

THE SEMIGROUP PROPERTY OF VALUE FUNCTIONS IN LAGRANGE PROBLEMS

PETER R. WOLENSKI

ABSTRACT. The Lagrange problem in the calculus of variations exhibits the principle of optimality in a particularly simple form. The binary operation of inf-composition applied to the value functions of a Lagrange problem equates the principle of optimality with a semigroup property. This paper finds the infinitesimal generator of the semigroup by differentiating at $t = 0$. The type of limit is epigraphical convergence in a uniform sense. Moreover, the extent to which a semigroup is uniquely determined by its infinitesimal generator is addressed. The main results provide a new approach to existence and uniqueness questions in Hamilton-Jacobi theory. When L is in addition finite-valued, the results are given in terms of pointwise convergence.

1. INTRODUCTION

In this paper we are concerned with properties of the so-called value function $V^{(T)}$ associated with the following Lagrange problem:

$$(1.1) \quad \inf \int_0^T L(x(t), \dot{x}(t)) dt \quad \text{over } x(\cdot) \in AC[0, T]$$

satisfying $x(0) = \xi$ and $x(T) = \eta$. Here $T > 0$, $AC[0, T]$ is the set of absolutely continuous functions from $[0, T]$ into \mathbb{R}^n , $\dot{x}(\cdot)$ signifies the derivative of $x(\cdot)$, and ξ and η are elements of \mathbb{R}^n . The precise assumptions on the extended real-valued integrand L (which is referred to as the Lagrangian) will be given below, but at this point, we merely emphasize that the admittance of infinite values to L allows for generality in the problem formulation (1.1) than is perhaps readily apparent. We refer to Rockafellar [22, §4, 24] or Clarke [5, §1.3] for a detailed discussion in the techniques for which standard control problems can be reformulated into the form (1.1). The optimal value in (1.1) is denoted by $V^{(T)}(\xi, \eta)$. This notation suggests that for fixed $T > 0$, we are viewing $V^{(T)}(\cdot, \cdot)$ as an (extended) real-valued function of the endpoints ξ and η . The goal of this paper is to characterize the collection $\{V^{(T)}\}_{T \geq 0}$ of value functions.

The approach to our characterization is based upon a (one-parameter) semigroup property which is satisfied by $\{V^{(T)}\}_{T \geq 0}$. The binary operation here is

Received by the editors April 27, 1990.

1991 *Mathematics Subject Classification*. Primary 49C20, 49C99.

Key words and phrases. Lagrange problems, principle of optimality, epigraphical convergence, Hamilton-Jacobi theory.

©1993 American Mathematical Society
0002-9947/93 \$1.00 + \$.25 per page

that of inf-composition, which is defined as follows: If U_1 and U_2 are extended real-valued functions defined on $\mathbb{R}^n \times \mathbb{R}^n$, the inf-composition $U_1 \circ U_2$ is defined on $\mathbb{R}^n \times \mathbb{R}^n$ into $\mathbb{R}^1 \cup \{\pm\infty\}$ by

$$(1.2) \quad U_1 \circ U_2(\xi, \eta) = \inf_{\gamma \in \mathbb{R}^n} \{U_1(\xi, \gamma) + U_2(\gamma, \eta)\}.$$

Hence we say that $\{U^{(T)}\}_{T \geq 0}$ is a semigroup if

$$(1.3) \quad U^{(s+t)} = U^{(s)} \circ U^{(t)} \quad \text{for all } s, t \geq 0.$$

The “principle of optimality” (as is relevant to (1.1)) is thus equivalent to the property that $\{V^{(T)}\}_{T \geq 0}$ is a semigroup.

The standard approach to the study of problems of form (1.1) often centers around the Hamilton-Jacobi equation (HJ)

$$(1.4) \quad \frac{\partial U^{(t)}(\xi, \eta)}{\partial t} + H(\eta, \nabla_\eta U^{(t)}(\xi, \eta)) = 0,$$

where the Hamiltonian $H(\cdot, \cdot)$ is given by

$$H(\xi, p) = \sup_{v \in \mathbb{R}^n} \{\langle v, p \rangle - L(\xi, v)\}.$$

The partial differential equation (1.4) can be derived (at least formally) from property (1.3); see Bellman [1] or Pontryagin et al. [20]. However to guarantee the existence of a solution to (HJ), restrictive smoothness assumptions must be imposed on the value function. This is the approach followed in [20] and other subsequent early work. Also, see Cesari [3, p. 203]. But because these assumptions fail to hold in many of the simplest problems, nonsmooth analysis has been developed to play an essential role in extending the theory. Loosely speaking, this takes the form of replacing the classical derivatives in (1.4) by some notion of generalized gradient. This approach is taken in work by Clarke and Vinter [6], Clarke [4], and Zeidan [33, 34] which employ Clarke generalized gradients in the derivation of optimality conditions. An alternative notion, that of viscosity solution, was introduced by Crandall and Lions [11] which in effect uses Dini subgradients coupled with certain inequalities. The crowning achievement of the viscosity approach is that under quite general assumptions, viscosity solutions exist and are unique. See Crandall, Evans and Lions [9], Lions [17], Crandall, Ishii, and Lions [10], and the papers [12, 16, 18] for further references. One further concept has recently been introduced by Frankowska [14], Berkovitz [2], and Vinter and Wolenski [28] in which Dini directional derivatives are featured rather than subgradients.

The main results of this paper are perhaps more primitive than those just mentioned. We are merely interested to what extent the semigroup property (1.3) determines the value functions. Indeed, the following two natural questions arise in connection with any (one-parameter) semigroup.

- (Q1) What is the infinitesimal generator of the semigroup, and in what sense is it the “derivative” at $t = 0$?
- (Q2) To what extent does the infinitesimal generator characterize the semigroup?

In the context of problem (1.1), these questions were posed but left open in Rockafellar [23]. We will provide answers to (Q1) and (Q2) assuming L satisfies convexity, coercivity, and Lipschitz-type assumptions. Before stating these

assumptions precisely, we will next review the derivation of (1.4) from (1.3). This will give the pedagogical connection of (Q1) and (Q2) with Hamilton-Jacobi theory.

First note that if $\{U^{(t)}\}$ satisfies (1.3), then for each $h > 0$, we have

$$U^{(t+h)}(\xi, \eta) - U^{(t)}(\xi, \eta) = \inf_{\gamma \in \mathbb{R}^n} \{U^{(t)}(\xi, \gamma) - U^{(t)}(\xi, \eta) + U^{(h)}(\gamma, \eta)\}.$$

Dividing by h and making the change of variables $\gamma = \eta - hv$, this becomes

$$(1.5) \quad \begin{aligned} & \frac{1}{h} \{U^{(t+h)}(\xi, \eta) - U^{(t)}(\xi, \eta)\} \\ &= \inf_{v \in \mathbb{R}^n} \left\{ \frac{U^{(t)}(\xi, \eta - hv) - U^{(t)}(\xi, \eta)}{h} + \frac{1}{h} U^{(h)}(\eta - hv, \eta) \right\}. \end{aligned}$$

By letting $h \searrow 0$ one sees that (1.4) arises from (1.5), provided the derivatives exist, the same $v \in \mathbb{R}^n$ achieves the inf in (1.5) for all small h , and the convergence

$$(1.6) \quad \frac{1}{h} U^{(h)}(\gamma, \gamma + hv) \rightarrow L(\eta, v)$$

holds as $h \searrow 0$. (Recall that $\gamma = \eta - hv$ and hence $\gamma \rightarrow \eta$ as $h \searrow 0$.) Of course these are severe provisions.

Note that the function

$$\Psi^{(h)}(\gamma, v) := \frac{1}{h} U^{(h)}(\gamma, \gamma + hv) \quad \left(= \frac{1}{h} \{U^{(h)}(\gamma, \gamma + hv) - U^{(0)}(\gamma, \gamma)\} \right)$$

which appears in (1.6) is the difference quotient used in finding the directional derivative of $(h, \eta) \mapsto U^{(h)}(\gamma, \eta)$ at $(0, \gamma) \in \mathbb{R}^1 \times \mathbb{R}^n$ in the direction $(1, v) \in \mathbb{R}^1 \times \mathbb{R}^n$. Thus one can interpret the limiting behavior of the functions $\Psi^{(h)}(\gamma, \cdot)$ as $h \searrow 0$ as the “derivative” of the semigroup $\{U^{(t)}\}_{t \geq 0}$ at $t = 0$. In other words, the limit of $\Psi^{(h)}(\gamma, \cdot)$ as $h \searrow 0$ is the natural candidate for the infinitesimal generator; the precise manner in which the limit should be taken will be given in §3—a uniform epigraphical limit. As the previous paragraph suggests, showing L is the infinitesimal generator of the value function semigroup (see Theorem 3.1 below) is the semigroup counterpart to showing the value function satisfies (HJ). The resolution of the question if this “derivative” at $t = 0$ characterizes the value functions (see Theorem 3.2 and Theorem 6.2) is analogous to the determination if (HJ) has a unique solution. Thus, at least pedagogically, (Q1) and (Q2) are the respective semigroup formulations of the existence and uniqueness problems in Hamilton-Jacobi theory.

We now state our basic assumptions on L . Throughout the rest of the paper, L will be assumed to satisfy:

- (H1) for each $\xi \in \mathbb{R}^n$, $L(\xi, \cdot)$ is a proper, closed, convex function,
- (H2) for each $\xi \in \mathbb{R}^n$, $L(\xi, \cdot)$ is coercive,
- (H3) There exists $\lambda > 0$ so that for each ξ, ξ' , and $p \in \mathbb{R}^n$, we have

$$|H(\xi, p) - H(\xi', p)| \leq \lambda(1 + |p|)|\xi - \xi'|.$$

Recall (cf. [21]) that $L(\xi, \cdot)$ “proper” means $L(\xi, \cdot)$ never takes the value $-\infty$ and is not identically $+\infty$; “closed” is equivalent to lower semicontinuity; and “coercive” is the property that $\frac{1}{|v|}L(\xi, v) \rightarrow +\infty$. Assumption (H3)

(known as the strong Lipschitz condition [4]) is equivalent to: there exists $\lambda > 0$ so that for all $\xi, \xi' \in \mathbb{R}^n$, we have

$$(1.7) \quad \text{epi } L(\xi, \cdot) \subseteq \text{epi } L(\xi', \cdot) + \lambda|\xi - \xi'|B,$$

where $\text{epi } L(\xi, \cdot) = \{(v, \alpha) \in \mathbb{R}^n \times \mathbb{R}^1 : \alpha \geq L(\xi, v)\}$ is the epigraph of $L(\xi, \cdot)$. (Throughout, the letter B denotes the *closed* unit ball in the appropriate dimension. For example, in (1.7) B has dimension $n+1$). The proof of the equivalence of (H3) with (1.7) consists of taking Legendre-Fenchel conjugates. These details are left to the reader.

The assumptions (H1)–(H3) are the global autonomous versions of those employed in the necessary conditions of Clarke [4, or 5, §4.2]. Early viscosity results [9, 16] also use similar assumptions, however more recent work [12, 10] allow for somewhat weaker conditions.

We give here further justification for our assumptions by considering a special case of (1.1). Suppose $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a multifunction (i.e. a set-valued map), and consider the differential inclusion

$$(1.8) \quad \begin{aligned} x(\cdot) &\in AC[0, T] \\ \dot{x}(t) &\in F(x(t)) \quad \text{a.e. } t \in [0, T], \\ x(0) &= \xi. \end{aligned}$$

Many essential features of (1.8) are covered within the formulation (1.1) by letting L be defined as the indicator of F :

$$(1.9) \quad L(\xi, v) = \begin{cases} 0 & \text{if } v \in F(\xi), \\ +\infty & \text{if } v \notin F(\xi). \end{cases}$$

The reachable set multifunction is defined by $R^{(T)}(\xi) := \{x(T) : x(\cdot) \text{ satisfies (1.8)}\}$. Note that with L as in (1.9), $V^{(T)}$ is then the indicator of $R^{(T)}$:

$$V^{(T)}(\xi, \eta) = \begin{cases} 0 & \text{if } \eta \in R^{(T)}(\xi), \\ +\infty & \text{if } \eta \notin R^{(T)}(\xi). \end{cases}$$

Hence a semigroup characterization of $\{V^{(T)}\}$ subsumes a characterization of $\{R^{(T)}\}$ in terms of the semigroup operation of multifunction composition. For instance, if the values of F are sets consisting of a single element, such a characterization becomes a uniqueness theorem in ordinary differential equations. For differential inclusions, there are also uniqueness theorems which characterize reachable set multifunctions. The first such we are aware of is Roxin [26, Theorem 7.1]. A version involving “the funnel equation” is given by Panasiuk and Panasiuk [19]. Still another version is given by this author [32] which strengthens the notion of limit in finding the infinitesimal generator, but enlarges the class over which the reachable set multifunction is unique by allowing noncontinuous semigroups. All of these versions (essentially) employ the same assumptions on F . In autonomous form, these are

(H1') for each $\xi \in \mathbb{R}^n$, $F(\xi)$ is nonempty, closed, and convex

(H2') for each $\xi \in \mathbb{R}^n$, $F(\xi)$ is bounded,

(H3') F is Lipschitz in \mathbb{R}^n with respect to the Hausdorff metric.

If L is given by (1.9), then the set of assumptions (H1)–(H3) are the natural extensions of (H1')–(H3') to the more general situation.

The plan of the rest of the paper is as follows: Section 2 has preliminaries; the main results are stated in §3; §4 develops properties of Yosida approximations, which will be used in a proof in §5; §5 has proofs; §6 considers finite-valued Lagrangians; and §7 contains pointwise limit theorems.

2. PRELIMINARIES

The word “function”, unless explicitly stated otherwise, will always refer to an extended real-valued function. If $f : \mathbb{R}^m \rightarrow \mathbb{R}^1 \cup \{\pm\infty\}$ is a function, then the (effective) domain of f is the set $\text{dom } f := \{\xi : f(\xi) \neq +\infty\}$. The epigraph of f , denoted by $\text{epi } f$, is given by $\text{epi } f := \{(\xi, \alpha) \in \mathbb{R}^m \times \mathbb{R}^1 : \alpha \geq f(\xi)\}$. Recall that f is lower semicontinuous (l.s.c.) if and only if $\text{epi } f$ is closed. A collection $\{U^{(t)}\}_{t \geq 0}$ is said to be a l.s.c. semigroup on $\mathbb{R}^n \times \mathbb{R}^n$ if (1.3) holds and for each $(t, \xi) \in [0, \infty) \times \mathbb{R}^n$, the function $U^{(t)}(\xi, \cdot)$ is l.s.c.

A metric d is defined on the subsets of \mathbb{R}^n as follows (the so-called Hausdorff ρ -metric; see Wets [29]): let C and D be subsets of \mathbb{R}^n . For each $m = 1, 2, \dots$, define

$$\begin{aligned} d_m^+(C, D) &:= \inf\{\delta \geq 0 : C \cap mB \subseteq D + \delta B\}, \\ d^+(C, D) &:= \sum_{m=1}^{\infty} 2^{-m} \frac{d_m^+(C, D)}{1 + d_m^+(C, D)}, \end{aligned}$$

and

$$d(C, D) := \max\{d^+(C, D), d^+(D, C)\}.$$

It can be easily shown that a net $\{C_h\}_{h>0}$ of subsets of \mathbb{R}^n satisfies $d^+(C_h, D) \rightarrow 0$ as $h \searrow 0$ (resp. $d(C_h, D) \rightarrow 0$) if and only if for each $m = 1, 2, \dots$ and $\varepsilon > 0$, there exists h_0 so that for $0 < h \leq h_0$, we have $d_m^+(C_h, D) < \varepsilon$ (resp. $d_m^+(C_h, D) < \varepsilon$ and $d_m^+(D, C_h) < \varepsilon$).

3. THE MAIN RESULTS

Again consider the problem (1.1). For $h > 0$ and $\xi, v \in \mathbb{R}^n$, define $\Phi^{(h)}(\xi, v) := \frac{1}{h}V^{(h)}(\xi, \xi + hv)$. The following theorem responds to (Q1) as posed in the introduction. Recall that L is assumed to satisfy (H1)–(H3).

Theorem 3.1. *The collection $\{V^{(t)}\}_{t \geq 0}$ is a l.s.c. semigroup on $\mathbb{R}^n \times \mathbb{R}^n$ such that*

$$(3.1) \quad d(\text{epi } \Phi^{(h)}(\xi, \cdot), \text{epi } L(\xi, \cdot)) \rightarrow 0 \quad \text{as } h \searrow 0$$

uniformly over ξ in a compact subset of \mathbb{R}^n .

Our second main result is

Theorem 3.2. *Suppose $\{U^{(t)}\}_{t \geq 0}$ is a l.s.c. semigroup for which*

$$(3.2) \quad d^+(\text{epi } L(\xi, \cdot), \text{epi } \Psi^{(h)}(\xi, \cdot)) \rightarrow 0 \quad \text{as } h \searrow 0$$

uniformly for ξ in a compact set, where $\Psi^{(h)}(\xi, v) := \frac{1}{h}U^{(h)}(\xi, \xi + hv)$. Then for all $t > 0$ and $\xi, \eta \in \mathbb{R}^n$, we have

$$(3.3) \quad U^{(t)}(\xi, \eta) \leq V^{(t)}(\xi, \eta).$$

Theorem 3.2. responds to (Q2). It is in the spirit of similar “upper envelope” results which appear in Gonzales [15], Vinter and Lewis [27], and Vinter and Wolenski [28].

Of course it would be desirable if the inequality (3.3) was an equality (under the additional assumption that d^+ in (3.2) is replaced by d). If L is given as in (1.9), and the $U^{(t)}$'s are restricted to be indicators only, then this is indeed the case. Theorem 3.2 then reduces to the uniqueness theorem in [32]. However in the more general case (1.1), we can only prove equality in (3.3) with finite Lagrangians and with further restrictions on the semigroups. This result is given in §6, where we also discuss the additional assumption. Of course L finite everywhere precludes the situation (1.9). Hence only for Lagrangians L for which $\text{dom } L(\xi, \cdot)$ is bounded or $\text{dom } L(\xi, \cdot)$ equals all of \mathbb{R}^n (in addition to (H1)–(H3)) that it is known the inequality in (3.3) can be strengthened to an equality. Our proofs entail time discretization, and it is pertinent that in these two situations, it is known that minimizers are necessarily Lipschitz (when $\text{dom } L(\xi, \cdot)$ is bounded, this is trivial; the case $\text{dom } L(\xi, \cdot) = \mathbb{R}^n$ is handled in Clarke and Vinter [7]; see also [8]). In fact, in the intermediate situation, it may happen that minimizers are *AC* but not Lipschitz, and our proofs will fail.

Pointwise limits cannot replace epigraph limits in (3.1), as the following simple example illustrates. Let $n = 1$, $F(\xi) = \{\xi\}$, and L be as in (1.9). Then $R^{(t)}(\xi) = \{e^T \xi\}$, and so for each ξ and $h > 0$, one has $V^{(h)}(\xi, \xi + h\xi) = +\infty$ while $L(\xi, \xi) = 0$. However there are situations where pointwise limits exist, the most notable being when $\text{dom } L(\xi, \cdot) = \mathbb{R}^n$. This topic will be addressed in §7. The relationship between epi and pointwise convergence for convex functions is given by Wets [29]. These results are not applicable here because $\Phi^{(h)}(\xi, \cdot)$ is not in general convex.

We finish this section by remarking that the assertion “ $\{V^{(t)}\}_{t \geq 0}$ is a l.s.c. semigroup” in Theorem 3.1 is well known. The semigroup property (1.3) is in fact valid under no assumptions on L provided undefined and infinite integrals are properly interpreted. See Rockafellar [23]. For proofs under standard calculus of variation and control problem assumptions, see [3, p. 27 and p. 202]. The l.s.c. property is also valid under weaker assumptions than in force here. Indeed, in existence theory, one is often interested in the lower semicontinuity of the functional $J(x) := \int_0^T L(x(t), \dot{x}(t))dt$. From (H3), it follows that H satisfies “the stronger growth condition” in Rockafellar [24], hence by the “semicontinuity theorem” in [24], J is l.s.c. It immediately follows from this that $V^{(T)}(\xi, \cdot)$ is l.s.c.

4. YOSIDA APPROXIMATIONS

As mentioned in the introduction the allowance of L to admit infinite values adds great generality to the basic problem (1.1). But it also creates technical obstacles that are not always easily overcome. A technique employed here is to sometimes approximate L by a penalty function L_α , which is called the Yosida approximation. For each $\alpha \geq 0$, L_α is defined by

$$(4.1) \quad L_\alpha(\xi, v) = \inf_{u \in \mathbb{R}^n} \{L(\xi, u) + \alpha|u - v|\}, \quad \text{for } \xi, v \in \mathbb{R}^n.$$

The advantage in employing Yosida approximations is that L_α is finite everywhere and approximates L nicely from below when α is large. We gather technical information on L_α in the following proposition. Recall that we are assuming L satisfies (H1)–(H3).

Proposition 4.1. *For each $\alpha \geq 0$, L_α possesses the following properties:*

- (i) *For each $\xi \in \mathbb{R}^n$, $L_\alpha(\xi, \cdot)$ is a proper, closed, convex function with $\text{dom } L_\alpha(\xi, \cdot) = \mathbb{R}^n$, and $\min_{v \in \mathbb{R}^n} L_\alpha(\xi, v) = \min_{v \in \mathbb{R}^n} L(\xi, v)$.*
- (ii) *For $\xi, v \in \mathbb{R}^n$ and $\alpha' > \alpha$, we have $L_\alpha(\xi, v) \leq L_{\alpha'}(\xi, v) \leq L(\xi, v)$.*
- (iii) *Assumption (H3) holds with L replaced by L_α , and with the same λ .*
- (iv) *For $\xi, \xi', v \in \mathbb{R}^n$, $L_\alpha(\xi, v) \geq L_\alpha(\xi', v) - \lambda(1 + \alpha)|\xi - \xi'|$.*
- (v) *$\lim_{|v| \rightarrow \infty} L_\alpha(\xi, v)/|v| = \alpha$, and the limit is uniform as ξ ranges over a compact set.*
- (vi) *For $m \geq 0$, define $\mu_0(m) := \sup\{|v| : L(0, v)/|v| \leq m\} < +\infty$, and for $\alpha > m$, define $\mu_0^\alpha(m) = \sup\{|v| : L_\alpha(0, v)/|v| \leq m\}$. Then for all $m \geq 0$, we have*

$$(4.2) \quad \lim_{\alpha \rightarrow \infty} \mu_0^\alpha(m) = \mu_0(m).$$

Remark. A much stronger assertion than (4.2) is true. It can be shown that in fact the sets $\{v : \frac{1}{|v|} L_\alpha(0, v) \leq m\}$ set-converge to $\{v : \frac{1}{|v|} L(0, v) \leq m\}$.

Proof. (i) These are general properties of the infimal convolution of the convex functions $L(\xi, \cdot)$ and $\xi \mapsto \alpha|\xi|$ (see [21, p. 76]).

(ii) Immediate from the definitions.

(iii) Let $\xi, \xi', v \in \mathbb{R}^n$. Then

$$\begin{aligned} L_\alpha(\xi, v) &= \inf_u \{L(\xi, u) + \alpha|u - v|\} \\ &\geq \inf_{\substack{u \in \mathbb{R}^n \\ b \in B}} \{L(\xi', u + \lambda|\xi - \xi'|b) + \alpha|u - v|\} - \lambda|\xi - \xi'| \quad (\text{by (H3)}) \\ &= \inf_{\substack{u \in \mathbb{R}^n \\ b \in B}} \{L(\xi', u') + \alpha|u' - v| - \lambda|\xi - \xi'||b| - \lambda|\xi - \xi'| \\ &= \inf_{b \in B} L_\alpha(\xi', v + \lambda|\xi - \xi'|b) - \lambda|\xi - \xi'| \end{aligned}$$

This shows L_α satisfies (H3) with the same λ .

(iv) Let $\xi, \xi', v \in \mathbb{R}^n$. Then

$$\begin{aligned} L_\alpha(\xi, v) &\geq \inf_{\substack{u \in \mathbb{R}^n \\ b \in B}} \{L(\xi', u) + \alpha|u - v - \lambda|\xi - \xi'||b| - \lambda|\xi - \xi'|\} \quad (\text{by (iii) above}) \\ &\geq \inf_{u \in \mathbb{R}^n} \{L(\xi', u) + \alpha|u - v|\} - \lambda(1 + \alpha)|\xi - \xi'| \\ &\quad (\text{by the triangle inequality}) \\ &= L_\alpha(\xi', v) - \lambda(1 + \alpha)|\xi - \xi'|. \end{aligned}$$

(v) First, fix $\xi \in \mathbb{R}^n$. For any $u_0 \in \text{dom } L(\xi, \cdot)$ and $v \in \mathbb{R}^n$, we have

$$(4.3) \quad \frac{1}{|v|} \{L(\xi', u_0) + \alpha|u_0 - v|\} \geq \frac{1}{|v|} L_\alpha(\xi, v).$$

As $|v| \rightarrow \infty$, the left-hand side of (4.3) approaches α . Hence

$$\limsup_{|v| \rightarrow \infty} \frac{1}{|v|} L_\alpha(\xi, v) \leq \alpha.$$

We next show that $\liminf_{|v| \rightarrow \infty} \frac{1}{|v|} L_\alpha(\xi, v) \geq \alpha$. If this is not the case, then there exists $\varepsilon > 0$ and a sequence $\{v_k\}$ with $|v_k| \rightarrow \infty$ so that $\frac{1}{|v_k|} L_\alpha(\xi, v_k) \leq$

$\alpha - \varepsilon$. From (H2), it follows that the inf defining $L_\alpha(\xi, v_k)$ is attained, say at u_k . We have

$$(4.4) \quad \frac{L(\xi, u_k)}{|v_k|} + \frac{\alpha|u_k - v_k|}{|v_k|} \leq \alpha - \varepsilon$$

for all k . If a subsequence of $\{u_k\}$ (which we do not relabel) satisfies $|u_k| \rightarrow \infty$ as $k \rightarrow \infty$, in virtue of (H2), we have $\frac{1}{|u_k|} L(\xi, u_k) \rightarrow \infty$. From (4.4) it is immediate that for all k ,

$$\frac{L(\xi, u_k)}{|u_k|} \frac{|u_k|}{|v_k|} \leq \alpha - \varepsilon,$$

and hence it must follow that $\frac{|u_k|}{|v_k|} \rightarrow 0$. But now for large k , $L(\xi, u_k) \geq 0$. By letting $k \rightarrow \infty$ in (4.4), we have that $\alpha \leq \alpha - \varepsilon$, a contradiction. So the sequence $\{u_k\}$ is bounded. Since $\{L(\xi, u_k)\}$ is bounded below, letting $k \rightarrow \infty$ in (4.4) again gives us the contradiction $\alpha \leq \alpha - \varepsilon$. We conclude $\lim_{|v| \rightarrow \infty} \frac{1}{|v|} L_\alpha(\xi, v) = \alpha$. That this limit is uniform over ξ in a compact set follows from directly from (iv) above.

(vi) Fix $m \geq 0$, if $\alpha > m \geq 0$, then $\mu_0^\alpha(m) < \infty$ by (v). By (ii) we always have $\mu_0^\alpha(m) \geq \mu_0(m)$. So we must show $\limsup_{\alpha \rightarrow \infty} \mu_0^\alpha(m) \leq \mu_0(m)$. If not, then there exists $\varepsilon > 0$, a sequence $\{\alpha_k\}_k$ with $\alpha_k \nearrow \infty$ as $k \rightarrow \infty$, and a sequence $\{v_k\}_k \subseteq \mathbb{R}^n$ so that

$$(4.5) \quad \frac{L_{\alpha_k}(0, v_k)}{|v_k|} \leq m < \alpha_k$$

and

$$(4.6) \quad \mu_0(m) + \varepsilon \leq |v_k|.$$

For $k < k'$, we have

$$\frac{1}{|v_{k'}|} L_{\alpha_k}(0, v_{k'}) \leq \frac{1}{|v_{k'}|} L_{\alpha_{k'}}(0, v_{k'}) \leq m$$

by (ii). If a subsequence $\{v_{k'}\}$ satisfies $|v_{k'}| \rightarrow \infty$, then from (v) we can deduce that $\alpha_k = \lim_{k' \rightarrow \infty} \frac{1}{|v_{k'}|} L_{\alpha_k}(0, v_{k'}) \leq m$, which contradicts $\alpha_k > m$. Hence the sequence $\{v_k\}$ is bounded, and without loss of generality we assume $v_k \rightarrow v_0$. Let $u_k \in \mathbb{R}^n$ be such that $L_{\alpha_k}(0, v_k) = L(0, u_k) + \alpha_k|u_k - v_k|$. Then

$$(4.7) \quad \begin{aligned} |v_k - u_k| &= \frac{1}{\alpha_k} (L_{\alpha_k}(0, v_k) - L(0, u_k)) \\ &\leq \frac{1}{\alpha_k} \left(|v_k|m - \inf_u L(0, u) \right) \quad (\text{by (4.5)}) \\ &\rightarrow 0 \quad \text{as } k \rightarrow \infty. \end{aligned}$$

Hence (4.7) implies $u_k \rightarrow v_0$. By the lower semicontinuity of $L(0, \cdot)$ we have

$$(4.8) \quad \frac{L(0, v_0)}{|v_0|} \leq \liminf_{k \rightarrow \infty} \frac{L(0, u_k)}{|u_k|} \leq \limsup_{k \rightarrow \infty} \frac{L_{\alpha_k}(0, v_k)}{|v_k|} \leq m.$$

The last two inequalities of (4.8) follow from the definition of u_k and from (4.5), respectively. Finally, (4.8) says by definition that $|v_0| \leq \mu_0(m)$, which contradicts (4.6) because $v_k \rightarrow v_0$. \square

The next simple lemma provides an estimate that will be used extensively below. In (4.9), note that L_α cannot be replaced by L because L may take on infinite values.

Lemma 4.2. *Fix $T > 0$ and $\xi \in \mathbb{R}^n$. Suppose $x(\cdot) \in AC[0, T]$, and define $\phi(t) = \int_0^t L(x(s), \dot{x}(s))ds$.*

Then for each $\alpha \geq 0$ and $0 \leq t_0 \leq t \leq T$, we have

$$(4.9) \quad \phi(t_1) \geq \phi(t_0) + (t_1 - t_0) \left\{ L_\alpha \left(\xi, \frac{x(t_1) - x(t_0)}{t_1 - t_0} \right) - \lambda(1 + \alpha) \sup_{t_0 \leq t \leq t_1} |x(t) - \xi| \right\}$$

Proof. We assume $\phi(t_1) < \infty$ and $t_0 < t_1$, for the result is trivial otherwise. We have

$$\begin{aligned} \phi(t_1) - \phi(t_0) &= \int_{t_0}^{t_1} L(x(t), \dot{x}(t))dt \\ &\geq \int_{t_0}^{t_1} L_\alpha(x(t), \dot{x}(t))dt \quad (\text{by Proposition 4.1(ii)}) \\ &\geq \int_{t_0}^{t_1} L_\alpha(\xi, \dot{x}(t))dt - \lambda(1 + \alpha) \int_{t_0}^{t_1} |x(t) - \xi|dt \\ &\hspace{40em} (\text{by Proposition 4.1(iv)}) \\ &\geq (t_1 - t_0) \left\{ L_\alpha \left(\xi, \frac{x(t_1) - x(t_0)}{t_1 - t_0} \right) - \lambda(1 + \alpha) \sup_{t_0 \leq t \leq t_1} |x(t) - \xi| \right\} \\ &\hspace{58em} (\text{by Jensen's inequality}). \quad \square \end{aligned}$$

We next give a new proof of a special case of a result due to Rockafellar [24]. This result will feature prominently in our analysis, and the proof given here will also illustrate the utility of Yosida approximations.

Proposition 4.3. *Suppose $T > 0$, $m > 0$, and $K \subseteq \mathbb{R}^n$ is compact. Then there exists a constant c (depending on T , m , K) so that if $x(\cdot) \in AC[0, T]$ satisfies $x(0) \in K$ and $\int_0^T L(x(t), \dot{x}(t))dt \leq m$, then $x(\cdot)$ also satisfies $\|x(\cdot) - x(0)\| \leq c$.*

Proof. We first assume $T < \frac{1}{\lambda}$. If the result is false, then there exist for each $k = 1, 2, \dots$, an absolutely continuous function $x_k(\cdot)$ on $[0, T]$ for which the following conditions (4.10)–(4.12) hold:

$$(4.10) \quad x_k(0) =: \xi_k \in K$$

Define $\phi_k(t) := \int_0^t L(x_k(s), \dot{x}_k(s))ds$.

$$(4.11) \quad \phi_k(T) \leq m.$$

Let $t_k \in [0, T]$ so that $|x_k(t_k) - \xi_k| = \sup_{0 \leq t \leq T} |x_k(t) - \xi_k|$.

$$(4.12) \quad |x_k(t_k) - \xi_k| \rightarrow \infty \quad \text{as } k \rightarrow \infty.$$

By Lemma 4.2 with $t_0 = 0$, $t_1 = t_k$, and $\xi = \xi_k$, we have for all $\alpha \geq 0$,

$$\phi_k(t_k) \geq t_k \left\{ L_\alpha \left(\xi_k, \frac{x_k(t_k) - \xi_k}{t_k} \right) - \lambda(1 + \alpha)|x_k(t_k) - \xi_k| \right\}.$$

Dividing by $|x_k(t_k) - \xi_k|$ gives

$$(4.13) \quad \frac{\phi_k(t_k)}{|x_k(t_k) - \xi_k|} \geq \frac{t_k}{|x_k(t_k) - \xi_k|} L_\alpha \left(\xi_k, \frac{x_k(t_k) - \xi_k}{t_k} \right) - t_k \lambda(1 + \alpha).$$

Since $0 < t_k \leq T$, from (4.12) we have that $\frac{|x_k(t_k) - \xi_k|}{t_k} \rightarrow \infty$ as $k \rightarrow \infty$. So letting $k \rightarrow \infty$ in (4.13), we deduce from (4.10) and Proposition 4.1(v) that

$$(4.14) \quad \liminf_{k \rightarrow \infty} \frac{\phi_k(t_k)}{|x_k(t_k) - \xi_k|} \geq \alpha - T\lambda(1 + \alpha) = (1 - T\lambda)\alpha - T\lambda.$$

Now $\alpha > 0$ is arbitrary in (4.14), and we assumed $(1 - T\lambda) > 0$, hence

$$(4.15) \quad \lim_{k \rightarrow \infty} \frac{\phi_k(t_k)}{|x_k(t_k) - \xi_k|} = \infty.$$

Again we use Lemma 4.2, this time with $\alpha = 0$, $t_0 = t_k$, $t_1 = T$ and $\xi = \xi_k$:

$$(4.16) \quad \phi_k(T) - \phi_k(t_k) \geq (T - t_k) \left\{ \inf_v L(\xi_k, v) - \lambda|x_k(t_k) - \xi_k| \right\}.$$

We have

$$\begin{aligned} -\infty &= \lim_{k \rightarrow \infty} \frac{\phi_k(T) - \phi_k(t_k)}{|x_k(t_k) - \xi_k|} \quad (\text{by (4.11), (4.12) and (4.15)}) \\ &\geq \lim_{k \rightarrow \infty} \frac{T - t_k}{|x_k(t_k) - \xi_k|} \inf_v L(\xi_k, v) - \lambda T \quad (\text{by (4.16)}) \\ &= -\lambda T \quad \left(\text{by (4.12) and since } \inf_{\substack{v \in \mathbb{R}^n \\ \xi \in K}} L(\xi, v) > -\infty \right). \end{aligned}$$

This contradiction finishes the proof whenever $T < \frac{1}{\lambda}$. However these values of T do not depend on the size of K or m , so we actually have the conclusion for all $T > 0$. \square

5. PROOFS

In this section, we prove Theorems 3.1 and 3.2. The Yosida approximations will be used in the first part of the proof of Theorem 3.1, and then the rest of the proofs will involve a reformulation of the basic problem into a differential inclusion. We begin with a definition.

Definition 5.1. Let $K \subseteq \mathbb{R}^n$ be compact, $m > 0$, and $T > 0$. Define

$$(5.1) \quad C(K, m, T) := \inf \left\{ c \geq 0 : \text{for all } x(\cdot) \in AC[0, T] \right. \\ \text{with } x(0) \in K \text{ and } \int_0^T L(x(t), \dot{x}(t)) dt \leq m, \\ \left. \text{we also have } \|x(\cdot) - x(0)\| \leq c \right\}.$$

The utility of Yosida approximations in this paper is contained in the proof of the following important proposition.

Proposition 5.2. *There exists a function $\mu : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ so that for all $m \geq 0$ and $\rho \geq 0$, we have*

$$(5.2) \quad \limsup_{T \searrow 0} \frac{C(\rho B, mT, T)}{T} \leq \mu(m, \rho).$$

Proof. Let $l_0 = \inf_v L(0, v)$ and $m_0 = |l_0| + 1$.

For $m \geq m_0$, define $\mu_0(m) = \max\{m_0, \sup\{|v| : \frac{L(0, v)}{|v|} \leq m\}\}$. This is the same μ_0 as defined in Proposition 4.1, except that we now ensure $\mu_0(m) \geq m_0$. For $m \geq m_0$ and $\rho \geq 0$, define

$$(5.3) \quad \mu(m, \rho) = \mu_0(m(1 + \lambda\rho) + \lambda\rho) + \lambda\rho + 1.$$

Define $\mu(m, \rho) = \mu(m_0, \rho)$ whenever $m < m_0$. We will show that (5.2) holds for this μ . It suffices to consider only the case when $m \geq m_0$.

Fix $m \geq m_0$ and $\rho \geq 0$. We claim that there exists $\alpha > 0$ so that

$$(5.4) \quad |\xi| \leq \rho \quad \text{and} \quad \frac{L_\alpha(\xi, v)}{|v|} \leq m \quad \text{implies} \quad |v| \leq \mu(m, \rho).$$

To prove (5.4), first define

$$\mu_0^\alpha(m') = \max \left\{ m_0, \sup \left\{ |v| : \frac{L_\alpha(0, v)}{|v|} \leq m' \right\} \right\}$$

for α large and $m' \geq m_0$. By Proposition 4.1(vi), there exists large α so that

$$(5.5) \quad \mu_0^\alpha(m(\lambda\rho + 1) + \lambda\rho) \leq \mu_0(m(\lambda\rho + 1) + \lambda\rho) + 1.$$

Now suppose $|\xi| \leq \rho$ and $\frac{L_\alpha(\xi, v)}{|v|} \leq m$. By Proposition 4.1(iii), there exists $v' \in v + \lambda\rho B$ so that

$$(5.6) \quad L_\alpha(0, v') \leq L_\alpha(0, v) + \lambda\rho.$$

If $|v'| \leq 1$, then $|v| \leq \lambda\rho + 1 \leq \mu(m, \rho)$. So suppose $|v'| > 1$. Then $\frac{|v|}{|v'|} \leq 1 + \frac{|v - v'|}{|v'|} \leq \lambda\rho + 1$, and hence

$$(5.7) \quad \begin{aligned} \frac{L_\alpha(0, v')}{|v'|} &\leq \frac{L_\alpha(\xi, v)}{|v|} \frac{|v|}{|v'|} + \frac{\lambda\rho}{|v'|} \quad (\text{by (5.6)}) \\ &\leq m(\lambda\rho + 1) + \lambda\rho. \end{aligned}$$

From (5.7) and the definition of μ_0^α , we have $|v'| \leq \mu_0^\alpha(m(\lambda\rho + 1) + \lambda\rho)$, which by (5.5) is $\leq \mu_0(m(\lambda\rho + 1) + \lambda\rho) + 1$. From (5.7), we have

$$|v| \leq |v'| + |v - v'| \leq \mu_0(m(\lambda\rho + 1) + \lambda\rho) + 1 + \lambda\rho = \mu(m, \rho).$$

Therefore the claim (5.4) is verified.

Now suppose the conclusion of the proposition is false for our choice of $\mu(m, \rho)$ as defined in (5.3). Then there exists $m \geq m_0$, $\rho \geq 0$, and sequences $T_k \searrow 0$ and $x_k(\cdot) \in AC[0, T_k]$ so that the following (5.8)–(5.10) hold. To simplify the notation, set $\xi_k = x_k(0)$, define $\phi_k(t) := \int_0^t L(x_k(s), x_k(s)) ds$, and let $t_k \in [0, T_k]$ be such that $|x_k(t_k) - \xi_k| = \sup \{|x_k(t) - \xi_k| : 0 \leq t \leq T_k\}$.

$$(5.8) \quad |\xi_k| \leq \rho,$$

$$(5.9) \quad \frac{\phi_k(T_k)}{T_k} \leq m,$$

$$(5.10) \quad \left| \frac{x_k(t_k) - \xi_k}{T_k} \right| > \mu(m, \rho).$$

Let $\alpha > 0$ be chosen so that (5.4) holds. By Lemma 4.2 applied to $t_1 = t_k$, $t_0 = 0$, and $\xi = \xi_k$, we have

$$(5.11) \quad \phi_k(t_k) \geq t_k \left\{ L_\alpha \left(\xi_k, \frac{x_k(t_k) - \xi_k}{t_k} \right) - \lambda(1 + \alpha)|x_k(t_k) - \xi_k| \right\}.$$

Dividing (5.11) by $|x_k(t_k) - \xi_k|$ gives

$$(5.12) \quad \frac{\phi_k(t_k)}{|x_k(t_k) - \xi_k|} \geq \frac{L_\alpha(\xi_k, v_k)}{|v_k|} - t_k \lambda(1 + \alpha),$$

where $v_k := \frac{x_k(t_k) - \xi_k}{t_k}$. Since $t_k < T_k$, from (5.10) we have that $|v_k| > \mu(m, \rho)$, whence from (5.4) it follows that $\frac{L_\alpha(\xi_k, v_k)}{|v_k|} > m$. Let k approach ∞ in (5.12). We conclude

$$(5.13) \quad \liminf_{k \rightarrow \infty} \frac{\phi_k(t_k)}{|x_k(t_k) - \xi_k|} \geq m.$$

Again we resort to Lemma 4.2, this time with $t_1 = T_k$, $t_0 = t_k$, $\xi = \xi_k$, and $\alpha = 0$;

$$(5.14) \quad \begin{aligned} \phi_k(T_k) - \phi_k(t_k) &\geq (T_k - t_k) \left\{ \inf_v L(\xi_k, v) - \lambda|x_k(t_k) - \xi_k| \right\} \\ &\geq T_k \{-|\ell_0| - \lambda\rho - \lambda|x_k(t_k) - \xi_k|\} \quad (\text{by H3}) \end{aligned}$$

Next we put things together:

$$\begin{aligned} \frac{m}{\mu(m, \rho)} - m &\geq \limsup_{r \rightarrow \infty} \left\{ \frac{mT_k}{|x_k(t_k) - \xi_k|} - \frac{\phi_k(t_k)}{|x_k(t_k) - \xi_k|} \right\} \\ &\quad (\text{by (5.10) and (5.13)}) \\ &\geq \limsup_{k \rightarrow \infty} \left\{ \frac{\phi_k(T_k) - \phi_k(t_k)}{|x_k(t_k) - \xi_k|} \right\} \quad (\text{by (5.9)}) \\ &\geq \limsup_{k \rightarrow \infty} \left\{ \frac{-T_k}{|x_k(t_k) - \xi_k|} (|\ell_0| + \lambda\rho) - T_k \lambda \right\} \quad (\text{by (5.14)}) \\ &\geq -\frac{1}{\mu(m, \rho)} (|\ell_0| + \lambda\rho) \quad (\text{by (5.10)}). \end{aligned}$$

Therefore $m(\mu(m, \rho) - 1) \leq |\ell_0| + \lambda\rho$, which contradicts the choice of $\mu(m, \rho)$ in (5.3). This concludes the proof of Proposition 5.2. \square

Proof of Theorem 3.1. Let $K \subseteq \mathbb{R}^n$ be compact, $m > 0$, and $\varepsilon > 0$. Recall that $\Phi^{(t)}(\xi, v) = \frac{1}{t}V^{(t)}(\xi, \xi + tv)$. We first show that there exists $T > 0$ so that

$$(5.15) \quad \text{epi } \Phi^{(t)}(\xi, \cdot) \cap mB \subseteq \text{epi } L(\xi, \cdot) + \varepsilon B$$

for all $\xi \in K$ and $0 < t \leq T$.

Set $\rho = \sup\{|\xi| : \xi \in K\}$. By Proposition 5.2, there exists $T_0 > 0$ so that $0 < t \leq T_0$ implies $C(\rho B, mt, t) \leq (\mu(m, \rho) + 1)t$. We take

$$T := \min \left\{ T_0, \frac{\varepsilon}{(\mu(m, \rho) + 1)\lambda} \right\}.$$

Now suppose $\xi \in K$, $0 < t \leq T$, and $\Phi^{(t)}(\xi, v) \leq m$. By the existence theory in [21], there exists $x(\cdot) \in AC[0, t]$ with $x(0) = \xi$, $x(t) = \xi + tv$ and satisfying

$$\int_0^t L(x(s), \dot{x}(s)) ds = V^{(t)}(\xi, \xi + tv) \leq mt.$$

By the definition (5.1) of $C(\rho B, mt, t)$ and the choice of T , we have

$$(5.16) \quad |x(s) - \xi| \leq C(\rho B, mt, t) \leq \frac{\varepsilon}{\lambda}$$

for all $0 \leq s \leq t$. Hence

$$(5.17) \quad \begin{aligned} \Phi^{(t)}(\xi, v) &= \frac{1}{t} \int_0^t L(x(s), \dot{x}(s)) ds \\ &\geq \frac{1}{t} \int_0^t \inf_{b \in B} \{L(\xi, \dot{x}(s) + \lambda|x(s) - \xi|b)\} ds \\ &\quad - \lambda \frac{1}{t} \int_0^t |x(s) - \xi| ds \quad (\text{by (H3)}) \\ &\geq \frac{1}{t} \int_0^t \inf_{b \in B} L(\xi, \dot{x}(s) + \varepsilon b) ds - \varepsilon \quad (\text{by 5.16}). \end{aligned}$$

The function $(s, b) \mapsto L(\xi, \dot{x}(s) + \varepsilon b)$ is a normal integrand on $[0, t] \times B$ in the sense of Rockafellar [25]. By [25, Theorem 2K], the map

$$s \mapsto \arg \min_{b \in B} L(\xi, \dot{x}(s) + \varepsilon b)$$

is a measurable multifunction, and by (H2) it has domain of full measure. Therefore there exists a measurable selection $b(s)$ (cf. [25, Corollary 1C]). Plugging $b(s)$ into (5.17), we obtain

$$(5.18) \quad \begin{aligned} \Phi^{(t)}(\xi, v) &\geq \frac{1}{t} \int_0^t L(\xi, \dot{x}(s) + \varepsilon b(s)) ds - \varepsilon \\ &\geq L\left(\xi, v + \varepsilon \frac{1}{t} \int_0^t b(s) ds\right) - \varepsilon. \end{aligned}$$

The last inequality follows from Jensen's inequality and that $x(\cdot)$ was chosen so that $v = \frac{x(t) - \xi}{t}$. By setting $u = v + \varepsilon \frac{1}{t} \int_0^t b(s) ds$, we have by (5.18) that

$$(5.19) \quad |u - v| \leq \varepsilon \quad \text{and} \quad \Phi^{(t)}(\xi, v) \geq L(\xi, u) - \varepsilon.$$

Finally, since the only restriction on v was that $\Phi^{(t)}(\xi, v) \leq m$, we conclude from (5.19) that (5.15) holds.

Next we show that there exists $T > 0$ for which the opposite inclusion

$$(5.20) \quad \text{epi } L(\xi, \cdot) \cap mB \subseteq \text{epi } \Phi^{(t)}(\xi, \cdot) + \varepsilon B$$

holds for all $0 < t \leq T$ and $\xi \in K$.

Let $K_m = \{\xi \in K : \text{epi } L(\xi, \cdot) \cap mB \neq \emptyset\}$. If $K_m = \emptyset$, then (5.20) is trivial. So assume $K_m \neq \emptyset$ and define $F_m: \mathbb{R}^n \times \mathbb{R}^1 \rightrightarrows \mathbb{R}^{n+1}$ by $F_m(\xi, \phi) = \text{epi } L(\xi, \cdot) \cap (\mathbb{R}^n \times (m+1)B)$. (F_m is independent of the last coordinate $\phi \in \mathbb{R}^1$). As a consequence of (H3), there exists $\delta > 0$ so that $F_m(\xi, \phi) \neq \emptyset$ for $(\xi, \phi) \in (K_m + \delta B) \times \mathbb{R}^1$. It is an easy matter to check that F_m satisfies the hypotheses in [32, §3]. We apply [32, Lemma 5.4] with (in the notation of [32])

$X = \text{interior of } (K_m + \delta B) \times \mathbb{R}^1$, $K = K_m \times \{0\}$, and $F = F_m$. We conclude that there exists $T > 0$ so that for any $\xi \in K_m$ and $(v, r) \in F_m(\xi, 0)$, there exists $(x(\cdot), \phi(\cdot)) \in AC[0, T]$ for which the following hold:

$$(5.21) \quad \begin{aligned} (\dot{x}(t), \dot{\phi}(t)) &\in F_m(x(t), \phi(t)) \quad \text{a.e. } t \in [0, T], \\ x(0) &= \xi, \quad \phi(0) = 0, \end{aligned}$$

$$(5.22) \quad \left| \frac{x(t) - \xi}{t} - v \right| < \varepsilon \quad \text{for all } t \in [0, T],$$

$$(5.23) \quad \left| \frac{\phi(t)}{t} - r \right| < \varepsilon \quad \text{for all } t \in [0, T].$$

By (5.21) and the definition of F_m , we have

$$\dot{\phi}(t) \geq L(x(t), \dot{x}(t)) \quad \text{a.e. } 0 \leq t \leq T,$$

which after integrating from 0 to t becomes:

$$(5.24) \quad \phi(t) \geq \int_0^t (x(s), \dot{x}(s)) ds \quad \text{for all } t \in [0, T].$$

Now let $\xi \in K$ and $(v, r) \in \text{epi } L(\xi, \cdot) \cap mB$ ($\subseteq F_m(\xi, 0)$). Then $\xi \in K_m$. Let $(x(\cdot), \phi(\cdot))$ be chosen so that (5.21)–(5.23) hold, and set $v(t) = \frac{x(t) - \xi}{t}$. Then for all $0 \leq t \leq T$, we have

$$\begin{aligned} r &\geq \frac{\phi(t)}{t} - \varepsilon \quad (\text{by (5.23)}) \\ &\geq \frac{1}{t} \int_0^t L(x(s), \dot{x}(s)) ds - \varepsilon \quad (\text{by (5.24)}) \\ &\geq \frac{1}{t} V^{(t)}(\xi, x(t)) - \varepsilon \\ &= \Phi^{(t)}(\xi, v(t)) - \varepsilon. \end{aligned}$$

Conjoining (5.22) with the last line, we deduce that

$$(v, r) \in \text{epi } \Phi^{(t)}(\xi, \cdot) + \varepsilon B$$

for all $0 < t \leq T$. Recall that T was chosen independantly of $\xi \in K$, hence (5.20) has been verified.

In light of the comments immediately succeeding the definition of d in §2, the verification of (5.15) and (5.20) has completed the proof of Theorem 3.1. \square

Proof of Theorem 3.2. We now suppose $\{U^{(t)}\}_{t \geq 0}$ is a l.s.c. semigroup for which (3.2) holds. Fix $\xi, \eta \in \mathbb{R}^n$ and $t > 0$. We must show $U^{(t)}(\xi, \eta) \leq V^{(t)}(\xi, \eta)$.

The method of proof uses a reformulation of (1.1) into a differential inclusion similar to the second half of the proof of Theorem 3.1. This proof resembles the proof of Theorem 3.2(i) in [32], but for the purposes of completeness and clarity, the full details are given here.

Define $F : \mathbb{R}^n \times \mathbb{R}^1 \rightrightarrows \mathbb{R}^{n+1}$ by $F(\xi, \phi) = \text{epi } L(\xi, \cdot)$. For $\xi \in \mathbb{R}^n$, define $R^{(T)}(\xi) := \{(x(T), \phi(T)) : (x(\cdot), \phi(\cdot)) \in AC[0, T], (\dot{x}(t), \dot{\phi}(t)) \in F(x(t), \phi(t)) \text{ a.e. } t \in [0, T], x(0) = \xi, \phi(0) = 0\}$. Let $\varepsilon > 0$.

Note that F has closed convex values, and satisfies $F(\xi, \phi) \subseteq F(\xi', \phi') + \lambda|\xi - \xi'|B$. We assume $V^{(T)}(\xi, \eta) < +\infty$, for otherwise there is nothing to prove. This assumption is equivalent to $(\eta, V^{(T)}(\xi, \eta)) \in R^{(T)}(\xi)$. We will apply a C^1 approximation result due to Filippov [13] (see [31, Proposition 3.2] for a straightforward proof of this result in the case where F takes compact values, and [30, Proposition 2.3] for the general case). It follows from [13, Theorem 6.1] that there exists C^1 functions $x(\cdot)$ and $\phi(\cdot)$ into \mathbb{R}^n and \mathbb{R}^1 respectively so that

$$(5.25) \quad x(0) = \xi, \quad \phi(0) = 0,$$

$$(5.26) \quad |x(T) - \eta| < \varepsilon,$$

$$(5.27) \quad \phi(T) - V^{(T)}(\xi, \eta) < \varepsilon,$$

$$(5.28) \quad \dot{\phi}(t) \geq L(x(t), \dot{x}(t)) \quad \text{for } t \in [0, T].$$

Set $\rho = \sup_{0 \leq t \leq T} |x(t)| + \frac{2\varepsilon}{\lambda} e^{\lambda T}$ and $m = \sup_{0 \leq t \leq T} |\dot{\phi}(t)| + 2\varepsilon e^{\lambda T}$. Recall that $\Psi^{(t)}(\gamma, v) := \frac{1}{t} U^{(t)}(\gamma, \gamma + tv)$. By assumption (3.2), there exists N large so that if $0 < t \leq \frac{T}{N}$ and $\gamma \in \rho B$, then we have

$$(5.29) \quad \text{epi } L(\gamma, \cdot) \cap mB \subseteq \text{epi } \Psi^{(t)}(\gamma, \cdot) + \varepsilon B.$$

Since $\dot{x}(\cdot)$ and $\dot{\phi}(\cdot)$ are continuous, N can be enlarged if necessary to ensure that if $0 \leq s \leq \frac{T}{N}$ and $0 \leq t \leq T - \frac{T}{N}$, then

$$(5.30) \quad \left| \frac{x(t+s) - x(t)}{s} - \dot{x}(t) \right| < \varepsilon$$

and

$$(5.31) \quad \left| \frac{\phi(t+s) - \phi(t)}{s} - \dot{\phi}(t) \right| < \varepsilon$$

both hold. Set $h = \frac{T}{N}$ and $t_j = jh$. Define sequences $\{y(t_j)\}_{j=0}^N$, $\{\dot{y}(t_j)\}_{j=0}^N$ in \mathbb{R}^n and $\{\psi(t_j)\}_{j=0}^N$, $\{\dot{\psi}(t_j)\}_{j=0}^N$ in \mathbb{R}^1 inductively as follows: (the dots on \dot{y} and $\dot{\psi}$ here do not denote derivatives, but do suggest how they will be used). Let $y(t_0) = x(t_0)$, $\dot{y}(t_0) = \dot{x}(t_0)$, $\psi(t_0) = \phi(t_0)$, $\dot{\psi}(t_0) = \dot{\phi}(t_0)$. Suppose for $j = k \leq N - 1$, the vectors and numbers are defined so that

$$(5.32) \quad \left| \dot{y}(t_{k-1}) - \frac{y(t_k) - y(t_{k-1})}{h} \right| \leq \varepsilon,$$

$$(5.33) \quad \frac{1}{h} U^{(h)}(y(t_{k-1}), y(t_k)) \leq L(y(t_{k-1}), \dot{y}(t_{k-1})) + \varepsilon,$$

$$(5.34) \quad \dot{\psi}(t_k) \geq L(y(t_k), \dot{y}(t_k)),$$

$$(5.35) \quad \max\{|\dot{x}(t_k) - \dot{y}(t_k)|, |\dot{\phi}(t_k) - \dot{\psi}(t_k)|\} \leq \lambda|x(t_k) - y(t_k)|,$$

$$(5.36) \quad \max\{|x(t_k) - y(t_k)|, |\phi(t_k) - \psi(t_k)|\} \leq 2h\varepsilon \left(\frac{1 - \alpha^k}{1 - \alpha} \right),$$

where $\alpha = 1 + h\lambda$.

When $k = 0$, (5.32)–(5.35) are either vacuous or trivial.

The right-hand side of (5.36) equals $\frac{2\epsilon}{\lambda}(1 + h\lambda)^k - 1$, which is always less than or equal to $\frac{2\epsilon}{\lambda}e^{\lambda T}$. Therefore (5.34), (5.36), and the definitions of m and ρ imply that $L(y(t_k), \dot{y}(t_k)) \leq m$ and $|y(t_k)| \leq \rho$. Consequently from (5.29) with $\gamma = y(t_k)$, $y(t_{k+1})$ can be chosen to satisfy

$$(5.37) \quad \left| \dot{y}(t_k) - \frac{y(t_{k+1}) - y(t_k)}{h} \right| \leq \epsilon$$

and

$$(5.38) \quad \frac{1}{h} U^{(h)}(y(t_k), y(t_{k+1})) \leq L(y(t_k), \dot{y}(t_k)) + \epsilon.$$

By assumption (H3),

$$\text{epi } L(x(t_{k+1}), \cdot) \subseteq \text{epi } L(y(t_{k+1}), \cdot) + \lambda|x(t_{k+1}) - y(t_{k+1})|B.$$

Since $(\dot{x}(t_{k+1}), \dot{\phi}(t_{k+1})) \in \text{epi } L(x(t_{k+1}), \cdot)$, there exists $\dot{y}(t_{k+1})$ and $\dot{\psi}(t_{k+1})$ so that

$$(5.39) \quad (\dot{y}(t_{k+1}), \dot{\psi}(t_{k+1})) \in \text{epi } L(y(t_{k+1}), \cdot)$$

and

$$(5.40) \quad \max\{|\dot{x}(t_{k+1}) - \dot{y}(t_{k+1})|, |\dot{\phi}(t_{k+1}) - \dot{\psi}(t_{k+1})|\} \leq \lambda|x(t_{k+1}) - y(t_{k+1})|.$$

Also we have

$$\begin{aligned} |x(t_{k+1}) - y(t_{k+1})| &\leq |x(t_{k+1}) - x(t_k) - h\dot{x}(t_k)| + |x(t_k) - y(t_k)| \\ &\quad + h|\dot{x}(t_k) - \dot{y}(t_k)| + |y(t_k) + h\dot{y}(t_k) - y(t_{k+1})| \\ &\leq h\epsilon + 2h\epsilon \left(\frac{1 - \alpha^k}{1 - \alpha} \right) + 2h^2\lambda\epsilon \left(\frac{1 - \alpha^k}{1 - \alpha} \right) + h\epsilon \end{aligned}$$

by (5.30), (5.36), (5.35) and (5.36), (5.32), respectively

$$= 2h\epsilon \left(\frac{1 - \alpha^{k+1}}{1 - \alpha} \right).$$

Replacing x'^s by ϕ'^s and y'^s by ψ'^s in the last calculation gives a corresponding estimate for $|\phi(t_{k+1}) - \psi(t_{k+1})|$. Hence

$$(5.41) \quad \max\{|x(t_{k+1}) - y(t_{k+1})|, |\phi(t_{k+1}) - \psi(t_{k+1})|\} \leq 2h\epsilon \left(\frac{1 - \alpha^{k+1}}{1 - \alpha} \right).$$

The lines (5.37)–(5.41) restate (5.32)–(5.36) for the case $j = k + 1$. Now

putting things together, we have

$$\begin{aligned}
 (5.42) \quad U^{(T)}(y(t_0), y(t_N)) &\leq \sum_{k=0}^{N-1} U^{(h)}(y(t_k), y(t_{k+1})) \quad (\text{by the semigroup property}) \\
 &\leq \left\{ \sum_{k=0}^{N-1} hL(y(t_k), \dot{y}(t_k)) \right\} + \varepsilon T \quad (\text{by (5.33)}) \\
 &\leq \left\{ \sum_{k=0}^{N-1} h\{\dot{\phi}(t_k) + \lambda|x(t_k) - y(t_k)|\} \right\} + \varepsilon T \\
 &\qquad\qquad\qquad (\text{by (5.34) and (5.35)}) \\
 &\leq \left\{ \sum_{k=0}^{N-1} h\dot{\phi}(t_k) \right\} + 2\varepsilon e^{\lambda T} T + \varepsilon T \quad (\text{by (5.36)}).
 \end{aligned}$$

Of course the y' s constructed above depend on ε and N . Let $y_{N,\varepsilon}$ denote the end point $y(t_N)$ so constructed, and let $(x_\varepsilon, \phi_\varepsilon)$ be the (x, ϕ) used in (5.25)–(5.28). We have that

$$\begin{aligned}
 (5.43) \quad \limsup_{N \rightarrow \infty} |y_{N,\varepsilon} - \eta| &\leq \limsup_{N \rightarrow \infty} |y_{N,\varepsilon} - x_\varepsilon(T)| + |x_\varepsilon(T) - \eta| \\
 &\leq \frac{2\varepsilon e^{\lambda T}}{\lambda} + \varepsilon \quad (\text{by (5.36) and (5.26)}).
 \end{aligned}$$

All the pieces are finally in place. We have

$$\begin{aligned}
 U^{(T)}(\xi, \eta) &\leq \liminf_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} U^{(T)}(\xi, y_{N,\varepsilon}) \quad (\text{by l.s.c. of } U^{(T)} \text{ and (5.43)}) \\
 &\leq \liminf_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \left\{ \sum_{k=0}^{N-1} \frac{T}{N} \dot{\phi}_\varepsilon(t_k) + 2\varepsilon e^{\lambda T} T + \varepsilon T \right\} \quad (\text{by (5.42)}) \\
 &\leq \liminf_{\varepsilon \rightarrow 0} \left\{ \phi_\varepsilon(T) + 2\varepsilon e^{\lambda T} + \varepsilon \right\} \\
 &\qquad\qquad\qquad (\text{since } \dot{\phi}(\cdot) \text{ is continuous and } \phi(0) = 0) \\
 &= V^{(T)}(\xi, \eta) \quad (\text{by (5.27)}).
 \end{aligned}$$

This concludes the proof of Theorem 3.2. \square

6. FINITE LAGRANGIANS

The problem (1.1) with L finite-valued is known as the basic problem in the calculus of variations. Its research history is indeed a long one. This section singles out finite Lagrangians for special treatment, although we draw upon the previous material in the proofs and still use epigraphical convergence. Pointwise limit results will appear in §7, and will be seen to be pertinent to finite Lagrangian problems. The manner of presentation is chosen so as to unify the treatment of classical finite-valued problems with possibly infinite-valued ones. In this section, we assume in addition to (H1)–(H3) that L is finite everywhere. It follows from (H3) and the convexity assumption in (H1) that L is locally Lipschitz on $\mathbb{R}^n \times \mathbb{R}^n$. Our main result here is Theorem 6.2, which reverses the

inequality in (3.3). Hence a characterization of the value function semigroup is obtained among all l.s.c. semigroups with the same infinitesimal generator, but which also must satisfy an additional property.

Definition 6.1. Suppose $\{U^{(T)}\}_{T \geq 0}$ is a l.s.c. semigroup and $\xi, \eta \in \mathbb{R}^n$ and $T > 0$. Then $x(\cdot) \in AC[0, T]$ is *optimal with respect to U* (or optimal w.r.t. U) at (ξ, η, T) if $x(0) = \xi$, $x(T) = \eta$, and

$$U^{(t-s)}(x(s), x(t)) = U^{(h)}(x(s), x(s+h)) + U^{(t-s-h)}(x(s+h), x(t))$$

for all $0 \leq s \leq t \leq T$ and $0 < h < t - s$.

Theorem 6.2. Assume L is finite valued and satisfies (H1)–(H3). Suppose that $\{U^{(T)}\}_{T \geq 0}$ is a l.s.c. semigroup for which

$$(6.1) \quad d(\text{epi } \Psi^{(h)}(\xi, \cdot), \text{epi } L(\xi, \cdot)) \rightarrow 0 \quad \text{as } h \searrow 0$$

uniformly for ξ in a compact set, where $\Psi^{(h)}(\xi, v) := \frac{1}{h}U^{(h)}(\xi, \xi + hv)$. Then for each $(\xi, \eta, T) \in \mathbb{R}^n \times \mathbb{R}^n \times [0, \infty]$ for which there exists an $x(\cdot)$ Lipschitz on $[0, T]$ that is optimal w.r.t. U at (ξ, η, T) , we have

$$U^{(T)}(\xi, \eta) = V^{(T)}(\xi, \eta).$$

It is difficult to ascertain the cruciality of hypothesizing the existence of a Lipschitz $x(\cdot)$ in the statement of the theorem. Indeed, it is only verified in the case $U = V$ by quoting a deep result of Clarke and Vinter. The regularity theorem [7, Corollary 3.1] states that an optimal solution to (1.1) is necessarily Lipschitz. Also, see [8] for related results of this type. The principle of optimality directly gives that an optimal solution to (1.1) is optimal w.r.t. V as defined in Definition 6.1. Hence the value function semigroup satisfies the conditions of the theorem for all $(\xi, \eta, T) \in \mathbb{R}^n \times \mathbb{R}^n \times [0, \infty)$. It seems to be no easy matter to construct a l.s.c. semigroup for which (6.1) holds but the added hypothesis fails. On the other hand, our proof that $U = V$ depends heavily upon it.

Another interesting question is whether an $x(\cdot)$ optimal w.r.t. V must solve (1.1). If $x(\cdot)$ is in addition Lipschitz, then this is true, and can be proved using Proposition 6.3 below and Theorem 3.1. As just recalled from Clarke and Vinter [7], all solutions of (1.1) are Lipschitz, but the question remains open whether an absolutely continuous $x(\cdot)$ that is not Lipschitz can be optimal w.r.t. V . We suspect the answer to be no, but do not believe the methods employed here can show this.

In the next proposition, we will only use the “only if” implication, but the equivalence shows that (6.2) could be assumed in Theorem 6.2 in place of the Lipschitz assumption on $x(\cdot)$ without changing the content of the theorem.

Proposition 6.3. Suppose L and $\{U^{(T)}\}_{T \geq 0}$ are as in Theorem 6.2, and suppose $x(\cdot) \in AC[0, T]$ is optimal w.r.t. U at $(\xi, \eta, T) \in \mathbb{R}^n \times \mathbb{R}^n \times [0, \infty]$. Then $x(\cdot)$ is Lipschitz if and only if there exists $m > 0$ so that for all $t \in [0, T]$ and $h \in [0, T-t]$, we have

$$(6.2) \quad \frac{1}{h}U^{(h)}(x(t), x(t+h)) \leq m.$$

Proof. First, suppose $x(\cdot)$ is Lipschitz on $[0, T]$ of order l . Let $m = \sup\{L(\xi, v) : (\xi, v) \in \{\text{range of } x(\cdot)\} \times \ell B\}$. Then for $t \in [0, T]$ and

$h \in [0, T-t]$, we have

$$\begin{aligned} \frac{1}{h} U^{(h)}(x(t), x(t+h)) &\leq \frac{1}{h} V^{(h)}(x(t), x(t+h)) \quad (\text{by Theorem 3.2}) \\ &\leq \frac{1}{h} \int_t^{t+h} L(x(s), \dot{x}(s)) ds \quad (\text{by definition of } V) \\ &\leq m \quad (\text{by definition of } m). \end{aligned}$$

Hence (6.2) holds.

Now suppose (6.2) holds for some $m > 0$. It follows from (6.1) that for $t \in [0, T]$ and all small $h > 0$ with $t+h \leq T$ that

$$\text{epi } \Psi^{(h)}(x(t), \cdot) \cap mB \subseteq \text{epi } L(x(t), \cdot) + B.$$

From this and (6.2), an element $v \in \mathbb{R}^n$ exists for which

$$\left| v - \frac{x(t+h) - x(t)}{h} \right| \leq 1$$

and

$$(6.3) \quad L(x(t), v) \leq m+1.$$

Let $\rho = \sup \{|v| : \xi \in \text{range of } x(\cdot) \text{ and } L(\xi, v) \leq m+1\}$. It immediately follows from (6.3) that

$$\left| \frac{x(t+h) - x(t)}{h} \right| \leq \rho + 1$$

for all $t \in [0, T]$ and all small h , hence $x(\cdot)$ is Lipschitz on $[0, T]$. \square

Proof of Theorem 6.2. Let $(\xi, \eta, T) \in \mathbb{R}^n \times \mathbb{R}^n \times [0, \infty]$. It has already been shown in Theorem 3.2 that $U^{(t)}(\xi, \eta) \leq V^{(t)}(\xi, \eta)$. Suppose $x(\cdot)$ is Lipschitz on $[0, T]$ which is optimal w.r.t. U at (ξ, η, T) . We must show $U^{(t)}(\xi, \eta) \geq V^{(t)}(\xi, \eta)$.

Let $\varepsilon > 0$ be arbitrary. We fix some notation. Let $K \subseteq \mathbb{R}^n$ be compact whose interior contains the range of $x(\cdot)$. Also, let $\ell :=$ Lipschitz constant of $x(\cdot)$, $\lambda_0 :=$ Lipschitz constant of L on $K \times (\ell + \varepsilon)B \subseteq \mathbb{R}^n \times \mathbb{R}^n$, and m be such that (6.2) holds. From (6.1) there exists $h_0 > 0$ so that

$$(6.4) \quad \text{epi } \Psi^{(h)}(\gamma, \cdot) \cap mB \subseteq \text{epi } L(\gamma, \cdot) + \varepsilon B$$

for all $\gamma \in K$ and $0 < h < h_0$. Choose N large enough so that $\frac{T}{N} < \min \{h_0, \varepsilon\}$, and set $h = \frac{T}{N}$ and $t_j = jh$, $j = 0, 1, \dots, N$. From (6.2) and (6.4), there exist $v_j \in \mathbb{R}^n$, $j = 0, 1, \dots, N-1$, so that

$$(6.5) \quad \left| v_j - \frac{x(t_{j+1}) - x(t_j)}{h} \right| < \varepsilon$$

and

$$(6.6) \quad L(x(t_j), v_j) < \frac{1}{h} U^{(h)}(x(t_j), x(t_{j+1})) + \varepsilon.$$

Now let $y(\cdot) : [0, T] \rightarrow \mathbb{R}^n$ be the piecewise linear interpolation of $\{x(t_j)\}_{j=0}^N$. That is

$$(6.7) \quad y(t) = x(t_j) + \frac{x(t_{j+1}) - x(t_j)}{h} (t - t_j)$$

whenever $t_j \leq t \leq t_{j+1}$. We may assume N is large enough so that $y(\cdot)$ is contained in K . We have

$$\begin{aligned}
V^{(T)}(\xi, \eta) &\leq \int_0^T L(y(t), \dot{y}(t))dt \\
&= \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} L\left(y(t), \frac{x(t_{j+1}) - x(t_j)}{h}\right) dt \\
&\leq \sum_{j=0}^{N-1} \left\{ hL(x(t_j), v_j) + \lambda_0 h \left| v_j - \frac{x(t_{j+1}) - x(t_j)}{h} \right| \right. \\
&\quad \left. + \lambda_0 \int_{t_j}^{t_{j+1}} |y(t) - x(t_j)| dt \right\} \\
&\quad (\text{since } L \text{ is Lipschitz of order } \lambda_0 \text{ on } K \times (l + \varepsilon)B) \\
&\leq \sum_{j=0}^{N-1} \left\{ U^{(h)}(x(t_j), x(t_{j+1})) + h\varepsilon + \lambda_0 h\varepsilon + \lambda_0 h\varepsilon l \right\} \\
&\quad (\text{by (6.6), (6.5), (6.7) and } h \leq \varepsilon \text{ respectively}) \\
&= U^{(T)}(\xi, \eta) + \varepsilon(1 + \lambda_0 + \lambda_0 l)T \\
&\quad (\text{since } x(\cdot) \text{ is optimal w.r.t. } U).
\end{aligned}$$

Since ε was arbitrary, the conclusion of Theorem 6.2 is verified. \square

7. POINTWISE CONVERGENCE

The interior of a subset $A \subseteq \mathbb{R}^n$ is denoted by $\text{int } A$. In this final section we show that $L(\xi, v)$ is the pointwise limit of $\Phi^{(h)}(\xi, v)$ as $h \searrow 0$ provided $v \in \text{int dom } L(\xi, \cdot)$. This would follow immediately from Wets [29, Corollary 4] if $\Phi^{(h)}(\xi, \cdot)$ was convex. Convexity, however, is generally not present. The particular nature of value functions allows for the same conclusion to be drawn. In the case of finite-valued Lagrangians, $\text{dom } L(\xi, \cdot) = \mathbb{R}^n$ and hence the above applies. Moreover, we will see in Theorem 7.4 that Theorem 6.2 can be rephrased in terms of pointwise limits.

Theorem 7.1. *Suppose L satisfies (H1)–(H3), and let K be any compact subset of $\{(\xi, v) : \xi \in \mathbb{R}^n, v \in \text{int dom } L(\xi, \cdot)\}$. Then*

$$(7.1) \quad \frac{1}{h} V^{(h)}(\xi, \xi + hv) \rightarrow L(\xi, v) \quad \text{as } h \searrow 0$$

uniformly over $(\xi, v) \in K$.

Proof. Let $\xi \in \mathbb{R}^n$, and suppose $v \in \text{int dom } L(\xi, \cdot)$. Let $\varepsilon > 0$, which we assume is ≤ 1 . It is sufficient to show there exists $\delta_0 > 0$ and $h_0 > 0$ so that $(\eta, u) \in (\xi, v) + \delta_0 B$ and $0 < h \leq h_0$ imply $|\Phi^{(h)}(\eta, u) - L(\eta, u)| < \varepsilon$ (where $\Phi^{(h)}(\eta, u) := \frac{1}{h} V^{(h)}(\eta, \eta + hu)$).

Recall that convex functions are locally Lipschitz on the interior of their domain [21, Theorem 24.7]. This fact and (H3) combine to yield that L is locally Lipschitz near (ξ, v) . Choose $0 < \delta_1 \leq 1$ and $\lambda_1 \geq 1$ such that L

is Lipschitz of order λ_1 on $(\xi, v) + \delta_1 B$. Set $h_0 = \frac{\varepsilon}{\lambda_1(|v| + \delta_1)}$. Then for each $(\eta, u) \in (\xi, v) + \delta_1 B$ and $0 < h \leq h_0$, we have

$$(7.2) \quad \begin{aligned} \Phi^{(h)}(\eta, u) &= \frac{1}{h} V^{(h)}(\eta, \eta + hu) \leq \frac{1}{h} \int_0^h L(\eta + tu, u) dt \\ &\leq L(\eta, u) + h_0 \lambda_1 |u| \leq L(\eta, u) + \varepsilon \end{aligned}$$

Conversely, let m be an integer larger than $|v| + 1$ and $\sup\{|L(\eta, u)| + 1 : (\eta, u) \in (\xi, v) + \delta_1 B\}$. As a consequence of Theorem 3.1, we can shrink h_0 if necessary so that

$$(7.3) \quad \text{epi } \Phi^{(h)}(\eta, \cdot) \cap mB \subseteq \text{epi } L(\eta, \cdot) + \frac{\varepsilon \delta_1}{2\lambda_1} B$$

for each $\eta \in \xi + \delta_1 B$ and $0 < h \leq h_0$. If $(\eta, u) \in (\xi, v) + \frac{\delta_1}{2} B$ and $0 < h \leq h_0$, then (7.2) and the definition of m imply that $(u, \Phi^{(h)}(\eta, u)) \in \text{epi } \Phi^{(h)}(\eta, \cdot) \cap mB$. Consequently from (7.3), one can find $u' \in u + \frac{\varepsilon \delta_1}{2\lambda_1} B$ such that

$$(7.4) \quad \Phi^{(h)}(\eta, u) \geq L(\eta, u') - \frac{\varepsilon \delta_1}{2\lambda_1}.$$

Observe that $\frac{\varepsilon}{\lambda_1} \leq 1$, so that $(\eta, u') \in (\eta, u) + \frac{\delta_1}{2} B \subseteq (\xi, v) + \delta_1 B$, and recall that L is Lipschitz of order λ_1 on $(\xi, v) + \delta_1 B$. Consequently, from (7.4) we have

$$(7.5) \quad \Phi^{(h)}(\eta, u) \geq L(\eta, u) - \lambda_1 |u - u'| - \frac{\varepsilon \delta_1}{2\lambda_1} \geq L(\eta, u) - \varepsilon.$$

We have shown (7.2) and (7.5) hold for each $(\eta, u) \in (\eta, v) + \frac{\delta_1}{2} B$ and $0 < h \leq h_0$, which concludes the proof. \square

Of course Theorem 7.1 has content only when $\text{int dom } L(\xi, \cdot) \neq \emptyset$, whereas Theorems 3.1 and 3.2 do not require this. For the case when L is finite valued everywhere, it is now immediate that epi convergence in Theorem 3.1 can be replaced by uniform pointwise convergence. At the end of the section, we show the same conclusion is valid with the assumptions (H1)–(H3) replaced by local Lipschitz and strict convexity. But next we give versions of Theorems 3.1 and 6.2 for finite-valued Lagrangians with pointwise convergence.

Corollary 7.2. *Suppose L satisfies (H1)–(H3) and in addition is finite-valued. Then for each compact subset $K \subseteq \mathbb{R}^n \times \mathbb{R}^n$,*

$$(7.6) \quad \frac{1}{h} V^{(h)}(\xi, \xi + hv) \rightarrow L(\xi, v) \quad \text{as } h \searrow 0$$

uniformly over $(\xi, v) \in K$.

Theorem 7.3. *Theorem 6.2 remains valid if (6.1) is replaced by: for each compact $K \subseteq \mathbb{R}^n \times \mathbb{R}^n$,*

$$(7.7) \quad \Psi^{(h)}(\xi, v) \rightarrow L(\xi, v) \quad \text{as } h \searrow 0$$

uniformly over $(\xi, v) \in K$.

The proof consists of showing that (7.7) implies (6.1). In fact, the next proposition asserts that (7.7) and (6.1) are equivalent.

Proposition 7.4. *Let L be as in Theorem 6.2, and suppose $\{U^{(T)}\}_{T \geq 0}$ is a l.s.c. semigroup. Then (7.7) and (6.1) are equivalent.*

Proof. First, assume (7.7) holds. We show the validity of (6.1). Let $\varepsilon > 0$, $\rho > 0$ and $m > 0$. By (7.7) there exists $h_0 > 0$ so that

$$|\Psi^{(h)}(\xi, v) - L(\xi, v)| < \varepsilon$$

whenever $0 < h \leq h_0$ and $(\xi, v) \in \rho B \times mB$. It then follows immediately

$$\text{epi } \Psi^{(h)}(\xi, \cdot) \cap mB \subseteq \text{epi } L(\xi, \cdot) + \varepsilon B$$

and

$$\text{epi } L(\xi, \cdot) \cap mB \subseteq \text{epi } \Psi^{(h)}(\xi, \cdot) + \varepsilon B$$

whenever $0 < h \leq h_0$ and $\xi \in \rho B$. This is sufficient to conclude that (6.1) holds.

Now suppose (6.1) holds. Let $K \subseteq \mathbb{R}^n \times \mathbb{R}^n$ be compact and $\varepsilon > 0$. From Theorem 3.2 we have $U^{(T)}(\xi, \eta) \leq V^{(T)}(\xi, \eta)$ for all T, ξ and η . Hence it follows from Theorem 7.1 that

$$(7.8) \quad \Psi^{(h)}(\xi, v) \leq \Phi^{(h)}(\xi, v) \leq L(\xi, v) + \varepsilon$$

whenever $(\xi, v) \in K$ and h is small. To obtain the lower bound $L(\xi, v) - \varepsilon$ of $\Psi^{(h)}(\xi, v)$, one can proceed as in the proof of Theorem 7.1 commencing at (7.3). Simply substitute $\Psi^{(h)}$ for $\Phi^{(h)}$. The conclusion is that (7.7) holds. \square

We finish by showing that Corollary 7.2 is valid if the hypotheses (H1)–(H3) are replaced by L locally Lipschitz in (ξ, v) and strictly convex in v . To see this, we need the following theorem of Clarke and Vinter, see [8, Proposition 5.1].

Theorem 7.5 [8]. *Suppose $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^1$ is locally Lipschitz in (ξ, v) and strictly convex in v . Let $K \subseteq \mathbb{R}^n \times \mathbb{R}^n$ be compact. Then there exists $h_0 > 0$ and $r_0 > 0$ so that whenever $0 < h \leq h_0$ and $(\xi, \eta) \in K$, a minimizer x of problem (1.1) with $T = h$ and end points (ξ, η) is C^1 and satisfies $|\dot{x}(t)| \leq r_0$ for all $0 < t \leq h$.*

Theorem 7.6. *Suppose $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^1$ is locally Lipschitz in (ξ, v) and strictly convex in v . Then for each compact $K \subseteq \mathbb{R}^n \times \mathbb{R}^n$, we have*

$$\frac{1}{h} V^{(h)}(\xi, \xi + hv) \rightarrow L(\xi, v) \quad \text{as } h \rightarrow \infty$$

uniformly over $(\xi, v) \in K$.

Proof. Let h_0 and r_0 be chosen as in Theorem 7.5. Let $\rho = |K| + h_0 r_0$, and define

$$\tilde{L}(\xi, v) = \begin{cases} L(\xi, v) & \text{if } (\xi, v) \in \rho B \times r_0 B, \\ L\left(\rho \frac{\xi}{|\xi|}, v\right) & \text{if } |\xi| > \rho \text{ and } |v| \leq r_0, \\ +\infty & \text{if } |v| > r_0. \end{cases}$$

Let $\tilde{V}^{(T)}(\xi, \eta)$ be the value function associated with \tilde{L} , and $\tilde{\Phi}^{(h)}(\xi, v) = \frac{1}{h} \tilde{V}^{(h)}(\xi, \xi + hv)$. If $(\xi, v) \in K$ and $0 < h \leq h_0$, then a minimizer x

of (1.1) with $T = h$ and endpoints $(\xi, \xi + hv)$ with satisfies $(x(t), \dot{x}(t)) \in \rho B \times r_0 B$. Therefore for these values of h , ξ , and v , we have $\tilde{\Phi}^{(h)}(\xi, v) = \Phi^{(h)}(\xi, v)$. Note that \tilde{L} satisfies (H1)–(H3), and hence Theorem 7.1 asserts that $\tilde{\Phi}^{(h)}(\xi, v) \rightarrow \tilde{L}(\xi, v)$ as $h \searrow 0$ uniformly over $(\xi, v) \in K$. We have seen that $\Phi^{(h)} = \tilde{\Phi}^{(h)}$ on K for small h , and $L = \tilde{L}$ on K . Therefore $\Phi^{(h)} \rightarrow L$ as $h \searrow 0$ uniformly on K .

ACKNOWLEDGMENT

The author wishes to express his gratitude for the assistance provided by Professor R. T. Rockafellar.

REFERENCES

1. R. Bellman, *Dynamic programming*, Princeton Univ. Press, Princeton, N.J., 1957.
2. L. D. Berkowitz, *Optimal feedback controls*, SIAM J. Control Optim. **27** (1989), 991–1006.
3. L. Cesari, *Optimization—Theory and Applications*, Springer-Verlag, New York, 1983.
4. F. H. Clarke, *Extremal arcs and extended Hamiltonian systems*, Trans. Amer. Math. Soc. **231** (1977), 349–367.
5. ———, *Optimization and nonsmooth analysis*, Wiley Interscience, New York, 1983.
6. F. H. Clarke and R. B. Vinter, *Local optimality conditions and Lipschitzian solutions to the Hamilton-Jacobi equation*, SIAM J. Control Optim. **21** (1983), 856–870.
7. ———, *Regularity properties of solutions to the basic problem in the calculus of variations*, Trans. Amer. Math. Soc. **289** (1985), 73–98.
8. ———, *Existence and regularity in the small in the calculus of variations*, J. Differential Equations **59** (1985), 336–354.
9. M. G. Crandall, L. C. Evans and P. L. Lions, *Some properties of viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc. **282** (1984), 487–502.
10. M. G. Crandall, H. Ishii and P. L. Lions, *Uniqueness of viscosity solutions of Hamilton-Jacobi equations revisited*, J. Math. Soc. Japan **39** (1987), 581–595.
11. M. G. Crandall and P. L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc. **277** (1983), 1–42.
12. ———, *On existence and uniqueness of solutions of Hamilton-Jacobi equations*, Nonlinear Analysis, Theory, Methods and Applications **10** (1986), 353–370.
13. A. F. Filippov, *Classical solutions of differential equations with multivalued right-hand side*, SIAM J. Control Optim. **5** (1967), 609–621.
14. H. Frankowska, *Optimal trajectories associated to a solution of the contingent Hamilton-Jacobi equation*, Appl. Math. Opt. **19** (1989), 291–311.
15. R. L. Gonzales, *Sur l'existence d'une solution maximale de l'équation de Hamilton-Jacobi*, C. R. Acad. Sci. Paris **282** (1976), 1287–1290.
16. H. Ishii, *Uniqueness of unbounded viscosity solution of Hamilton-Jacobi equations*, Indiana Univ. Math. J. **33** (1984), 721–748.
17. P. L. Lions, *Generalized solutions of Hamilton-Jacobi equations*, Pitman, London, 1982.
18. P. L. Lions and P. E. Souganidis, *Differential games, optimal control, and directional derivatives of viscosity solutions of Bellman's and Isaac's equations*, SIAM J. Control Optim. **23** (1985), 566–583.
19. A. I. Panasiuk and V. I. Panasiuk, *On one equation resulting from a differential inclusion*, Mat. Zametki **27** (1980), 429–445. (Russian)
20. L. S. Pontryagin et al., *The mathematical theory of optimal processes*, Interscience, New York, 1962.

21. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N. J., 1970.
22. ———, *Optimal arcs and the minimum value function in problems of Lagrange*, Trans. Amer. Math. Soc. **180** (1973), 53–83.
23. ———, *Semigroups of convex bifunctions generated by Lagrange problems in the calculus of variations*, Math. Scand. **36** (1975), 137–158.
24. ———, *Existence theorems for general control problems of Bolza and Lagrange*, Adv. in Math. **15** (1975), 312–333.
25. ———, *Integral functionals, normal integrands, and measurable selections*, Nonlinear Operators and the Calculus of Variations, (L. Waelbroeck, Ed.), Lecture Notes in Math., vol. 543, Springer-Verlag, 1976, pp. 157–207.
26. E. Roxin, *On the generalized dynamical systems defined by contingent equation*, J. Differential Equations **1** (1965), 188–205.
27. R. B. Vinter and R. M. Lewis, *A necessary and sufficient condition for optimality of dynamic programming type, making no a priori assumptions on the controls*, SIAM J. Control Optim. **16** (1978), 571–583.
28. R. B. Vinter and P. R. Wolenski, *Hamilton-Jacobi theory for optimal control problems with data measurable in time*, SIAM J. Control Optim. (to appear).
29. R. J. B. Wets, *Convergence of convex functions, variational inequalities, and convex optimization problems*, Variational Inequalities and Complementarity Problems, (R. Cottle, F. Giannessi, and J. L. Lions, Eds.), Wiley, 1980, pp. 375–403.
30. P. R. Wolenski, *Semigroups of multifunctions and properties of the value function*, Thesis, Univ. of Washington, 1988.
31. ———, *The exponential formula for the reachable set of a Lipschitz differential inclusion*, SIAM J. Control Optim. (to appear).
32. ———, *A uniqueness theorem for differential inclusions*, J. Differential Equations (to appear).
33. V. Zeidan, *Sufficient conditions for the generalized problem of Bolza*, Trans. Amer. Math. Soc. **275** (1983), 561–586.
34. ———, *A modified Hamilton-Jacobi approach in the generalized problem of Bolza*, Appl. Math. Optim. **11** (1984), 97–109.

DEPARTMENT OF MATHEMATICS, LOUISIANA STATE UNIVERSITY, BATON ROUGE, LOUISIANA
 70803
E-mail address: wolenski@marais.math.lsu.edu