backslash \sigma B) \neq \emptyset\} \cup \{a\}\right),$$

$$t_1 = \inf\left(\{t \in (\tau, b] : \text{ Ess } \dot{x}(t) \cap (\mathbb{R}^n \backslash \sigma B) \neq \emptyset\} \cup \{b\}\right),$$

$$S = \{v \in \mathbb{R}^n : |v| = 1\}.$$

Then the interval $[t_0, t_1]$ *contains a neighbourhood of* τ *relative to* $[a, b]$ *and one has*

(i) Ess $\dot{x}(t) \cap \sigma\overline{B} \neq \emptyset$ *for every* $t \in [t_0, t_1]$;

(ii) Ess $\dot{x}(t) \subseteq \sigma B$ *for every* $t \in (t_0, t_1)$;

(iii) *Either* $\overline{co}$ Ess $\dot{x}(t_0) \cap \sigma S \neq \emptyset$, *or else* $t_0 = a$ *and* Ess $\dot{x}(a) \subseteq \sigma B$;

(iv) *Either* $\overline{co}$ Ess $\dot{x}(t_1) \cap \sigma S \neq \emptyset$, *or else* $t_1 = b$ *and* Ess $\dot{x}(b) \subseteq \sigma B$.

PROOF. Since $\dot{x}$ is essentially bounded, the multifunction Ess $\dot{x}$ has nonempty compact values. Thus the inclusion Ess $\dot{x}(\tau) \subseteq \sigma B$ implies that Ess $\dot{x}(\tau) + \varepsilon B \subseteq \sigma B$ for some $\varepsilon > 0$. By the upper semicontinuity of Ess $\dot{x}$ (Cor. 2.6), there is a neighbourhood of τ relative to $[a, b]$ on which this inclusion persists. This neighbourhood is clearly a subset of $[t_0, t_1]$.

A similar argument proves assertion (i). Indeed, if (i) were false then there would be some $t \in [t_0, t_1]$ at which Ess $\dot{x}(t) \subseteq \mathbb{R}^n \backslash \sigma\overline{B}$. The upper semicontinuity of Ess $\dot{x}$ would then supply a neighbourhood of t throughout which this inclusion persists, and the definition of t_0 or t_1 would be contradicted.

Assertion (ii) is an obvious consequence of our choices of t_0 and t_1. To prove (iii), we consider two cases. If Ess $\dot{x}(t_0)$ contains a point outside σB, then condition (i) implies that $\overline{co}$ Ess $\dot{x}(t_0)$ must contain a point on the boundary of σB. This is the first alternative. The only other possibility is Ess $\dot{x}(t_0) \subseteq \sigma B$, in which case the arguments of the first paragraph apply to t_0 as well as to τ, and a contradiction ensues unless $t_0 = a$.

The proof of (iv) is similar to that of (iii). ∎

2.3 *Essential values and extremals*

The multifunction $W(t,x,p)$ used to define strict convexity at infinity comes together with the concept of essential values when we study extremals of a given Lagrangian. We prefer not to define extremality precisely here, but observe that each of the three potential definitions, namely [2], [4, Thm. 4.2.2], and [4, Thm. 5.2.1], implies the condition (2.6) used in the next result.

PROPOSITION 2.8. *Let* $L : \Omega \times \mathbb{R}^n \to \mathbb{R}$ *be continuous, and suppose* $L(t,x,\cdot)$ *is a convex function for each fixed* $(t,x) \in \Omega$. *Furthermore, suppose two arcs* x *and* p *are given, obeying*

$$p(t) \in \partial_v L(t, x(t), \dot{x}(t)) \quad \text{a.e.} \tag{2.6}$$

Then for all t *without exception, one has*

$$\overline{co}\ \text{Ess}\ \dot{x}(t) \subseteq W(t, x(t), p(t)). \tag{2.7}$$

PROOF. For any fixed t and any $w \in$ Ess $\dot{x}(t)$, there must be a sequence $t_i \to t$ along which (2.6) holds while $\dot{x}(t_i) \to w$. Now at each time t_i, (2.6) is equivalent to

$$\dot{x}(t_i) \in W(t_i, x(t_i), p(t_i)).$$

Consequently $w \in W(t, x(t), p(t))$ by Proposition 2.3. We conclude that

$$\text{Ess}\ \dot{x}(t) \subseteq W(t, x(t), p(t)).$$

To deduce (2.7) from this, it suffices to note that the right side is itself a closed convex set, being simply $\partial_p H(t, x(t), p(t))$. ∎

To see why Proposition 2.8 is fundamental to the regularity of solutions in the calculus of variations, suppose x is a Lipschitzian extremal for some continuous Lagrangian L which happens to be strictly convex in v. The Lipschitz character of x implies that Ess $\dot{x}(t)$ is compact and nonempty for each t by Cor. 2.6, while the strict convexity of L implies that $W(t,x,p)$ never contains more than one point. So by Proposition 2.8, Ess $\dot{x}(t)$ consists of exactly one point for every t, which is to say that the Lipschitz extremal x is actually smooth!

DEFINITION 2.9. For all scalars $0 < r < s$, we define the nonnegative numbers

$$\Delta_R(r,s) = \inf\{t_1 - t_0\},$$

where the infimum is taken over all intervals $[t_0, t_1]$ in $[a, b]$ on which there is a Lipschitz arc x obeying

(a) Ess $\dot{x}(\tau) \cap r\overline{B} \neq \emptyset$ for some $\tau \in [t_0, t_1]$;

(b) Ess $\dot{x}(\sigma) \cap (\mathbb{R}^n \backslash \rho(s)B) \neq \emptyset$ for some $\sigma \in [t_0, t_1]$;

(c) Ess $\dot{x}(t) \subseteq sB$ for all $t \in (t_0, t_1)$;

(d) $|x(t)| < R$ for all $t \in [t_0, t_1]$;

(e) x solves the following problem:

$$\min \left\{ \int_{t_0}^{t_1} L(t, y, \dot{y}) dt : y(t_i) = x(t_i),\ |y(t)| < R\ \forall t \in [t_0, t_1],\ |\dot{y}(t)| \leq s \text{ a.e.} \right\}.$$

(Note that conditions (a) and (b) must be understood as involving the appropriate onesided essential value in cases where either σ or τ falls at an endpoint of $[t_0, t_1]$).

Note that for very large values of s, condition (b) above may be incompatible with (a), (d), and (e). In cases like this $\Delta_R(r, s)$ equals $+\infty$, the infimum over the empty set. Examples of this desirable behaviour are presented in Section 7 in connection with global existence theorems. Even when one can only assert $\lim_{s\to\infty} \Delta_R(r, s) > 0$ for all $R, r > 0$, local existence results will follow, as we demonstrate in Section 6.

DEFINITION 2.10. The Lagrangian L is said to satisfy the *extremal growth condition* (ECG) provided that for every R and $r > 0$, one has $\lim_{s\to\infty} \Delta_R(r, s) = +\infty$.

In the arguments to follow, the functions ρ and Δ_R provide quantitative control over changes in the size of Lipschitz extremals' derivatives. Here is how they work. The index ρ monitors possible jumps in the derivative, in view of Proposition 2.4 (ii). Consider a Lipschitz extremal x. In view of Proposition 2.8, we see that if Ess $\dot{x}(t)$ contains some vector v with $|v| \geq s$, then the whole set $\overline{co}$ Ess $\dot{x}(t)$ must lie outside the ball of radius $\rho(s)$. It is impossible for $|\dot{x}|$ ever to jump from a value inside $\rho(s)B$ to a value outside sB. (Note that if L is strictly convex in v, then $W(t, x, p)$ is a singleton and $\rho(s) = s$ for all $s > 0$: thus, the previous statement asserts that $|\dot{x}|$ is continuous).

As for $\Delta_R(r, s)$, it measures the shortest time in which a Lipschitzian extremal x can move from $|\dot{x}(\tau)| \leq r$ to $|\dot{x}(\sigma)| \geq s$. This is clearest in the strictly convex case, where one has $x \in C^1$ and $\rho(s) = s$. When we postulate only strict convexity at infinity, the possibility that $|\dot{x}|$ might reach level s by taking a jump at either t_0 or t_1 forces us to be conservative with condition (b). All we can assert in case of an endpoint jump to level s is that $|\dot{x}|$ must reach level $\rho(s)$ inside $[t_0, t_1]$. This principle is justified and used in the proof of Thm. 4.1, step 3.

It is immediate from the definition that $\Delta_R(r, s)$ is nonincreasing in r; we now verify that $\Delta_R(r, s)$ is eventually nondecreasing in s.

PROPOSITION 2.11. *For fixed* $r, R > 0$, *the mapping* $s \to \Delta_R(r, s)$ *is nondecreasing on* $(\rho^{-1}(r), +\infty)$.

PROOF. Consider any $s' \leq s$, with $r < \rho(s')$. We fix $\varepsilon > 0$ and select a Lipschitz arc x on an interval $[t_0, t_1]$ such that $t_1 - t_0 \leq \Delta_R(r, s) + \varepsilon$ and conditions (a)-(e) hold. From (a), Ess $\dot{x}(\tau) \cap r\overline{B} \neq \emptyset$ for some $\tau \in [t_0, t_1]$. Applying (e), we find that *co* Ess $\dot{x}(\tau) \subseteq W(\tau, x(\tau), p(\tau))$ for some $p(\tau)$. (cf. Prop. 2.8). For any σ such that $\rho(\sigma) > r$, it follows from the definition of $\rho_0(r)$ that $W(\tau, x(\tau), p(\tau))$ is contained in σB, whence *co* Ess $\dot{x}(\tau) \subseteq \rho^{-1}(r)\overline{B} \subseteq s'B$. Apply Prop. 2.7 to produce a subinterval $[t_0', t_1']$ of $[t_0, t_1]$ which is a neighbourhood of τ relative to $[t_0, t_1]$ and in which one has

(ii) Ess $\dot{x}(t) \subseteq s'B$ for $t \in (t_0', t_1')$

(iii) Either Ess $\dot{x}(t_0')$ contains points outside $s'B$ or else $t_0' = t_0$

(iv) Either Ess $\dot{x}(t_1')$ contains points outside $s'B$ or else $t_1' = t_1$.

Let us compare x on $[t_0', t_1']$ to the definition of $\Delta_R(r, s')$. Conditions (a′) and (d′) are inherited from the definition of $\Delta_R(r, s)$. Condition (c′) comes from (ii) above. The principle of optimality gives (e′) from (e), so only (b′) remains to check. Now if (b′) were false, then one would have Ess $\dot{x}(\sigma) \subseteq \rho(s')B \subseteq s'B$ for all $\sigma \in [t_0', t_1']$, whence $[t_0', t_1'] = [t_0, t_1]$ by (iii) and (iv). But then Ess $\dot{x}(\sigma) \subseteq \rho(s')B \subseteq \rho(s)B$ for all $\sigma \in [t_0, t_1]$ and (b) is contradicted. So (b′) must hold. Consequently $[t_0', t_1']$ and x obey (a′)-(e′) and $\Delta_R(r, s') \leq t_1' - t_0' \leq t_1 - t_0 \leq \Delta_R(r, s) + \varepsilon$. Since $\varepsilon > 0$ is arbitrary the proposition holds. ∎

3. - The main result

The next three sections contain the statement and proof of our main result. Throughout these sections, $[a, b]$ is a given interval and R, a given positive constant. We are also given a mapping $L : [a, b] \times R\overline{B} \times \mathbb{R}^n \to \mathbb{R}$. Writing $\Omega := [a, b] \times R\overline{B}$ as in §2, we assume the following basic hypotheses throughout:

(H1) L is Lipschitz on every compact subset of $\Omega \times \mathbb{R}^n$, and for every $(t, x) \in \Omega$, $L(t, x, \cdot)$ is convex

(H2) for every $(t, x) \in \Omega$, $L(t, x, \cdot)$ is strictly convex at infinity.

For any choice of endpoints $x_a, x_b \in RB$ we consider the following basic problem (P_R) in the calculus of variations:

$$(P_R) \qquad \min \left\{ \Lambda(x) := \int_a^b L(t, x(t), \dot{x}(t)) dt : x \in AC[a, b],\ x(a) = x_a,\ x(b) = x_b, \right. \\ \left. |x(t)| < R\ \forall t \in [a, b] \right\}$$

Note that problem (P_R) incorporates a state constraint requiring the solution to lie in the interior of RB. We recall that H and Δ_R were defined in §2.

3.1 *Statement*

THEOREM 3.1. *Let L satisfy* (H1) (H2), *and let m be a number satisfying* $\min(|x_a|, |x_b|) < m < R$. *Suppose that there is a Lipschitz arc $\overline{x}$ admissible for (P_R) such that*

(H3) *for some $\overline{\alpha} > 0$, for all $(t, x) \in \Omega$, for all p with $|p| \le \overline{\alpha}$, we have*

$$H(t, x, p) \le \frac{(R - m)\overline{\alpha} - \Lambda(\overline{x})}{b - a};$$

(H4) $b - a \le \Delta_R(r, s)$ *for some positive numbers r, s satisfying*

$$r > \frac{R - m}{b - a}, \quad \rho(s) > r.$$

Then problem (P_R) has at least one solution, and every solution to (P_R) is Lipschitzian.

Note that (H4) above and (H7) below are certainly satisfied if the extremal growth condition (ECG) (Def. 2.10) holds, a condition we shall study in §7. The following Lagrangian form of the theorem can be shown to be equivalent to the one above, although in appearance it seems more special.

THEOREM 3.2. *Let L satisfy* (H1), (H2) *as well as*

(H5) $\qquad L(t, x, v) \ge \overline{\alpha}|v|$ *for all* $(t, x, v) \in \Omega \times \mathbb{R}^n$ $(\overline{\alpha} > 0)$,

and suppose that for some α in $(0, \overline{\alpha})$ and some Lipschitz arc $\overline{x}$ admissible for

(P_R),

(H6) $$\frac{\Lambda(\overline{x})}{\alpha} + \min(|x_a|, |x_b|) < R,$$

(H7) $b - a \leq \Delta_R(\overline{r}, \overline{s})$ *for some* $\overline{s} > 0$ *obeying* $\rho(\overline{s}) > \overline{r}$, *where* $\overline{r} := \dfrac{\Lambda(\overline{x})}{\alpha(b-a)}$.

Then problem (P_R) *has at least one solution, and every solution to* (P_R) *is Lipschitzian.*

The proof of Theorem 3.2 takes up all of Sections 4 and 5. In fact, we establish a somewhat more general result. To prove that (P_R) has at least one Lipschitzian solution, we can weaken our hypothesis of strict convexity at infinity to the simple assumption that (2.3) holds when $r = \overline{r}$. That is, for some $M > 0$, one has

(3.1) $$W(t, x, p) \cap \overline{r}\, \overline{B} \neq \emptyset \text{ implies } W(t, x, p) \subseteq M\overline{B}.$$

Having chosen such an M, we will show that any Lipschitzian solution x obeys $\|\dot{x}\|_\infty \leq \overline{s}$, provided that $\overline{s} > \max\{M, \overline{r}, \|\dot{\overline{x}}\|_\infty\}$ satisfies (H7). And to obtain the additional conclusion that all solutions to (P_R) are Lipschitzian, we will assume only that for some such value of $\overline{s}$, (2.3) holds for $r = \overline{s}$. That is, for some $\mu > 0$, one has

(3.2) $$W(t, x, p) \cap \overline{s}\, \overline{B} \neq \emptyset \text{ implies } W(t, x, p) \subseteq \mu\overline{B}.$$

Let us now show how Theorem 3.1 follows from Theorem 3.2, the latter being the form we shall prove. The fact that Theorem 3.1 in turn implies Theorem 3.2 is left as an exercise. From (H3) we deduce that for any p with $|p| \leq \overline{\alpha}$, for any v, for any $(t, x) \in \Omega$, we have

$$L(t, x, v) + \frac{(R - m)\overline{\alpha} - \Lambda(\overline{x})}{(b - a)} \geq \langle p, v \rangle.$$

Let the left side of this inequality define $\overline{L}(t, x, v)$. Setting $p = \overline{\alpha}\dfrac{v}{|v|}$, we deduce that $\overline{L}$ satisfies (H5). Once we have proceeded to verify (H6) and (H7), Theorem 3.2 applied to $\overline{L}$ yields Theorem 3.1 for L, since L and $\overline{L}$ differ only by a constant. Letting $\overline{\Lambda}$ correspond to $\overline{L}$, we have $\overline{\Lambda}(\overline{x}) = (R - m)\overline{\alpha}$, so that (H6) becomes

$$(R - m)\frac{\overline{\alpha}}{\alpha} + \min(|x_a|, |x_b|) < R.$$

This will hold for any α sufficiently near $\overline{\alpha}$. Similarly, (H7) becomes

$$b - a \leq \Delta_R\left(\frac{(R - m)\overline{\alpha}}{(b - a)\alpha}, \overline{s}\right),$$

and this will follow from (H4) when α is close to $\overline{\alpha}$, as required.

3.2 *Examples*

EXAMPLE 1. We illustrate here the application of Theorem 3.1 to a problem in which the local level sets of Λ (see §1) fail to be compact. The problem (P_R) has data

$$n = 1,\ a = 0,\ x_a = x_b = 0,\ R = 2$$

$$L(t, x, v) = \sqrt{1+v^2} - \frac{x^2}{8} - 1$$

where for now b is an unspecified number in $\left(0, \frac{2}{3}\right)$. As for the parameters of Theorem 3.1, we define

$$m = 2 - 3b,\ \overline{\alpha} = \frac{1}{2},\ r = 4,\ \overline{x} \equiv 0.$$

The Hamiltonian is readily calculated:

$$H(t, x, p) = \begin{cases} \frac{x^2}{8} - \sqrt{1-p^2} + 1, & \text{if } |p| \leq 1 \\ +\infty, & \text{otherwise} \end{cases}$$

and it is easily checked that hypotheses (H1)-(H3) are satisfied. By a result to come (Lemma 6.1) we know that $\Delta_2(4, s)$ (defined relative to the interval $[0, \frac{2}{3}]$) is positive for s large. We choose such an $s > r$ and also specify $b < \Delta_2(4, s)$. It is easily seen from the definition that $\Delta_2(4, s)$ defined relative to the smaller interval $[0, b]$ majorizes $\Delta_2(4, s)$ defined relative to $[0, \frac{2}{3}]$, thus confirming (H4) and the applicability of Theorem 3.1. We deduce therefore the existence of a solution to this instance of problem (P_R).

We remark that, if the constraint $|x| < 2$ is deleted, the resulting (global) problem has no solution, for if x_i is the piecewise linear arc whose graph consists of the three line segments joining the points $(0,0)$, $(\frac{1}{i}, i)$, $(b - \frac{1}{i}, i)$, $(b, 0)$, then $\Lambda(x_i) \to -\infty$.

Our final comments concerning this example will show that *all* the relevant local level sets of Λ fail to be compact. We shall demonstrate that $\min(P_R) = 0$, that the unique solution to (P_R) is $\overline{x} \equiv 0$, and that for any positive numbers λ and ρ, the level set

$$\{x \in AC[0, b] : \Lambda(x) \leq \lambda,\ |x| \leq \rho,\ x(0) = x(b) = 0\}$$

fails to have the compactness properties invoked in the classical existence theories. Let x be an admissible arc with $\|x\|_\infty = c \in [0, 2)$. It is geometrically evident that we have

$$\int_0^b \sqrt{1+\dot{x}^2}\, dt \geq \sqrt{4c^2 + b^2},$$

from which we deduce that $\Lambda(x)$ majorizes $\sqrt{4c^2+b^2}-\frac{bc^2}{8}-b$. It is readily verified that this last expression is nonnegative for b and c in the indicated range, and vanishes only for $c=0$. These remarks prove that $\min(P_R)=0$, and that the unique solution to (P_R) is $\overline{x}\equiv 0$. Now let λ and ρ be positive numbers; choose any ε in $(0,\rho)$ such that $2\varepsilon-\frac{b\varepsilon^2}{8}<\lambda$. Consider the sequence of admissible arcs x_i whose graphs consist of the three line segments joining $(0,0)$, $(\frac{1}{i},\varepsilon)$, $(b-\frac{1}{i},\varepsilon)$, $(b,0)$. Routine calculation establishes that $\Lambda(x_i)$ tends to $2\varepsilon-\frac{b\varepsilon^2}{8}$, a number less than λ by choice of ε. Thus the tail of the sequence $\{x_i\}$ belongs to the level set

$$\{x\in AC[0,b]:\Lambda(x)\le\lambda,\ \ |x|\le\rho,\ \ x(0)=x(b)=0\}.$$

It is clear however that the sequence has no subsequence converging appropriately to an admissible arc x.

We now adduce examples illustrating the necessity of the various hypotheses to obtain the stated conclusion.

EXAMPLE 2. This will show that Theorem 3.2 is false in the sole absence of strict convexity at infinity. We set

$$n=1,\ \ a=0,\ \ b=6,\ \ x_a=x_b=0,\ \ R=2,\ \ \overline{\alpha}=1,$$
$$L(t,x,v)=|v|-|x|+x^2+\frac{1}{4}.$$

Observe that (H1) and (H5) hold. We claim that the extremal growth condition holds, in fact for any $r>0$, $\Delta_R(r,s)$ is $+\infty$ whenever $\rho(s)>0$ (whence (H7) holds). This will follow by showing that there exists no function x satisfying (a)-(e) in the definition of $\Delta_R(r,s)$ (Definition 2.9). To see this, we write the necessary conditions that any such x would satisfy in view of condition (e) (see [4, Theorem 4.4.3]): there exists an arc p on $[t_0,t_1]$ such that a.e. (among other things):

$$\dot p=\begin{cases}2x-1, & \text{if } x>0\\ 2x+1, & \text{if } x<0,\end{cases}$$

and such that

$$p\in\begin{cases}\{-1\}, & \text{if } \dot x<0\\ [-1,1], & \text{if } \dot x=0\\ \{1\}, & \text{if } \dot x>0.\end{cases}$$

Except for isolated points, the interval $[0,6]$ consists of intervals in which either $p=-1$, $p=+1$, or else $|p|<1$. It follows from the above that $\dot x=0$ on any such interval, whence x is constant, a contradiction of (a)-(e).

Finally, (H6) holds for $\alpha=\frac{7}{8}$, if we choose $\overline{x}\equiv 0$. Thus all the hypotheses of Theorem 3.2 hold, except that L fails to be strictly convex at infinity. To see that the conclusions of the theorem fail, we argue exactly as in estimating Δ_R above to deduce that the only possible Lipschitzian solution of (P_R) is

$x \equiv 0$. But calculation shows that (for example), the admissible arc y which has slope $1/2$ between 0 and 1, has slope $-1/2$ between 5 and 6, and which is equal to $1/2$ between 1 and 5 assigns a lower value to Λ than does $x \equiv 0$, a contradiction.

EXAMPLE 3. This example will demonstrate the need for (H6) in Theorem 3.2. We set

$$n = 1, \ a = 0, \ b = \frac{3}{2}\pi, \ x_a = x_b = 0, \ \overline{\alpha} = 1,$$

$$L(t, x, v) = v^2 - x^2 + R^2 + \frac{1}{4},$$

where R has been chosen large enough so that

$$\frac{3}{2}\pi\left(R^2 + \frac{1}{4}\right) > R.$$

We take $\overline{x} \equiv 0$, so that $\Lambda(\overline{x}) = \frac{3}{2}\pi\left(R^2 + \frac{1}{4}\right)$. Then (H1), (H2), (H5) are verified, but the inequality above implies that (H6) is not. To see that (H7) holds, observe first that all the extremals (solutions of the Euler equation) generated by L are of the form $x(t) = c \ \sin(t + k)$. Now any x described as in the definition of $\Delta_R(r, s)$ is necessarily an extremal, and we deduce that $|c|$ is at least $s - r$ (from (a), (b), noting $\rho(s) = s$). Further $x^2 + \dot{x}^2$ is always equal to c^2, so at some point τ in $[t_0, t_1]$ we must have $x(\tau)^2 \geq c^2 - r^2$, in view of (a). This yields

$$x(\tau)^2 \geq (s - r)^2 - r^2,$$

which will imply $|x(\tau)| > R$ if s is chosen big enough, a contradiction. This proves that $\Delta_R(r, s) = +\infty$ for s sufficiently large, so that (H7) holds.

Let us now verify that the conclusion of Theorem 3.2 fails. If x were a Lipschitz solution to (P_R), then x would be an admissible extremal. But the only such is $x \equiv 0$. On the other hand, $x \equiv 0$ fails to be a local minimum for Λ, since the point π is conjugate to 0 by the classical calculus of variations (see [1], [4]). This is the required contradiction.

EXAMPLE 4. We demonstrate now the necessity of (H7) for Theorem 3.2, using a version of the Ball-Mizel problem. We set

$$n = 1, \ a = 0, \ b = 1, \ x_a = 0, \ x_b = k > 0, \ \overline{\alpha} = 1,$$

$$L(t, x, v) = \frac{1}{4\varepsilon} + \varepsilon v^2 + (x^3 - t^2)^2 v^{14}.$$

It is easily seen that (H1), (H2), (H5) hold. It is shown in [7] that for suitable choices of the positive constants ε and k, the unique global minimizer for Λ over $AC[0, 1]$ subject to the given boundary condition is

$$\hat{x}(t) := kt^{2/3}.$$

Let us take any suitable $\bar{x}$ and then choose R large enough so that (H6) holds (for some appropriate choice of α), and also so that $|\hat{x}(t)| < R$ for all t in $[0,1]$. Then all the hypotheses of Theorem 3.2 are present (except (H7)), yet the conclusion fails. For the only solution to (P_R) is $\hat{x}$ itself, and $\hat{x}$ is not Lipschitzian. This example hints at the rôle that (H7) can play as a tool to deduce regularity of the solution, a theme to be developed in section 7.

4. - Existence in the Lipschitz problem

The proof of Theorem 3.2 proceeds in two main stages. The first of these, described in this section, is to find a solution for the Lipschitz version of problem (P_R). This auxiliary problem, which we denote (P_R^∞), is identical to (P_R) in every way except that the minimization is over $AC^\infty[a,b]$ instead of $AC[a,b]$. The second stage, set out in Section 5, is to prove that the solution we find for (P_R^∞) is also optimal in (P_R).

THEOREM 4.1. *Under* (H1), (H5)-(H7) *and* (3.1), *problem* (P_R^∞) *has a solution.*

PROOF. Suppose $\alpha \in (0,\bar{\alpha})$ and endpoints x_a and x_b are given for which (H1), (H5)-(H7) and (3.1) hold. Let $\bar{x}$ and $\bar{r}$ be the feasible Lipschitz arc and constant participating in (H7) and (3.1).

Step 1. *Introduce Auxiliary Problems*

Minor refinements of [9, Lemma 5.1] and [8, Prop. 2.1] imply that for every $s > 0$ there exists an auxiliary Lagrangian $L_s : \Omega \times \mathbb{R}^n \to \mathbb{R}$ with the following properties:

(a) L_s obeys (H1) and (H2);

(b) $L_s(t,x,v) = L(t,x,v)$, for all $(t,x,v) \in \Omega \times s\overline{B}$;

(c) $L_s(t,x,v) \geq \max\{\alpha|v|, |v|^2 - s^2\}$, for all $(t,x,v) \in \Omega \times \mathbb{R}^n$;

(d) There is a finite number $\nu(s) > 0$ such that

$$L_s(t,x,v) = |v|^2 - s^2, \text{ for all } (t,x) \in \Omega,\ |v| \geq \nu(s).$$

In terms of these auxiliary Lagrangians, we formulate the auxiliary problems:

$$(P^s)\quad \min\left\{\Lambda_s(x) := \int_a^b L_s(t,x(t),\dot{x}(t))dt : x \in AC[a,b],\ x(a) = x_a,\ x(b) = x_b,\ |x(t)| \leq R\ \forall t \in [a,b]\right\}.$$

Note that the minimization here is over all absolutely continuous functions (not just the Lipschitz ones), and that the state constraint set $R\overline{B}$ is closed. Both these features are critical in the classical existence theory.

Step 2. *Existence and Regularity of Interior Solutions*

Every problem (P^s), $s > 0$, has an admissible arc simply because (H1) holds for L. Since each Lagrangian L_s, $s > 0$, is jointly continuous in all its arguments while being both convex and quadratically coercive in v, Tonelli's classical existence theorem [1, 2.20.i] implies that each problem (P^s) has a solution x_s.

Note that whenever $s \geq \|\dot{\overline{x}}\|_\infty$, Step 1 (b) implies

$$\Lambda_s(\overline{x}) := \int_a^b L_s\left(t, \overline{x}(t), \dot{\overline{x}}(t)\right) dt = \Lambda(\overline{x}),$$

whence

$$\begin{aligned}(4.1) \qquad \Lambda(\overline{x}) = \Lambda_s(\overline{x}) \geq \Lambda_s(x_s) \geq \int_a^b \alpha|\dot{x}_s(t)|dt \\ \geq \alpha \max\left\{\|x_s - x_a\|_\infty, \|x_s - x_b\|_\infty\right\}.\end{aligned}$$

Consequently $\|x_s\|_\infty \leq \frac{\Lambda(\overline{x})}{\alpha} + \min\{|x_a|, |x_b|\}$ for every $s \geq \|\dot{\overline{x}}\|_\infty$ and (H6) implies that x_s lies in the interior of the state constraint region $R\overline{B}$. In particular, x_s must provide a strong local solution for the unconstrained version of (P^s).

For fixed s, the auxiliary Lagrangian L_s is independent of (t, x) outside the compact set $\Omega \times \nu(s)\overline{B}$. And inside this compact set, $L_s(\cdot,\cdot,\cdot)$ is Lipschitz. Hence some large ball contains all the sets $\partial_{(t,x)}L_s(t, x, v)$, $(t, x, v) \in \Omega \times \mathbb{R}^n$. By [4, Prop. 3.1], it follows that x_s is Lipschitz on $[a, b]$ and satisfies the Euler-Lagrange inclusion for L_s (see [2]). Of course, the Lipschitz rank of x_s may depend on s. But in any case, for each $s > 0$ the multifunction Ess $\dot{x}_s$ is upper semicontinuous, with nonempty compact values.

Step 3. *Apply Necessary Conditions*

Calculation (4.1) above implies that (when s is sufficiently large)

$$(b-a)^{-1} \int_a^b |\dot{x}_s(t)|dt \leq \overline{r}.$$

Hence for each $s \geq \|\dot{\overline{x}}\|_\infty$ there is some point $\tau \in [a, b]$, depending on s, such that

$$(4.2) \qquad \text{Ess } \dot{x}_s(\tau) \cap \overline{r}\,\overline{B} \neq \emptyset.$$

The Euler-Lagrange inclusion for L_s certainly implies that the s-version of inclusion (2.6) holds, from which we deduce via Proposition 2.8 that, for the adjoint arc p_s,

$$\overline{co}\ \text{Ess}\ \dot{x}_s(t) \subseteq W_s\left(t, x_s(t), p_s(t)\right), \text{ for all } t \in [a, b]. \tag{4.3}$$

Observe that for any $(t, x, p) \in [a, b] \times R\overline{B} \times \mathbb{R}^n$, one has

$$W_s(t, x, p) \cap rB \subseteq W(t, x, p), \text{ for all } r \in (0, s]. \tag{4.4}$$

Indeed, if v belongs to the left side then Step 1 (b) implies

$$p \in \partial_v L_s(t, x, v) = \partial_v L(t, x, v),$$

so that v belongs to the right side.

Together (4.2) and (4.3) imply that $W_s(\tau, x_s(\tau), p_s(\tau))$ contains some point – say w – of $\overline{r}\ \overline{B}$.

Provided $s > \overline{r}$, (4.4) implies that w also lies in $W(\tau, x_s(\tau), p_s(\tau))$. Making use of condition (3.1), we deduce from (4.4) that

$$W_s(\tau, x_s(\tau), p_s(\tau)) \cap sB \subseteq W(\tau, x_s(\tau), p_s(\tau)) \subseteq M\overline{B}.$$

Finally, since $W(\tau, x_s(\tau), p_s(\tau))$ is a closed convex set, it follows that whenever $s > \max\{M, \overline{r}, \|\dot{\overline{x}}\|_\infty\}$ we have $W_s(\tau, x_s(\tau), p_s(\tau)) \subseteq M\overline{B}$, and (4.3) implies

$$\text{Ess}\ \dot{x}_s(\tau) \subseteq M\overline{B}. \tag{4.5}$$

Step 4. Estimate $\|\dot{x}_s\|_\infty$

Let us now choose any $\sigma > \max\{M, \overline{r}, \|\dot{\overline{x}}\|_\infty\}$ such that $\Delta_R(\overline{r}, \sigma) \geq b - a$. Such a choice is possible by (H7) and Prop. 2.11. We will show that

$$s > \sigma \text{ implies Ess}\ \dot{x}_s(t) \subseteq \sigma B, \text{ for all } t \in (a, b). \tag{4.6}$$

Indeed, fix any $s > \sigma$ and apply Proposition 2.7 to the arc x_s and point τ identified in Step 3. The result is a nonempty interval $[t_0, t_1] \subseteq [a, b]$ containing a neighbourhood of τ relative to $[a, b]$, on which Proposition 2.7 (i)-(iv) hold. Noting conclusion (ii), we see that (4.6) will follow if we can show $t_0 = a$, $t_1 = b$. To do this, let us compare the properties of x_s and $[t_0, t_1]$ with conditions (a)-(e) defining $\Delta_R(\overline{r}, \sigma)$ (Definition 2.9). We have shown above that x_s is a Lipschitz arc, for which condition (a) holds by (4.2). Condition (c) follows from Proposition 2.7 (ii), and condition (d) is verified in Step 2. As for condition (e), the principle of optimality implies that x_s solves

$$\min \left\{ \int_{t_0}^{t_1} L_s(t, y, \dot{y}) dt : y(t_i) = x(t_i), \ |y(t)| < R\ \forall t \in [t_0, t_1] \right\}.$$

But we have $|\dot{x}_s(t)| < \sigma < s$ a.e. in $[t_0, t_1]$ by Proposition 2.7 (ii), so x_s continues to solve the above problem in the presence of the constraint $|\dot{y}(t)| \leq s$ a.e. in $[t_0, t_1]$. Along every arc obeying this condition, L_s coincides with L. Thus x_s satisfies condition (e) defining $\Delta_R(\overline{r}, \sigma)$. Finally we turn to condition (b). If (b) is satisfied, then we have $t_1 - t_0 \geq \Delta_R(\overline{r}, \sigma) \geq b - a$ by our choice of σ, and the desired conclusion follows.

Suppose therefore that (b) is false. We will show that this implies $t_0 = a$; the proof that $t_1 = b$ is similar. If (b) fails, then we have $|v| < \rho(\sigma)$ for any $v \in \text{Ess } \dot{x}_s(t_0+)$. Fix any such v, and note that $v \in \text{Ess } \dot{x}_s(t_0)$. By (4.3) and (4.4), it follows that $v \in W(t_0, x_s(t_0), p_s(t_0))$ (recall $\rho(\sigma) \leq \sigma$); the definition of ρ and Proposition 2.4 (ii) then imply that $W(t_0, x_s(t_0), p_s(t_0)) \subseteq \sigma B$. Applying (4.3) and (4.4) again, we find

$$\overline{co} \text{ Ess } \dot{x}_s(t_0) \cap sB \subseteq \sigma B.$$

Since $s > \sigma$ by assumption, it follows that $\overline{co}$ Ess $\dot{x}_s(t_0) \subseteq \sigma B$, and this in turn forces $t_0 = a$ by Proposition 2.7 (iii).

Thus, whether (b) is true or false, (4.6) holds.

Step 5.

Combining (4.6) with Step 1(b), we find that $\Lambda_s(x_s) = \Lambda(x_s)$ for all $s > \sigma$. Thus for any pair $r, s > \sigma$ the optimality of x_r in (P^r) implies

$$\Lambda(x_r) = \Lambda_r(x_r) \leq \Lambda_r(x_s) = \Lambda(x_s).$$

Clearly r and s are interchangeable here, so it follows that all the values $\Lambda(x_s)$ for $s > \sigma$ are equal.

We now claim that x_s solves (P_R^∞). Indeed, choose any arc $x \in AC^\infty[a, b]$ admissible for (P_R^∞) and fix $s > \max\{\|\dot{x}\|_\infty, \sigma\}$. Then by the optimality of x_s for (P^s) and Step 1(b), we find

$$\Lambda(x_s) = \Lambda_s(x_s) \leq \Lambda_s(x) = \Lambda(x).$$

Thus $\Lambda(x_s) = \inf\{\Lambda(x) : x \in AC^\infty[a, b],\ x(a) = x_a,\ x(b) = x_b\}$ as required. Moreover, condition (4.6) proves that $\|\dot{x}_s\|_\infty \leq \sigma$. ∎

Robustness

The arguments and conclusions of this section are unperturbed by certain changes to the Lagrangian L. Indeed, let us consider the same L, α, x_a, and x_b discussed in the preceding proof. For any $r > 0$, the values of $\Delta_R(r, \cdot)$ in the interval (r, s) are completely determined by the values of L on the set $\Omega \times s\overline{B}$. So once $\sigma > \max\{M, \overline{r}, \|\dot{\overline{x}}\|_\infty\}$ has been chosen so that $b - a \leq \Delta_R(\overline{r}, \sigma)$, the key quantities $M, \overline{r}$, and $\Delta_R(\overline{r}, \sigma)$ remain unchanged for any Lagrangian $\tilde{L}$ which coincides with L on $\Omega \times \sigma\overline{B}$. In particular, if $\tilde{L}$ also obeys (H1) and (H5) (for the same $\overline{\alpha}$, R, and $[a,b]$) then the arc x_s of Step 5 above will also solve $(\tilde{P}_R^\infty)$.

To verify this, first apply Thm. 4.1 to $\tilde{L}$ to obtain a solution $\tilde{x}$ for $(\tilde{P}_R^\infty)$ which obeys $\|\dot{\tilde{x}}\|_\infty \leq \sigma$. Then observe that

$$\Lambda(x_s) = \tilde{\Lambda}(x_s) \geq \tilde{\Lambda}(\tilde{x}) = \Lambda(\tilde{x}) \geq \Lambda(x_s).$$

Thus $\Lambda(\tilde{x}) = \Lambda(x_s)$, as claimed.

5. - On the Lavrentiev phenomenon

We must now prove that the solution to (P_R^∞) described in Section 4 remains optimal among the much larger class of arcs admissible for (P_R). This could only fail to be the case if the infimum of $\Lambda(\cdot)$ over the absolutely continuous functions were strictly less than its infimum over the Lipschitz functions. This pathology, called the *Lavrentiev phenomenon*, has been observed in a variety of reasonable-looking problems - see [1, Section 18.5], [8]. In this section we show that the hypotheses (H1), (H5)-(H7), and (3.1) of Section 3 preclude the Lavrentiev phenomenon, and that if (3.2) holds then all solutions to (P_R) are actually Lipschitz.

In the presence of (H1) and (H5), the Lavrentiev phenomenon cannot occur in problems whose Lagrangian grows at most linearly in v. Here is a general result of this sort, due to Tonelli [14, Vol. I, 141(b)].

PROPOSITION 5.1. *Let* $L : \Omega \times \mathbb{R}^n \to \mathbb{R}$ *obey* (H1) *and* (H5), *and suppose there exist constants* $\beta, \lambda, \nu \geq 0$ *such that if*

$$(t,x) \in \Omega, \quad |v| \geq \nu \text{ then } |L(t,x,v)| \leq \beta|v| + \lambda.$$

Then for any $x \in AC[a,b]$ *with* $\|x\|_\infty < R$, *there is a sequence* $\{x_i\}$ *in* $AC^\infty[a,b]$ *obeying*

(a) $x_i(a) = x(a), \ x_i(b) = x(b)$

(b) $x_i \to x$ *uniformly on* $[a,b]$

(c) $\dot{x}_i \to \dot{x}$ *a.e. and in* $L^1[a,b]$

(d) $\Lambda(x_i) \to \Lambda(x)$.

The remainder of this section completes the proof of Thm. 3.2. Thus, let us assume that (H1) and (H5)-(H7) hold, that some $\alpha \in (0,\overline{\alpha})$ and endpoints x_a and x_b in RB are given, and that (3.1)-(3.2) hold. Fix any $\overline{s}$ as in (H7) obeying $\overline{s} > \max\{M, \overline{r}, \|\dot{\overline{x}}\|_\infty\}$. Later in this section, we will produce constants $\mu > \overline{s}$, $\varepsilon \geq 0$, β, and λ and a family of auxiliary Lagrangians $L_k : \Omega \times \mathbb{R}^n \to \mathbb{R}$ with the following properties. For each $k \geq \mu$, $(t,x) \in \Omega$,

(i) L_k obeys (H1) and (H5), except that α replaces $\overline{\alpha}$

(ii) $|v| \leq \overline{s} \Rightarrow L_k(t, x, v) = L(t, x, v)$

(iii) $\overline{s} < |v| \leq \mu \Rightarrow L_k(t, x, v) \leq L(t, x, v)$

(iv) $\mu < |v| \leq k \Rightarrow L_k(t, x, v) \leq L(t, x, v) - \varepsilon$

(v) $k < |v| \Rightarrow L_k(t, x, v) \leq \beta|v| + \lambda$.

First, however, we show how these auxiliary elements will complete the proof of Theorem 3.2.

THEOREM 5.2. *As described above, assume* (H1), (H5)-(H7) *and* (3.1). *Then there exists a family of auxiliary Lagrangians obeying* (i)-(v) *above with* $\varepsilon = 0$, *and any solution of* (P_R^∞) *also solves* (P_R). *If* (3.2) *also holds, then there exists a family of auxiliary Lagrangians obeying* (i)-(v) *above with* $\varepsilon > 0$. *In this case all solutions of* (P_R) *are Lipschitz arcs.*

PROOF. The existence of suitable auxiliary Lagrangians is taken up later in this section. Here we show how conditions (i)-(v) imply the other conclusions of Thm. 5.2.

Each auxiliary Lagrangian L_k gives rise to a corresponding problem $(P_R^{k,\infty})$ with the same endpoint constraints as (P_R^∞). According to the robustness remarks at the end of Section 4, the agreement of L_k with L on $\Omega \times \sigma B$ (for $k > \mu$) implies that some fixed solution y of (P_R^∞) also solves $(P_R^{k,\infty})$: that is, for any $x \in AC^\infty[a, b]$ with $\|x\|_\infty < R$, $x(a) = x_a$, $x(b) = x_b$, one has

$$\Lambda(y) = \Lambda_k(y) \leq \Lambda_k(x). \tag{5.1}$$

Suppose now that we are given a *non-Lipschitz* arc x satisfying the endpoint conditions. Fix $k > \mu$ and apply Prop. 5.1 to obtain a sequence $\{x_i\}$ of admissible Lipschitz arcs such that

$$\Lambda_k(x) = \lim_{i \to \infty} \Lambda_k(x_i).$$

In view of (5.1), it follows from this that

$$\Lambda(y) \leq \Lambda_k(x), \text{ for all } k > \mu. \tag{5.2}$$

Now the arc x is not Lipschitz, so the set $E = \{t \in [a, b] : |\dot{x}(t)| > \mu\}$ has positive measure. Write $U = [a, b] \backslash E$, and partition $E = F_k \cup G_k$, where

$$F_k = \{t \in [a, b] : \mu < |\dot{x}(t)| \leq k\}$$
$$G_k = \{t \in [a, b] : k < |\dot{x}(t)|\}.$$

Then

$$\Lambda_k(x) = \int_{U \cup F_k \cup G_k} L(t, x, \dot{x})\mathrm{d}t$$

$$\leq \int_U L(t, x, \dot{x})\mathrm{d}t + \int_{F_k} L(t, x, \dot{x})\mathrm{d}t - \varepsilon m(F_k) + \int_{G_k} (\beta|\dot{x}(t)| + \lambda)\,\mathrm{d}t.$$

As $k \to \infty$, the right side tends to $\Lambda(x) - \varepsilon m(E)$, so (5.2) gives

$$\Lambda(y) \leq \Lambda(x) - \varepsilon m(E). \tag{5.3}$$

Now if $\varepsilon = 0$, then (5.3) shows that the solution y of (P_R^∞) does at least as well as any non-Lipschitz arc x, and hence solves (P_R). And if $\varepsilon > 0$, then any non-Lipschitz arc x has $\Lambda(x) > \Lambda(y)$ by (5.3), so all solutions to (P) must be Lipschitz. ■

The remainder of this section is devoted to the construction of the auxiliary Lagrangians described above. We shall first produce a suitable family with $\varepsilon > 0$ by assuming (3.2), and then show how (3.2) may be dropped if only $\varepsilon = 0$ is required. The only use of (3.2) is to support the following Lemma.

LEMMA 5.3. *Assume* (3.2). *Then the constant η defined below is positive*:

$$\eta = \inf\{L(t, x, w) - L(t, x, v) - \langle p, w - v\rangle : (t, x) \in \Omega, |v| \leq \bar{s},$$
$$p \in \partial_v L(t, x, v), |w| \geq \mu\}.$$

PROOF. It is an exercise in convex analysis to show that the infimum defining η can actually be taken over only $|w| = \mu$. Then it is easy to recognize η as the infimum of a continuous function over a compact set of (t, x, v, p, w)-values. It therefore suffices to prove that one has

$$L(t, x, w) - L(t, x, v) - \langle p, w - v\rangle > 0$$

for all $(t, x) \in \Omega$, $|v| \leq \bar{s}$, $p \in \partial_v L(t, x, v)$ and $|w| = \mu$. To prove this, note that the left side is nonnegative by the subgradient inequality. If the left side were to equal zero, one could easily derive that $p \in \partial_v L(t, x, v) \cap \partial_v L(t, x, w)$. In other words, both v and w would lie in $W(t, x, p)$, contradicting (3.2). ■

LEMMA 5.4. *Given any $\mu > \bar{s}$, there are constants $\beta > 0, \lambda$, and $\rho > \mu$ such that the function*

$$F(t, x, v) = \max\{L(t, x, v)\chi\{|v| \leq \rho\},\ \beta|v| + \lambda\}$$

(χ *denotes characteristic function*) *obeys* (H1) *and* (H5) *and satisfies, for all* $(t, x) \in \Omega$:

(a) *if* $|v| \leq \bar{s}$ *then* $F(t, x, v) = L(t, x, v)$*;*

(b) *if* $\mu \leq |v|$ *then* $F(t, x, v) = \beta|v| + \lambda$.

PROOF. See [5, Lemma 6.2]. ∎

Now assume (3.2) and recall the positive η supplied by Lemma 5.3. Choose any $\varepsilon \in (0, \eta)$ for which $\varepsilon < (\bar{\alpha} - \alpha)\mu$.

LEMMA 5.5. *Assume* (3.2). *Then for each* $k > \mu$ *there exists* $F_k : \Omega \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *such that*

(a) $F_k(t, x, v)$ *is Lipschitz on* $\Omega \times k\overline{B}$ *and convex in* v

(b) $F_k(t, x, v) \geq \alpha|v|$, *for all* $(t, x, v) \in \Omega \times \mathbb{R}^n$

(c) $|v| \leq \bar{s} \Rightarrow F_k(t, x, v) = L(t, x, v)$

(d) $|v| \leq k \Rightarrow F_k(t, x, v) \leq L(t, x, v)$

(e) $\mu < |v| \leq k \Rightarrow F_k(t, x, v) \leq L(t, x, v) - \varepsilon$

(f) $k < |v| \Rightarrow F_k(t, x, v) = +\infty$.

PROOF. Introduce $\tilde{f}_k : \Omega \times \mathbb{R}^n \to \mathbb{R}^n$ as the convexification in v of the function

$$f_k(t, x, v) = \begin{cases} L(t, x, v) & \text{if } |v| \leq \mu \\ L(t, x, v) - \varepsilon & \text{if } \mu < |v| \leq k \\ +\infty & \text{if } k < |v|. \end{cases}$$

Then let F_k equal $\tilde{f}_k$ when $|v| \leq k$ and $+\infty$ otherwise. Verification of (a)-(f) is given in [5, Lemma 6.3].

Under assumption (3.2), we may now produce the promised auxiliary Lagrangians by defining, for each $(t, x) \in \Omega$,

$$L_k(t, x, \cdot) = co\{F(t, x, \cdot),\ F_k(t, x, \cdot)\}.$$

That is,

$$L_k(t, x, v) = \inf\{\lambda F(t, x, v_1) + (1-\lambda)\ F_k(t, x, v_0) : \lambda \in [0, 1],\ v = \lambda v_1 + (1-\lambda) v_0\}.$$

Since $F_k(t, x, \cdot) = F(t, x, \cdot) = L(t, x, \cdot)$ on $\bar{s}\ \overline{B}$, it follows that $L_k(t, x, \cdot) = L(t, x, \cdot)$ on $\bar{s}\ \overline{B}$ ([5, Lemma 6.4]). Also, L_k is convex in v by construction, as well as locally Lipschitz on $\Omega \times \mathbb{R}^n$. Since both F and F_k majorize $\alpha|v|$, so does L_k. And since the convex hull of two functions is majorized by the two functions themselves, the desired properties (iii)-(v) follow from the conclusions of Lemmas 5.4 and 5.5. Note that $\varepsilon > 0$ in Lemma 5.5.

Suppose, finally, that (3.2) is unavailable. In this case Lemma 5.3 cannot be used, so we simply take $\mu = \bar{s} + 1$ and apply Lemma 5.4. While the full force of Lemma 5.5 is unavailable, we note that the functions $F_k(t, x, v)$ which equal $L(t, x, v)$ when $|v| \leq k$ and $+\infty$ otherwise satisfy all the conclusions of Lemma 5.5. with $\varepsilon = 0$. Given these preliminary choices, the definition of L_k given in the previous paragraph once again leads to a family of auxiliary Lagrangians obeying (i)-(v), except that this time $\varepsilon = 0$.

This completes the proof of Theorem 3.2. ∎

6. - Existence in the small

We may now present the promised generalization of [9, Thm. 2.1], in which the assumption of strict convexity is replaced by the much weaker requirement of strict convexity at infinity. The key arguments appear in our first Lemma, which concerns (H7).

LEMMA 6.1. *Assume* (H1)-(H2). *Then for any* $r > 0$ *one has* $\Delta_R(r, s) > 0$ *for all* s *sufficiently large.*

PROOF. For any $r > 0$ and $s > 0$, we will prove that if

$$\rho(\rho(s)) > r \text{ then } \Delta_R(r, s) > 0. \tag{6.1}$$

Since $\rho(s)$ is an unbounded and nondecreasing function (Prop. 2.4 (iv)), the desired conclusion will then follow.

Suppose (6.1) is false. That is, fix positive constants r and s for which $\rho(\rho(s)) > r$ but $\Delta_R(r, s) = 0$. Then the definition of Δ_R implies that for each i there must be an interval $[t_{0i}, t_{1i}] \subseteq [a, b]$ with $t_{1i} - t_{0i} < \frac{1}{i}$ and an arc $x_i \in AC[t_{0i}, t_{1i}]$ for which conditions (a)-(e) of Def. 2.9 hold. In particular x_i solves a certain optimal control problem on $[t_{0i}, t_{1i}]$, and necessary conditions such as [4, Thm. 5.2.1] apply. Condition 2.9 (c) together with these conditions implies that there is an arc p_i on $[t_{0i}, t_{1i}]$ for which

$$\dot{p}_i(t) \in \partial_x L(t, x_i(t), \dot{x}_i(t)) \text{ and } p_i(t) \in \partial_v L(t, x_i(t), \dot{x}_i(t)) \text{ a.e.} \tag{6.2}$$

Relying upon 2.9 (a), we find a point $\tau_i \in [t_{0i}, t_{1i}]$ where (6.2) holds, $\dot{x}_i(\tau_i)$ exists, and $|\dot{x}_i(\tau_i)| \leq r + \frac{1}{i}$. Condition 2.9 (b) furnishes a point $\sigma_i \in [t_{0i}, t_{1i}]$ where Ess $\dot{x}_i(\sigma_i) \cap (\mathbb{R}^n \backslash \rho(s)B) \neq \emptyset$. According to Prop. 2.7, $\overline{co}$ Ess $\dot{x}_i(\sigma_i)$ contains a point w_i for which $\rho(s) \leq |w_i| \leq s$.

For each i, condition 2.9 (c) implies that

$$|x_i(\sigma_i) - x_i(\tau_i)| \leq s|\sigma_i - \tau_i| \leq \frac{s}{i}.$$

Since $|x_i(\tau_i)| < R$ for all i by (e), it follows that along a suitable subsequence both $x_i(\tau_i)$ and $x_i(\sigma_i)$ converge to a common limit x. We may also assume that $\dot{x}_i(\tau_i)$ converges to some v with $|v| \leq r$, that $w_i \to w$ for some w outside $\rho(s)B$, and that both σ_i and τ_i converge to some point $t \in [a, b]$.

Defining

$$M = \sup\{|\varsigma| : \varsigma \in \partial_x L(t, x, v),\ (t, x) \in \Omega,\ |v| \leq s\},$$

we deduce from (H1) that $M < +\infty$, while inclusion (6.2) and condition 2.9 (c) imply that Ess $\dot{p}_i(t) \subseteq M\overline{B}$ for all $t \in (t_{0i}, t_{1i})$. Since we also have $p_i(\tau_i) \in \partial_v L(\tau_i, x_i(\tau_i), \dot{x}_i(\tau_i))$, with the right side bounded independently of i, we may pass to a further subsequence along which $p_i(\sigma_i)$ and $p_i(\tau_i)$ converge to a common limit p.

Consider now the two inclusions

$$\dot{x}_i(\tau_i) \in W(\tau_i, x_i(\tau_i), p_i(\tau_i)) \text{ and } w_i \in W(\sigma_i, x_i(\sigma_i), p_i(\sigma_i)),$$

which follow from (6.2) and Prop. 2.8, respectively. Taking the limit as $i \to \infty$, Prop. 2.3 gives

$$v \in W(t, x, p) \text{ and } w \in W(t, x, p). \tag{6.3}$$

But $|v| \leq r$ and $|w| \geq \rho(s)$. Thus (6.3) implies $\rho(\rho(s)) \leq r$, a contradiction. This confirms (6.1) and completes the proof. ∎

The results of this section are called *local* existence theorems because they apply to problems posed on sufficiently small subsets of the region Ω where L is defined. For a given interval $[a', b'] \subseteq [a, b]$, we denote by (P'_R) the problem

$$(P'_R) \quad \min\left\{\Lambda'(x) = \int_{a'}^{b'} L(t, x(t), \dot{x}(t))\,dt : x(a') = x'_a, x(b') = x'_b, \right.$$
$$\left. |x(t)| < R \ \forall t \in [a', b']\right\}.$$

To pass from (P_R) to (P'_R), it suffices to replace Ω by its subset $\Omega' = [a', b'] \times R\overline{B}$. Note that the definition of Δ_R depends implicitly on the basic interval $[a, b]$, so that the appropriate analogue for problem (P'_R) is a function Δ'_R. It is evident that $\Delta'_R(r, s) \geq \Delta_R(r, s)$ for every r, s. Shrinking the domain of the problem also affects the hypotheses: we emphasize that (H1)-(H7) introduced above always refer to the given set Ω, which we regard as fixed throughout what follows. When it becomes necessary to restrict these hypotheses to Ω', we will denote them by (H1′)-(H7′).

Our first local existence result makes explicit reference to the linear growth hypothesis (H5). This highlights its relationship to the intermediate existence Theorem 3.2, while simultaneously taking a big step toward the generalization of [9, Thm. 2.1] to be given later.

THEOREM 6.2. *Assume* (H1), (H2) *and* (H5). *Then for every* $M > 0$ *and* $\rho \in (0, R)$, *there exists* $\delta > 0$ *with the following properties. For every pair of endpoints* (a', x'_a) *and* (b', x'_b) *in* Ω *obeying*

$$0 < b' - a' < \delta, \quad |x'_b - x'_a| \leq M(b' - a'), \quad \min\{|x'_a|, |x'_b|\} < \rho, \tag{6.4}$$

problem (P_R') has a Lipschitz solution. In fact, all the solutions of (P_R') are Lipschitz.

PROOF. We will apply Thm. 3.2 to the problem (P_R') arising as in the theorem's statement. Let any $M > 0$ and ρ in $(0, R)$ be given, and define

$$\gamma = \max\{L(t,x,v) : (t,x) \in \Omega, |v| \leq M\}.$$

Given any $\overline{\alpha}$ obeying (H5), fix $\alpha \in (0, \overline{\alpha})$. Then define

$$\delta = \min\left\{\frac{\alpha(R-\rho)}{\gamma}, \lim_{s\to\infty} \Delta_R\left(\frac{\gamma}{\alpha}, s\right)\right\}.$$

Note that $\delta > 0$ by Lemma 6.1 and Proposition 2.11.

Now fix any endpoints (a', x_a') and (b', x_b') in Ω obeying (6.4). In $\Omega' = [a', b'] \times R\overline{B}$, hypotheses (H1'), (H2'), and (H5') follow immediately from (H1), (H2), and (H5). As for (H6'), we may choose the Lipschitz admissible arc

$$\overline{x}(t) = x_a' + (x_b' - x_a')\,\frac{t-a'}{b'-a'}$$

and observe that $\Lambda'(\overline{x}) \leq \gamma(b'-a')$. Since $b' - a' < \delta \leq \frac{\alpha(R-\rho)}{\gamma}$, we obtain

$$\frac{\Lambda'(\overline{x})}{\alpha} + \rho < R,$$

from which (H6') follows via (6.4).

To check (H7'), we observe that $\overline{r}' = \frac{\Lambda'(\overline{x})}{\alpha(b'-a')} \leq \frac{\gamma}{\alpha}$, so

$$\Delta_R'(\overline{r}', s) \geq \Delta_R'\left(\frac{\gamma}{\alpha}, s\right) \geq \Delta_R\left(\frac{\gamma}{\alpha}, s\right).$$

Our choice of δ implies that $\Delta_R\left(\frac{\gamma}{\alpha}, s\right) > b' - a'$ for all s sufficiently large, so (H7') follows.

All the hypotheses of Thm. 3.2 hold for problem (P_R'), and the conclusions of that result complete the present proof. ∎

Note that if $L(t,x,\cdot)$ happens to be strictly convex for each $(t,x) \in \Omega$, then (H2) certainly holds. Hence the conclusions of Thm. 6.2 are available, and the solution set of (P_R') for each appropriately chosen pair of endpoints is a nonempty subset of AC^∞. Every Lipschitzian solution is an extremal in the sense of [2], and must therefore be smooth by the remarks following Prop. 2.8. Thus in the strictly convex case, the conclusions of Thm. 6.2 remain valid when "continuously differentiable" is written in place of "Lipschitz".

We turn now to a local existence theorem in the spirit of [9]. Here Ω is not assumed to be fixed in advance – all we are given are a point $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$, a neighbourhood U of (t_0, x_0), and a function $L : U \times \mathbb{R}^n \to \mathbb{R}^n$ satisfying the conditions

(6.5) $\quad L$ is locally Lipschitz on $U \times \mathbb{R}^n$;

(6.6) $\quad L(t, x, \cdot)$ is strictly convex at infinity for each $(t, x) \in U$.

Here is the result.

THEOREM 6.3. *Under the conditions of the previous paragraph, there exist* $\varepsilon > 0$ *and* $R > 0$ *for which* $\Omega = [t_0 - \varepsilon, t_0 + \varepsilon] \times (x_0 + R\overline{B})$ *is a subset of* U *on which one has the following. For every* $M > 0$ *and* $\rho \in (0, R)$, *there exists* $\delta > 0$ *so small that for every pair of endpoints* (a', x'_a) *and* (b', x'_b) *in* Ω *obeying*

$$0 < b' - a' < \delta, \quad |x'_b - x'_a| \le M(b' - a'),$$
$$\min\{|x'_b - x_0|,\ |x'_a - x_0|\} < \rho,$$

the following problem has a Lipschitz solution:

$$\min\left\{\int_{a'}^{b'} L\,(t, x(t), \dot{x}(t))\,dt : x(a') = x'_a,\ \ x(b') = x'_b,\right.$$
$$\left.|x(t) - x_0| < R\ \ \forall t \in [a', b']\right\}.$$

In fact, all solutions of this problem are Lipschitz.

PROOF. Without loss of generality, we take $x_0 = 0$. Then we recognize the problem described in the theorem's statement as (P'_R), and observe that the desired conclusions will follow immediately from Thm. 6.2 if we can verify that Ω can be chosen so that (H1), (H2), and (H5) hold. It actually suffices to confirm these conditions for an integrand of the form

$$\tilde{L}(t, x, v) = L(t, x, v) - \langle \varsigma, v\rangle - \mu,$$

since the solution sets corresponding to L and $\tilde{L}$ in (P'_R) are identical. To summarize, we need only to show that for some small $\varepsilon > 0$ and $R > 0$ and some $\varsigma \in \mathbb{R}^n$, $\mu \in \mathbb{R}$, the Lagrangian $\tilde{L}$ obeys (H1), (H2), (H5) on Ω.

Now (6.5) and (6.6) give (H1) and (H2), provided $\Omega \subseteq U$. Only (H5) presents a serious challenge. To overcome it, fix $\varsigma \in \partial_v L(t_0, 0, 0)$. The convexity of $L(t_0, 0, \cdot)$ implies that for any $M > 0$, the infimum defining

$$\alpha(M) = \inf\left\{\frac{L(t_0, 0, v) - L(t_0, 0, 0) - \langle \varsigma, v\rangle}{|v|} : |v| \ge M\right\}$$

is attained by some v with $|v| = M$. Moreover, the strict convexity at infinity of $L(t_0, 0, \cdot)$ allows us to fix M so large that $\alpha(M) > 0$. We then take $\overline{\alpha} = \alpha(M)/2$. Since L is continuous on compact subsets of $U \times \mathbb{R}^n$, there is a neighbourhood of $(t_0, 0)$ throughout which

$$\overline{\alpha} \leq \frac{L(t,x,v) - L(t,x,0) - \langle \varsigma, v \rangle}{|v|} \quad \text{for all } |v| \geq M.$$

For sufficiently small values of ε and R, Ω will be a subset of this neighbourhood.

By continuity and compactness, the following quantity is finite:

$$\omega = \inf\{L(t,x,v) - \langle \varsigma, v \rangle - \overline{\alpha}|v| : (t,x) \in \Omega, \ |v| \leq M\}.$$

Upon choosing $\mu = \min\{\omega, L(t,x,0) : (t,x) \in \Omega\}$, it is easy to check that $\tilde{L}(t,x,v) \geq \overline{\alpha}|v|$ for every $(t,x,v) \in \Omega \times \mathbb{R}^n$, as required. ∎

COROLLARY 6.4. *Assume* (6.5) *and* (6.6). *For every* $M > 0$, *there exist* $\eta > 0$ *and* $R > 0$ *such that for every pair of endpoints* (a', x'_a) *and* (b', x'_b) *in* $(t_0 - \eta, t_0 + \eta) \times (x_0 + \eta B)$ *obeying*

$$0 < b' - a', \quad |x'_b - x'_a| \leq M(b' - a'),$$

the problem in the statement of Theorem 6.3 *has a Lipschitz solution. In fact, all its solutions are Lipschitzian.*

PROOF. Let ε and R be given by Thm. 6.3. For any given $M > 0$, choose $\rho = \frac{R}{2}$ and let δ be given by Thm. 6.3. The choice $\eta = \min\left\{\frac{\delta}{2}, \rho, \varepsilon\right\}$ reveals the statement of the Corollary as an instance of the Theorem. ∎

Of all the results in this section, Corollary 6.4 is the most readily comparable to [9, Thm. 2.1]. Its hypotheses are weaker in that strict convexity at infinity is required instead of strict convexity, but its conclusions are weaker too: the solutions are only asserted to be Lipschitz instead of smooth. We have noted above, however, that in the presence of strict convexity every Lipschitz solution is automatically smooth. Thus Corollary 6.4 is a proper generalization of [9, Thm. 2.1].

7. - Existence and regularity in the large

When the Lagrangian L is defined and locally Lipschitz on $[a,b] \times \mathbb{R}^n \times \mathbb{R}^n$, we can consider the variant of problem (P_R) in which the constraint $\|x\| < R$ is removed. For a given choice of endpoints (a, x_a) and (b, x_b), this problem is labeled (P):

$$(P) \qquad \min\{\Lambda(x) : x \in AC[a,b], \ x(a) = x_a, \ x(b) = x_b\}.$$

This section is devoted entirely to (P). Accordingly, we posit throughout the basic hypotheses (H1) (H2) for $R = +\infty$.

7.1 *Existence for* (P)

Criteria for existence of (Lipschitz) solutions to (P) can be derived from the results of §3 in terms of the following quantities measuring the growth of L:

$$\alpha(R,V) := \inf\left\{\frac{L(t,x,v)}{|v|} : t \in [a,b],\ |x| \le R,\ |v| \ge V\right\}$$

$$\beta(R,V) := \inf\{0, L(t,x,v) : t \in [a,b],\ |x| \le R,\ |v| < V\}.$$

THEOREM 7.1. *Suppose that for some* $R > 0$, *for some Lipschitz arc* $\bar{x}$ *feasible for* (P) *with* $\|\bar{x}\| < R$, *the following two conditions hold*:

(H8) $b - a \le \Delta_R(r,s)$, *where* $r > \dfrac{R - \min(|x_a|, |x_b|)}{b-a}$ *and* $\rho(s) > r$,

(H9) *For every* $R' \ge R$ *there exists* $V(R') > 0$ *such that* $\alpha(R', V(R')) > 0$ *and*

$$\begin{aligned}\Lambda(\bar{x}) < \alpha(R', V(R'))\ &\{R' - \min(|x_a|, |x_b|) - (b-a)V(R')\}\\ &+ (b-a)\beta(R', V(R')).\end{aligned}$$

Then the solution set for problem (P) *is a nonempty subset of* $AC^\infty[a,b]$.

PROOF. In studying existence for (P), it clearly suffices to limit attention to arcs x satisfying $\Lambda(x) \le \Lambda(\bar{x})$. We shall now show that such arcs necessarily satisfy $\|x\| < R$. Set

$$R' = \|x\|,\ \ V' = V(R'),\ \ \alpha' = \alpha(R', V'),\ \ \beta' = \beta(R', V'),$$

$$C = \{t \in [a,b] : |\dot{x}(t)| \ge V'\}, \quad D = \{t \in [a,b] : |\dot{x}(t)| < V'\}.$$

Then the definitions of α and β yield:

$$\begin{aligned}\Lambda(\bar{x}) \ge \Lambda(x) &\ge \alpha' \int_C |\dot{x}(t)|\mathrm{d}t + \int_D \beta' \mathrm{d}t\\ &\ge \alpha' \left\{\int_a^b |\dot{x}(t)|\mathrm{d}t - \int_D |\dot{x}(t)|\mathrm{d}t\right\} + (b-a)\beta'\\ &\ge \alpha'\{R' - \min(|x_a|, |x_b|) - (b-a)V'\} + (b-a)\beta'.\end{aligned}$$

In view of (H9), this inequality can only hold if $R' < R$, which is the required conclusion.

Having established that adding to (P) the constraint $\|x\| < R$ makes no difference, it suffices to consider the resulting problem (P_R). We shall complete the proof by demonstrating that Theorem 3.2 applies to the Lagrangian

$$\tilde{L}(t,x,v) := L(t,x,v) + \alpha(R, V(R))\ V(R) - \beta(R, V(R)).$$

Since $\tilde{L}$ and L differ by a constant, the conclusion for the corresponding (P_R) yields the required result for (P).

We set $\overline{\alpha} = \alpha(R, V(R))$, upon which (H5) is an immediate consequence of the definitions of α and β. To see that (H7) holds for all α in $(0, \overline{\alpha})$ sufficiently near $\overline{\alpha}$, observe that

$$\tilde{\Lambda}(\overline{x}) = \Lambda(\overline{x}) + (b-a)[\overline{\alpha}V(R) - \beta(R, V(R))],$$

and that the inequality in (H9) can be rewritten as

$$\frac{R - \min(|x_a|, |x_b|)}{b-a} > \frac{\tilde{\Lambda}(\overline{x})}{\overline{\alpha}(b-a)}.$$

Thus (H8) implies (H7). Finally, this last inequality is identical to (H6) for $\alpha = \overline{\alpha}$, which shows that (H6) is satisfied for all α sufficiently near $\overline{\alpha}$. ∎

Slow-growth Lagrangians

Let us recall from §2 the following *extremal growth condition*:

(EGC) for all $R > 0$, for all $r > 0$, $\lim_{s\to\infty} \Delta_R(r,s) = \infty$.

Two examples of §3 were analyzed in part by verifying this condition. Later we shall examine other situations in which (EGC) can be shown to hold. For now we pause to note a corollary of the theorem. The proof consists merely of verifying that (H9) holds for any sufficiently large R, which we leave as an exercise. (Recall that (H1), (H2) are in force throughout).

COROLLARY 7.2. *Let L be a function bounded below which satisfies the extremal growth condition as well as*

$$\liminf_{\substack{|v|\to\infty \\ (t,x)\in[a,b]\times\mathbb{R}^n}} \frac{L(t,x,v)}{|v|} > 0. \tag{7.1}$$

Then the solution set of (P) is a nonempty subset of $AC^\infty[a,b]$.

The growth condition (7.1) is considerably milder ("slower") than the coercivity usually invoked in existence theory, which we shall examine later. As a first example of a class of slow-growth Lagrangians which automatically satisfy the extremal growth condition, we adduce the following, familiar from the theory of parametric problems.

PROPOSITION 7.3. *Let $L(t,x,v)$ have the form $\varphi(x)(1+|v|^2)^{1/2}$, where φ is a positive-valued locally Lipschitz function bounded away from 0 on R^n.*

Then L satisfies the extremal growth condition, as well as the other hypotheses of Corollary 7.2. *Thus the solution set for* (P) *is a nonempty subset of* $C^1[a,b]$.

PROOF. Let r, s and x be as in the definition of $\Delta_R(r,s)$ (Def. 2.9). The solution x of the subproblem in question satisfies the second Erdmann condition [1], which amounts in this case to the existence of a constant c such that

$$\varphi(x)^2 = c^2(1+\dot{x}^2), \text{ for every } t. \tag{7.2}$$

In view of (a) in the definition of $\Delta_R(r,s)$, we deduce

$$c^2 \geq \frac{\varepsilon_R^2}{1+r^2},$$

where $\varepsilon_R > 0$ is a lower bound on $\varphi(x)$ for $|x| \leq R$. But then, for any t in $[t_0, t_1]$ we have (in view of (7.2)):

$$|\dot{x}(t)|^2 \leq M_R^2 \, \frac{1+r^2}{\varepsilon_R^2} - 1$$

where M_R is an upper bound for $\varphi(x)$ for $|x| \leq R$. Thus if s is chosen so that $\rho(s)^2$ is greater than the right side of the last inequality, we have $\Delta_R(r,s) = +\infty$, which verifies (EGC). The other hypotheses of Corollary 7.2 follow readily. ■

To illustrate a situation rather similar to the above but in which the extremal growth condition fails, we consider Lagrangians $L(t,x,v)$ having the form $\varphi(t)(1+|v|^2)^{1/2}$, where φ is a continuous function such that $m := \min\{\varphi(t) : a \leq t \leq b\}$ is positive. It is known [1, 14.3.iv] that when L has this form, the problem (P) admits a solution iff

$$|x_b - x_a| \leq \int_a^b \frac{m}{\sqrt{\varphi(t)^2 - m^2}} \, dt.$$

When this condition fails (as it can of course), then the extremal growth condition must also fail, for (P) has no solution (Lipschitz or otherwise), yet all the other hypotheses of Corollary 7.2 are present.

7.3 *Coercive (fast-growth) Lagrangians*

The Lagrangian L is said to be *coercive* if there exists a convex function $\theta : [0,\infty) \to \mathbb{R}$ satisfying $\lim \frac{\theta(r)}{r} = +\infty$ such that

$$L(t,x,v) \geq \theta(|v|), \text{ for all } (t,x,v) \in [a,b] \times \mathbb{R}^n \times \mathbb{R}^n.$$

This has been a familiar growth assumption in existence theory since the work of Tonelli. In view of Definition 2.1 and the remarks immediately thereafter, it is clear that convexity and coercivity imply hypothesis (H2). The proof of the following consequence of Theorem 7.1 requires only verifying that (H9) holds for sufficiently large R, and is left as an exercise.

COROLLARY 7.4. *Let L be coercive and satisfy the extremal growth condition* (EGC). *Then the solution set of (P) is a nonempty subset of $AC^{\infty}[a, b]$.*

This result resembles closely Tonelli's original global existence theorem; it has an extra hypothesis (EGC) but asserts that the solutions are Lipschitz. This highlights (EGC) as a condition leading to *regularity* of the solutions to (P). (We remark that once a solution x is known to be Lipschitz, further regularity conclusions follow easily from supplementary hypotheses on L (see [8, §2])). We shall see that coercivity, besides providing Lagrangian growth and strict convexity at infinity, has a bearing on (EGC). The first such result which we now give leads to an alternate proof of the fact (due to Clarke and Vinter [8]) that under Tonelli's hypotheses the solutions to autonomous problems are Lipschitz. (The problem (P) is termed autonomous when the Lagrangian $L(t, x, v)$ has no explicit dependence on t).

PROPOSITION 7.5. *Let L be coercive and independent of t. Then* (EGC) *holds, so that the conclusions of Corollary* 7.4 *are valid.*

PROOF. Fix R and r; we shall show that $\Delta_R(r, s) = +\infty$ for s sufficiently large. We begin with a technical result.

LEMMA 7.6. *For every $m > 0$ there exists $M > 0$ such that if $|x| \leq R$, $|p| \leq m$, then $\partial_p H(x, p) \subseteq M\overline{B}$.*

PROOF. Begin by choosing M so large that $s > M$ implies $\dfrac{\theta(s) - L(x, 0)}{s} > m$ whenever $|x| \leq R$.

Then, if $\partial_p H(x, p)$ contains a point v with $|v| > M$, the subgradient inequality corresponding to $p \in \partial_v L(x, v)$ implies

$$|p|\ |v| \geq \langle p, v\rangle \geq L(x, v) - L(x, 0) \geq \theta(|v|) - L(x, 0).$$

Upon dividing across by $|v|$ we arrive at $|p| > m$, a contradiction which completes the proof of the lemma.

We now define

$$\begin{aligned} m_1 := \max\{L(x, v) : |x| \leq R,\ |v| \leq 1\} + \max\{L(x, v) : |x| \leq R,\ |v| \leq r\} \\ + r\ \max\{|\varsigma| : \varsigma \in \partial_v L(x, v),\ |x| \leq R,\ |v| \leq r\}, \end{aligned}$$

and we let M_1 be the number corresponding to m_1 as in Lemma 7.6. We will complete the proof of Prop. 7.5 by showing that if

$$\rho(s) > \max[r, M_1] \text{ then } \Delta_R(r, s) = +\infty. \qquad (7.3)$$

We place ourselves in the notational context of the definition of $\Delta_R(r,s)$. From condition 2.9 (d) and [4, Theorem 5.2.1] we deduce that corresponding to x are an arc $p \in AC[t_0,t_1]$ and a constant $c \in \mathbb{R}$ for which

$$p(t) \in \partial_v L(x(t),\dot{x}(t)) \qquad \text{a.e. } [t_0,t_1] \tag{7.4}$$

$$\langle p(t),\dot{x}(t)\rangle - L(x(t),\dot{x}(t)) = c \qquad \text{a.e. } [t_0,t_1]. \tag{7.5}$$

Now condition (a) states that Ess $\dot{x}(\tau) \cap r\overline{B} \neq \emptyset$; it follows from (7.4) that for this τ, $|p(\tau)| \leq \sigma_1$, where

$$\sigma_1 := \max\{|\varsigma| : \varsigma \in \partial_v L(x,v),\ |x| \leq R,\ |v| \leq r\}.$$

If we now define

$$\sigma_2 := r\sigma_1 + \max\{L(x,v) : |x| \leq R,\ |v| \leq r\},$$

then (7.5) implies $|c| \leq \sigma_2$. From the subgradient inequality corresponding to (7.4), together with (7.5), we deduce that

$$\langle p(t),v\rangle - L(x(t),v) \leq c,\ \forall v \in \mathbb{R}^n,\ \forall t \in [t_0,t_1].$$

Therefore

$$|p(t)| \leq |c| + \max\{L(x(t),v) : |v| \leq 1\} \leq m_1,\ \forall t \in [t_0,t_1].$$

now (7.4) is equivalent to

$$\dot{x}(t) \in \partial_p H(x(t),p(t)), \qquad \text{a.e. } [t_0,t_1],$$

so the definition of M_1 together with the preceding bound on $|p(t)|$ yields

$$\text{Ess } \dot{x}(t) \subseteq M_1\overline{B},\ \text{ for all } t \in [t_0,t_1].$$

If $\rho(s) > M_1$, this contradicts condition (b) in the definition of $\Delta_R(r,s)$, which confirms that $\Delta_R(r,s)$ is $+\infty$ for such s. This completes the proof of (7.3) and establishes Proposition 7.5. ■

Morrey-type conditions

We turn now to another criterion allowing t-dependence which, when combined with coercivity, gives the extremal growth condition. We make the following assumption: there exist a positive constant c and an integrable function $\gamma \in L^1[a,b]$ such that $|\varsigma_1| \leq c|\varsigma_2| + \gamma(t)$ for all $(\varsigma_1,\varsigma_2) \in \partial L(t,x,v)$, for all $(t,x,v) \in [a,b] \times \mathbb{R}^n \times \mathbb{R}^n$, where ∂L denotes the generalized gradient in (x,v) variables. (Actually it would suffice to require this for $|x| \leq R$, where the latter is an a priori bound on arcs x solving (P) imposed by the coercivity hypothesis). We refer to this as the *Morrey growth condition* (see [8]).

PROPOSITION 7.7. *If L is coercive and satisfies the Morrey growth condition, then* (EGC) *holds, so that the conclusions of Corollary* 7.4 *are valid.*

PROOF. We let R be a natural a priori bound on $\|x\|_\infty$ for any solution x to (P), one generated by the boundary conditions, the coercivity and the inequality $\Lambda(x) \leq \Lambda(\overline{x})$. With this noted, we proceed to verify that Lemma 7.6 still holds in the nonautonomous case. The rest of the proof also parallels that of Proposition 7.5: given $r > 0$, we shall produce a number M_1 having the property that $\Delta_R(r,s) = +\infty$ whenever $\rho(s) > \max[r, M_1]$ (cf. (7.3)). Placing ourselves once again in the context of the definition of $\Delta_R(r,s)$, we recall that corresponding to the solution x is an arc p satisfying

$$(\dot{p}(t), p(t)) \in \partial L\left(t, x(t), \dot{x}(t)\right), \text{ a.e. } t \in [t_0, t_1].$$

In view of condition (a), one has $|p(\tau)| < \sigma_1$, where

$$\sigma_1 := \max\{|\varsigma_2| : (\varsigma_1, \varsigma_2) \in \partial L(t, x, v),\ t \in [a,b],\ |x| \leq R,\ |v| \leq r\}.$$

The Morrey growth condition implies

$$|\dot{p}(t)| \leq c|p(t)| + \gamma(t), \text{ a.e. } t \in [t_0, t_1],$$

which via Gronwall's inequality yields

$$|p(t)| \leq (\sigma_1 + \|\gamma\|_1)(1 + e^{c(b-a)}), \text{ for all } t \in [t_0, t_1].$$

As in the proof of Proposition 7.5, this uniform x-independent bound on $|p(t)|$, together with Lemma 7.6, produces a similar corresponding bound on $|\dot{x}(t)|$ and implies that $\Delta_R(r,s) = +\infty$ for s sufficiently large. ∎

8. - Periodic Hamiltonian trajectories

The subject of periodic trajectories of Hamiltonian systems has seen considerable activity in recent years. We shall illustrate in this section the use of intermediate existence in connection with the *dual action principle* [3] to deduce the existence of such trajectories in a nonautonomous context somewhat different from those considered previously. The results of this section are best compared to those of [4] and [5].

Given a function $H(t,z) : \mathbb{R} \times \mathbb{R}^{2n} \to \mathbb{R}$ (the Hamiltonian), we study the existence of absolutely continuous functions $z(t)$ satisfying the Hamiltonian boundary-value problem

(8.1) $$J\dot{z}(t) \in \partial H(t, z(t)),\ 0 \leq t \leq T,\ z(0) = z(T),$$

where J is the $2n \times 2n$ matrix

$$\begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$$

($I_n = n \times n$ identity matrix) and where ∂H signifies the generalized gradient of H with respect to the z variable. It is now well-known [4], [5] that when H is convex in z, certain such trajectories z result from "critical points" of the dual action $D(z)$ defined by

$$D(z) := \int_0^T \left\{ \frac{1}{2} \langle Jz(t), \dot{z}(t) \rangle + G(t, J\dot{z}(t)) \right\} \mathrm{d}t,$$

where $G(t, \cdot)$ is the conjugate of $H(t, \cdot)$ in the sense of convex analysis:

$$G(t, \varsigma) := \sup\{\langle \varsigma, z \rangle - H(t, z) : z \in \mathbb{R}^{2n}\}.$$

In some circumstances the required critical point is obtained by (globally) minimizing D; in others, critical point theory serves. We shall now describe circumstances in which intermediate existence applies to D in the definite absence of a global minimum.

The hypotheses on the Hamiltonian are the following:

(a) $H(t, z)$ is locally Lipschitz, convex in z, and satisfies

$$H(t, z) \geq \theta(|z|)$$

where θ is a convex function whose conjugate θ^* is everywhere finite (equivalently, $\frac{\theta(r)}{r} \to \infty$ as $r \to \infty$).

(b) For some constant c, at any point (t, z) such that the partial derivative $H_t(t, z)$ exists, we have

$$|H_t(t, z)| \leq c[1 + |H(t, z)|].$$

(c) For some $\gamma > 0$, for all z with $|z| \leq \gamma$,

$$H(t, z) \leq \frac{\gamma^2}{2T} - \theta^*(0), \;\; 0 \leq t \leq T.$$

THEOREM 8.1. *Under hypotheses* (a)-(c) *above, the Hamiltonian boundary-value problem* (8.1) *admits a solution.*

PROOF. We begin with some observations about G.

LEMMA 8.2. *$G(t, \varsigma)$ is bounded for (t, ς) bounded, as is $\partial G(t, \varsigma)$. G is locally Lipschitz.*

PROOF. We observe first $G(t,\varsigma) \geq -H(t,0)$ (from the definition), and also (by (a))

$$G(t,\varsigma) \leq \theta^*(|\varsigma|).$$

It follows that G is locally bounded. The fact that $\partial G(t,\varsigma)$ (the subdifferential in ς) is locally bounded then follows from convex analysis. The supremum defining G is attained at those z satisfying $\varsigma \in \partial H(t,z)$, or equivalently $z \in \partial G(t,\varsigma)$. The set of such z is therefore bounded when (t,ς) is bounded, so that the supremum can just as well be taken over a compact set. This readily yields that G inherits the locally Lipschitz property from H. ∎

LEMMA 8.3. *For every $M > 0$, there exist $N_1, N_2 > 0$ such that for all t in $[0,T]$*

$$\text{if } \partial H(t,z) \cap MB \neq \emptyset \text{ then } |z| \leq N_1$$
$$\text{if } H(t,z) \leq M \quad \text{then } |\varsigma| \leq N_2, \text{ for all } \varsigma \in \partial H(t,z).$$

PROOF. If the first implication were false for all N_1, there would be sequences t_i, z_i, ς_i such that $|z_i| \to \infty$, $\varsigma_i \in \partial H(t_i, z_i)$, and $|\varsigma_i| \leq M$. Thus $z_i \in \partial G(t_i, \varsigma_i)$, where (t_i, ς_i) is bounded. This contradicts the previous lemma and proves the existence of N_1. The existence of N_2 follows from the fact that (in view of (a)) $H(t,z) \leq M$ implies a bound on $|z|$. ∎

LEMMA 8.4. *The Lagrangian L figuring in the dual action functional D satisfies the extremal growth condition* (EGC).

PROOF. To see this, we place ourselves in the notational context of the definition of $\Delta_R(r,s)$. The solution x to the subproblem described in condition 2.9 (e) satisfies the Euler equation for L, which is known to imply [4, p. 282] that for some constant k one has

$$J\dot{x}(t) \in \partial H(t, x(t) + k), \text{ a.e. } t_0 \leq t \leq t_1.$$

It follows from (a) of the definition and from Lemma 8.3 that $|x(\tau) + k|$ is bounded by some number N_1 depending only on r. The necessary conditions for the subproblem go beyond just the Euler equation, since we have Lipschitz dependence in t. In fact, from [4, §3.6] we deduce that the (Lipschitz) function

$$h(t) = H(t, x(t) + k)$$

satisfies

$$(\dot{h}(t), J\dot{x}(t)) \in \partial_{t,z} H(t, x(t) + k), \text{ a.e. } t_0 \leq t \leq t_1.$$

Now we invoke hypothesis (b) to deduce

$$|\dot{h}(t)| \leq c[1 + |h(t)|], \text{ a.e. } t_0 \leq t \leq t_1,$$

which combined with the known bound on $|x(\tau)+k|$ (and hence on $h(\tau)$) and Gronwall's lemma leads to a bound for $h(t)$ on $[t_0, t_1]$ depending only on r. We now invoke the second half of Lemma 8.3 to derive a bound N_2 on $|x(t)+k|$, depending as always only on r. This gives a bound for $|\dot{x}(t)|$ on $[t_0, t_1]$ in view of (8.2), and shows that if s were chosen sufficiently large, condition (b) of the definition of $\Delta_R(r,s)$ could not hold. Thus for such s one has $\Delta_R(r,s) = +\infty$, establishing the extremal growth condition. ∎

We now set the stage for applying Theorem 3.1. The Lagrangian L is that of the dual action, and the endpoint data is $a = 0$, $b = T$, $x_a = x_b = 0$. We set $R = \gamma$, $\overline{\alpha} = \frac{\gamma}{2}$ and $m = \frac{R}{2}$, where γ is given in hypothesis (c). For $\overline{x}$ we choose the admissible arc which is identically zero. Observe

$$\Lambda(\overline{x}) = \int_0^T G(t,0)dt \leq T\theta^*(0).$$

The Hamiltonian corresponding to L is easily calculated to be $H\left(t, \frac{z}{2} + Jp\right)$, which implies that L satisfies the hypothesis (H2) of strict convexity at infinity. (H1) is a consequence of Lemma 8.2, and (EGC) provides (H4). The final hypothesis of Theorem 3.1 to verify is (H3), which is an easy consequence of (c) for the data as defined. Consequently the problem (P_R) at hand admits a Lipschitz solution x satisfying the Euler equation for the dual action. But then for some constant k, $x(\cdot)+k$ is a solution of the Hamiltonian boundary-value problem being studied. This completes the proof of Theorem 8.1. ∎

REMARKS.

1) The hypotheses of Theorem 8.1 do not imply that the dual action D admits a global minimum (subject to the given boundary conditions), as may easily be verified directly in the case $n = 1$ with

$$H(t,z) = \frac{|z|^3}{3}, \quad G(t,\varsigma) = \frac{2}{3}\,|\varsigma|^{3/2}.$$

The arcs $z(t)$ of the form $\left[r\,\sin\left(\frac{2\pi t}{T}\right),\ r\left(\cos\left(\frac{2\pi t}{T}\right) - 1\right)\right]$ are such that $D(z)$ goes to $-\infty$ as r goes to ∞. We note also that the results of §7 fail to apply here since L is neither bounded below nor coercive.

2) The hypothesis (c) is the only one directly constraining the period T. It is automatically satisfied for example when H is globally subquadratic in z (see [5]), or when $\theta^*(0) = 0$ and H is superquadratic in z near the origin.

3) L is not necessarily strictly convex in v under the given hypotheses although, as we have shown, it is strictly convex at infinity. Nor is it differentiable, even if one adds the hypothesis that the given Hamiltonian H is continuously differentiable.

REFERENCES

[1] L. CESARI, *Optimization—Theory and Applications*, Springer-Verlag, New York, 1983.

[2] F.H. CLARKE, *The Euler-Lagrange differential inclusion*, J. Differential Equations **19** (1975), 80-90.

[3] F.H. CLARKE, *Solutions périodiques des équations hamiltoniennes*, Comptes Rendus Acad. Sci. Paris **287** (1978), 951-952.

[4] F.H. CLARKE, *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York, 1983.

[5] F.H. CLARKE, *Hamiltonian trajectories and local minima of the dual action*, Trans. Amer. Math. Soc. **287** (1985), 239-251.

[6] F.H. CLARKE - I. EKELAND, *Nonlinear oscillations and boundary value problems for Hamiltonian systems*, Arch. Rational Mech. Anal. **78** (1982), 315-333.

[7] F.H. CLARKE - R.B. VINTER, *On the conditions under which the Euler equation or the maximum principle hold*, Appl. Math. Optim. **12** (1984), 73-79.

[8] F.H. CLARKE - R.B. VINTER, *Regularity properties of solutions to the basic problem in the calculus of variations*, Trans. Amer. Math. Soc. **289** (1985), 73-98.

[9] F.H. CLARKE - R.B. VINTER, *Existence and regularity in the small in the calculus of variations*, J. Differential Equations **59** (1985), 336-354.

[10] P.D. LOEWEN, *On the Lavrentiev phenomenon*, Canad. Math. Bull. **30** (1987), 102-108.

[11] R.T. ROCKAFELLAR, *Convex Analysis*, Princeton Univ. Press, Princeton, N.J., 1970.

[12] R.T. ROCKAFELLAR, *Existence theorems for general control problems of Bolza and Lagrange*, Adv. in Math. **15** (1975), 312-333.

[13] L. TONELLI, *Sur une méthode directe du calcul des variations*, Rend. Circ. Mat. Palermo **39** (1915), 233-264; also in *Opere Scelte* (Vol. 2), 289-333, Cremonese, Rome, 1961.

[14] L. TONELLI, *Fondamenti di Calcolo delle Variazioni* (2 Vols.), Zanichelli, Bologna, 1921, 1923.

Centre de recherches mathématiques
Université de Montréal
C.P. 6128-A
Montréal, QC
Canada H3C 3J7

Department of Mathematics
University of British Columbia
Vancouver, B.C.
Canada V6T 1Y4