

DUALITY AND UNIQUENESS OF CONVEX SOLUTIONS TO STATIONARY HAMILTON-JACOBI EQUATIONS

RAFAL GOEBEL

ABSTRACT. Value functions for convex optimal control problems on infinite time intervals are studied in the framework of duality. Hamilton-Jacobi characterizations and the conjugacy of primal and dual value functions are of main interest. Close ties between the uniqueness of convex solutions to a Hamilton-Jacobi equation, the uniqueness of such solutions to a dual Hamilton-Jacobi equation, and the conjugacy of primal and dual value functions are displayed. Simultaneous approximation of primal and dual infinite horizon problems with a pair of dual problems on finite horizon, for which the value functions are conjugate, leads to sufficient conditions on the conjugacy of the infinite time horizon value functions. Consequently, uniqueness results for the Hamilton-Jacobi equation are established. Little regularity is assumed on the cost functions in the control problems, correspondingly, the Hamiltonians need not display any strict convexity and may have several saddle points.

1. INTRODUCTION

A key issue in dynamic programming is the description of the optimal value function as the unique solution to an associated Hamilton-Jacobi equation. When the control problem defining the value function is convex, duality theory suggests pairing it with a dual problem, possibly leading to a conjugacy relationship between the value function of the original problem and that of the dual. In the first part of this work, we explore the connections between these seemingly unrelated issues, for convex control problems on unbounded time intervals. We obtain the equivalence of the following statements:

- The value function of the original problem is the unique convex solution of the Hamilton-Jacobi equation.
- The value function of the dual problem is the unique convex solution of the dual Hamilton-Jacobi equation.
- The value functions of the original and the dual problems are conjugate to each other.
- The value functions of the original and dual problems can be simultaneously approximated by value functions of a pair of dual finite time horizon problems (with the latter value functions being conjugate to each other).

In the second part of the paper, we use the approximation technique of the last statement above to obtain sufficient conditions for all of the equivalent statements

Received by the editors April 3, 2003.

2000 *Mathematics Subject Classification*. Primary 49N15, 49L99; Secondary 49M29.

Key words and phrases. Stationary Hamilton-Jacobi equation, convex value function, optimal control, infinite horizon, dual problem, concave-convex Hamiltonian.

©2005 American Mathematical Society
Reverts to public domain 28 years from publication

to hold. In particular, this leads to results on uniqueness of convex solutions to the stationary Hamilton-Jacobi equation. The sufficient conditions turn out to be closely related to zero level sets of the value functions and to the growth of the Hamiltonian outside the set of its saddle points. We conclude the paper with applications of the developed theory to some important problems of control engineering.¹

Main objects under investigation in the paper are as follows. The value function $V : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is defined as the optimal value in a generalized problem of Bolza:

$$(1.1) \quad V(\xi) = \inf \left\{ \int_0^{+\infty} L(x(t), \dot{x}(t)) dt \mid x(0) = \xi \right\}.$$

The minimization is carried out over all locally absolutely continuous arcs $x : [0, +\infty) \mapsto \mathbb{R}^n$ (that is, $\dot{x}(\cdot)$ is integrable on bounded intervals), subject to the initial condition $x(0) = \xi$. Above and in what follows, $\overline{\mathbb{R}} = (-\infty, +\infty]$. The Lagrangian $L : \mathbb{R}^{2n} \rightarrow \overline{\mathbb{R}}$ is allowed to be nonsmooth and infinite-valued, and the key assumption on it is the full convexity: $L(x, v)$ is convex in (x, v) . It is well appreciated that such problems of Bolza can model control problems with explicit linear dynamics, convex running costs, and control constraints. For a general exposition see Clarke [6], for details in the convex setting see Rockafellar [19] or Goebel and Rockafellar [13]. A dual value function, defined by a problem of Bolza closely related to that of (1.1), will be described in Section 1.2.

In order for the value function to be well defined, we assume that $L \geq 0$, $L(0, 0) = 0$. This is a standard feature in problems arising from control engineering applications (like the linear-quadratic regulator and its extensions; see Section 4) and in “associated problems of Lagrange” used by Carlson, Haurie, and Leizarowitz [5] and in some of the references therein to study the existence of optimal solutions (in a weaker, for example overtaking, sense) for problems where the value function is never finite. Detailed technical growth assumptions on L , equivalent to those stated at the end of this section in terms of the corresponding Hamiltonian

$$(1.2) \quad H(x, y) = \sup_{v \in \mathbb{R}^n} \{y \cdot v - L(x, v)\},$$

are given in Section 2. A key feature of Hamiltonians corresponding to fully convex Lagrangians is that they are not only convex in y , but also concave in x .

As the value function in our setting is guaranteed to be convex, we will be interested in convex functions satisfying the following Hamilton-Jacobi equation:

$$(1.3) \quad H(x, -\partial f(x)) = 0 \quad \text{for all } x \in \mathbb{R}^n.$$

In this paper, ∂f denotes the subdifferential of f , in the sense of convex analysis:

$$\partial f(x) = \{y \mid f(x') \geq f(x) + y \cdot (x' - x) \text{ for all } x' \in \mathbb{R}^n\},$$

and the equation $H(x, -\partial f(x)) = 0$ is understood as $H(x, -y) = 0$ for all $y \in \partial f(x)$. The notion of a viscosity solution (see a detailed exposition by Bardi and Capuzzo-Dolcetta [1]) for continuous functions and the notion of a generalized solution of Clarke [6] for locally Lipschitz functions agree with the one just defined for the class of finite-valued convex functions (which are necessarily locally Lipschitz). This follows from the fact that for convex functions, the convex and Clarke subdifferentials

¹Some of the results of this paper were reported in the proceedings of the Louisiana Conference on Mathematical Control Theory 2003, [11].

agree, and from a result of Frankowska [7]. We restrict our attention to convex solutions of (1.3), but do not require that they be finite-valued. Issues of uniqueness of solutions to (1.3) are discussed in Section 1.1.

Detailed assumptions, in place throughout the paper, are as follows.

Assumption 1.1 (Hamiltonian assumptions). The function $H : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ is everywhere finite, such that $x \mapsto H(x, y)$ is concave for any fixed y , and $y \mapsto H(x, y)$ is convex for any fixed x . Moreover, there exists a constant α and a finite convex function ϕ with $\phi \geq 0$, $\phi(0) = 0$, such that

$$(1.4) \quad -\phi(x) - \alpha|x||y| \leq H(x, y) \leq \phi(y) + \alpha|y||x|$$

for all $(x, y) \in \mathbb{R}^{2n}$.

In particular, $H(0, 0) = 0$ and $(0, 0)$ is a saddle point of H : for any $x \in \mathbb{R}^n$, any $y \in \mathbb{R}^n$, $H(x, 0) \leq H(0, 0) \leq H(0, y)$. (These properties correspond to $L \geq 0$ and $L(0, 0) = 0$.) However, we do not require that the saddle point be unique. We add that our assumptions are closely related to those used by Rockafellar and Wolenski [25] in the development of duality and Hamilton-Jacobi theories for finite horizon convex control problems (detailed discussion of those assumptions can be found in [25], [8]; for special cases see also [13], [9]). Our H satisfies Assumption 1.1 if and only if it fits the framework of [25] and meets the normalizing conditions of $(0, 0)$ being its saddle point with $H(0, 0) = 0$.

1.1. Value function and the Hamilton-Jacobi equation. The value function (1.1) does provide us with a solution to (1.3); see Corollary 2.5. In general, one should not expect it to be the unique solution, even subject to some boundary conditions. This is well appreciated for the other generalized notions of solutions; see Bardi and Capuzzo-Dolcetta [1]. Let us argue that the assumption of convexity does not automatically lead to uniqueness.

Equation (1.3) may have several solutions, even in the classical sense, and subject to a boundary condition $f(x_0) = a_0$. For example, consider $H_1(x, y) = -\frac{1}{2}x^2 + \frac{1}{2}y^2$, corresponding to the Lagrangian $L_1(x, v) = \frac{1}{2}x^2 + \frac{1}{2}v^2$. There are four continuously differentiable solutions f satisfying $H_1(x, -\nabla f(x)) = 0$ and $f(0) = 0$. Only one of them, $f(x) = \frac{1}{2}x^2$, is convex; in fact, only this one is nonnegative. Both properties are expected from the value function corresponding to L , and indeed the value function is $V_1(\xi) = \frac{1}{2}\xi^2$. Note though that looking for solutions to $H(x, -M(x)) = 0$ among maximal monotone operators M (gradients, and more generally subdifferentials of convex functions are not just maximal monotone, but maximal cyclically monotone) does not exclude “bad” solutions. Consider $H(x, y) = -\frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2$ on \mathbb{R}^4 . The solution of our main interest – the value function – is as expected $V(\xi) = \frac{1}{2}\|\xi\|^2$. However, any rotation M of \mathbb{R}^2 by an angle in $[-\pi/2, \pi/2]$ is monotone and satisfies $H(x, -M(x)) = 0$ for all $x \in \mathbb{R}^2$. Such rotations are not related to gradients of any convex functions, unless the angle is 0.

This leads to the question: is the value function V the unique proper, lsc, and convex solution to (1.3), subject to $f \geq 0$, $f(0) = 0$. (A function $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is proper if it is finite somewhere.) In general, the answer is still negative. Consider $H_2(x, y) = xy$ on \mathbb{R}^2 , this indeed being a concave-convex Hamiltonian, corresponding to $L_2(x, v) = \delta_x(v)$. The value function $V(\xi) = 0$ is the unique classical solution, but there are three other solutions among proper, lsc, and convex

functions: δ_0 , $\delta_{(-\infty,0]}$, and $\delta_{[0,+\infty)}$. Here, and in what follows, δ_S denotes the indicator of the set S : $\delta_S(x) = 0$ if $x \in S$ and $\delta_S(x) = +\infty$ otherwise.

Even worse situations can appear for reasonable control problems. For $s \in \mathbb{R}$, let $\phi(s) = \max\{0, |s| - 1\}$, and consider $H_3(x, y) = -\phi(x) + \phi(y)$. There is a continuum of classical convex solutions to (1.3), and “even more” of nondifferentiable ones. Smoothing out the corners of H_3 does not change this situation. Detailed analysis of Hamiltonians with such large saddle sets was carried out by Goebel [10]; here we only add that the Lagrangian corresponding to H_3 is $L_3(x, v) = \phi(x) + |v| + \delta_{[-1,1]}(v)$.

Positive results on uniqueness of solutions to (1.3) cover only special classes of convex problems. The Riccati equation, describing quadratic solutions to the Hamilton-Jacobi equation for the linear-quadratic regulator, has a unique positive semidefinite solution under stabilizability and detectability assumptions (the Hamilton-Jacobi equation may even then possess other convex but not quadratic solutions; see the discussion in Section 4). More general problems with “separable” Hamiltonians $H(x, y) = y \cdot Ax - g(x) + h(y)$ were treated by Cannarsa and Da Prato [4], Iftode [15] (focusing, respectively, on viscosity solutions and classical solutions, and not permitting constraints), and Di Blasio [3] (allowing for hard, but not state-dependent, constraints). All of the mentioned works required some strict convexity assumptions, and neither took advantage of those available in the convex setting (and powerful) duality methods. The latter is the path we pursue here, and obtain strong uniqueness results for convex solutions. (These do not guarantee the uniqueness of viscosity solutions.) Following Da Lio [18], we add that many other uniqueness results developed for viscosity solutions rarely apply to the convex setting, as their growth conditions are not compatible with convexity.

1.2. Dual problems and conjugacy of value functions. The work of Rockafellar [19] suggested pairing a given control problem with a dual one, having exactly the same structure as the original. Before proceeding, we recall that for a proper, lsc, and convex function $k : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, its conjugate $k^* : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, defined by

$$k^*(b) = \sup_{a \in \mathbb{R}^n} \{b \cdot a - k(a)\},$$

is also proper, lsc, and convex, and the conjugate of k^* is in turn exactly k . Following Rockafellar [22], where dual control problems on infinite time horizons were introduced, we consider

$$(1.5) \quad W(\eta) = \inf \left\{ \int_0^{+\infty} M(y(t), \dot{y}(t)) dt \mid y(0) = \eta \right\}.$$

The dual Lagrangian $M : \mathbb{R}^{2n} \rightarrow [0, +\infty)$ is given as the convex conjugate of L with switched variables:

$$(1.6) \quad M(y, w) = L^*(w, y) = \sup_{(x,v) \in \mathbb{R}^{2n}} \{y \cdot v + w \cdot x - L(x, v)\}.$$

(Note that the Lagrangian dual to M is exactly L .) Equivalently, M can be constructed directly from the Hamiltonian H . Just as L corresponds to H through a formula reciprocal to (1.2),

$$(1.7) \quad L(x, v) = \sup_{y \in \mathbb{R}^n} \{v \cdot y - H(x, y)\},$$

the dual Lagrangian M corresponds to a Hamiltonian $\overline{H}(y, x) = -H(x, y)$ (note that \overline{H} is indeed a concave-convex function satisfying Assumption 1.1). That is,

$$(1.8) \quad M(y, w) = \sup_{x \in \mathbb{R}^n} \{w \cdot x - \overline{H}(y, x)\} = \sup_{x \in \mathbb{R}^n} \{w \cdot x + H(x, y)\}.$$

Similarly to L , the Lagrangian M is convex in both variables, nonnegative, with $M(0, 0) = 0$. The last two properties are an instance of the following fact, which we will often use in what follows: a proper, lsc, and convex function k satisfies $k \geq 0$ and $k(0) = 0$ if and only if $k^* \geq 0$ and $k^*(0) = 0$. Indeed, $k \geq 0$ and $k(0) = 0$ implies that $0 \in \partial k(0)$ which is equivalent to $0 \in \partial k^*(0)$, which is equivalent to saying that k^* attains its minimum at 0, while $f(0) = 0$ is equivalent to 0 being the minimum of $k^* -$ one has $k(0) = (k^*)^*(0) = \sup_y \{0 \cdot y - k^*(y)\} = -\inf_y k^*(y)$.

As $\overline{H}(y, -\partial g(y)) = 0$ is exactly $H(-\partial g(y), y) = 0$, the Hamilton-Jacobi equation for the dual problem, being a dual version of (1.3), is

$$H(-\partial g(y), y) = 0 \quad \text{for all } y \in \mathbb{R}^n.$$

In any optimization framework, dual problems are of particular interest if the optimal value of the original problem and that of the dual can be meaningfully related to each other. Rockafellar [22] showed, for strictly concave, strictly convex Hamiltonians (these conditions rule out hard constraints in the underlying control problems) that V and W are conjugate to each other, subject to a minus sign (the two chains of equalities below are equivalent):

$$(1.9) \quad \begin{aligned} W(\eta) &= \sup_{\xi \in \mathbb{R}^n} \{-\eta \cdot \xi - V(\xi)\} = V^*(-\eta), \\ V(\xi) &= \sup_{\eta \in \mathbb{R}^n} \{-\xi \cdot \eta - W(\eta)\} = W^*(-\xi). \end{aligned}$$

For an example of a similar duality relationship appearing without a direct reference to optimality, see Goebel et al. [14], where convex Lyapunov functions, and more generally, convex storage functions for linear differential inclusions, are discussed.

Hamiltonian H_1 from the previous section is strictly concave, strictly convex and thus, for the related control problems, (1.9) holds. Indeed, the dual Lagrangian in this case is the same as L_1 , and the dual value function is $W_1(\eta) = \frac{1}{2}\eta^2$. Now $\eta \mapsto W_1(-\eta) = \frac{1}{2}\eta^2$ is indeed conjugate to V_1 . For one-dimensional problems but with no strict convexity present, conjugacy-like formulas were obtained by Goebel [10]. These included a correction for a possible duality gap between $W(\eta)$ and $V^*(-\eta)$, measured in terms of the size of the saddle set of the Hamiltonian. Consequently, these conjugacy-like formulas were applicable to H_3 of the previous section – note that for this Hamiltonian, both value functions are 0 on $[-1, 1]$, and thus are not conjugate to each other. No reasonable duality relationship can be expected for Hamiltonians like H_2 . Both value functions are then equal to 0 everywhere (the original Lagrangian is $L(x, v) = \delta_x(v)$, the dual is $M(y, w) = \delta_{-y}(w)$), while the conjugate of the zero function is the indicator of 0.

1.3. Finite horizon problems. The study of value functions in the framework of duality, for finite time horizon convex control problems, was initiated by Rockafellar and Wolenski [25]. Their conjugacy result, adjusted for our current purposes, is as follows. Consider two proper, lsc, and convex functions $\phi, \psi : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, and define

two value functions $V_\phi, W_\psi : [0, +\infty) \times \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ by

$$(1.10) \quad \begin{aligned} V_\phi(\tau, \xi) &= \inf \left\{ \int_0^\tau L(x(t), \dot{x}(t)) dt + \phi(x(\tau)) \mid x(0) = \xi \right\}, \\ W_\psi(\tau, \eta) &= \inf \left\{ \int_0^\tau M(y(t), \dot{y}(t)) dt + \psi(-y(\tau)) \mid y(0) = \eta \right\}. \end{aligned}$$

Minimization is carried out over absolutely continuous arcs $x, y : [0, \tau] \rightarrow \mathbb{R}^n$. For any $\tau \geq 0$, the functions $V_\phi(\tau, \cdot)$, $W_\psi(\tau, \cdot)$ are proper, lsc, and convex. Moreover, when $\psi = \phi^*$, they are conjugate to each other (subject to a minus sign), that is,

$$(1.11) \quad \begin{aligned} W_{\phi^*}(\tau, \eta) &= \sup_{\xi \in \mathbb{R}^n} \{-\eta \cdot \xi - V_\phi(\tau, \xi)\}, \\ V_\phi(\tau, \xi) &= \sup_{\eta \in \mathbb{R}^n} \{-\xi \cdot \eta - W_{\phi^*}(\tau, \eta)\}. \end{aligned}$$

Compare this to (1.9). Vaguely speaking, letting $\tau \rightarrow +\infty$ in the equations above should yield the desired conjugacy of the infinite time horizon value functions. The catch is, that while choosing a function ϕ so that $V_\phi(\tau, \cdot)$ converges in some sense to V is not difficult, it is much less obvious (and not always true) that W_{ϕ^*} will then converge to the desired limit W .

As the functions in question may be infinite-valued, and since we need a notion of convergence with respect to which conjugacy is well behaved, we call upon the convergence of functions in the epigraphical sense (epi-convergence). A sequence of functions $k_i : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $i = 1, 2, \dots$, epi-converges to k if, for every point $x \in \mathbb{R}^n$,

$$(1.12) \quad \begin{aligned} (a) \quad &\liminf_{i \rightarrow \infty} k_i(x_i) \geq k(x) \quad \text{for every sequence } x_i \rightarrow x, \\ (b) \quad &\limsup_{i \rightarrow \infty} k_i(x_i) \leq k(x) \quad \text{for some sequence } x_i \rightarrow x. \end{aligned}$$

Consult Rockafellar and Wets [24], Chapter 7, for details. Conjugacy of convex functions is continuous with respect to epi-convergence, that is, for proper, lsc, and convex functions k_i , $i = 1, 2, \dots$, and k ,

$$(1.13) \quad \text{e-}\lim_{i \rightarrow \infty} k_i = k \quad \text{is equivalent to} \quad \text{e-}\lim_{i \rightarrow \infty} k_i^* = k^*.$$

We note that Cannarsa and Da Prato [4], Di Blasio [3], and Iftode [15] in their work on the stationary Hamilton-Jacobi equation did rely on finite-horizon approximates. The task we will face here is producing approximations that work simultaneously for the primal and the dual value function.

2. EQUIVALENCE

We are now ready to state the equivalence promised in the Introduction. Assumption 1.1 is still in force; in this section we also give an equivalent formulation in terms of the Lagrangian.

Theorem 2.1 (equivalence). *The following statements are equivalent:*

- (a) *The value function V is the unique proper, lsc, and convex function satisfying*

$$(2.1) \quad H(x, -\partial f(x)) = 0 \text{ for all } x \in \mathbb{R}^n, \quad f \geq 0, \quad f(0) = 0.$$

(b) *The dual value function W is the unique proper, lsc, and convex function satisfying*

$$(2.2) \quad H(-\partial g(y), y) = 0 \text{ for all } y \in \mathbb{R}^n, \quad g \geq 0, \quad g(0) = 0.$$

(c) *The conjugacy relationship $W(\eta) = V^*(-\eta)$ for all $\eta \in \mathbb{R}^n$ holds. Equivalently, for all $\xi \in \mathbb{R}^n$, $V(\xi) = W^*(-\xi)$.*

(d) *There exists a proper, lsc, and convex function ϕ for which*

$$(2.3) \quad \text{e-}\lim_{\tau \rightarrow +\infty} V_\phi(\tau, \cdot) = V, \quad \text{e-}\lim_{\tau \rightarrow +\infty} W_{\phi^*}(\tau, \cdot) = W.$$

We do not assert that any of the equivalent statements holds in the general convex setting; falseness of this was pointed out in the Introduction. Sufficient conditions will be obtained in the next section, with statement (d) being the vehicle.

To prove Theorem 2.1, we need several preliminary results. Each lemma below has its dual counterpart, that is, any statement involving L or V is also valid, respectively, for M or W (the Hamiltonian H would need to be replaced by \bar{H}). As already mentioned, the growth conditions we use are those of Rockafellar and Wolenski [25] specified to the case of $L \geq 0$, $L(0, 0) = 0$. Theorem 2.3 of [25] translated to the current setting states that the Lagrangian $L : \mathbb{R}^{2n} \rightarrow \bar{\mathbb{R}}$ corresponds through (1.2) (equivalently, through (1.7)) to a Hamiltonian H satisfying Assumption 1.1 if and only if L is a proper, lsc, and convex function, such that:

- (i) the set $F(x) = \{v \mid L(x, v) < \infty\}$ is nonempty for all x , and there exists a constant ρ such that $\text{dist}(0, F(x)) \leq \rho|x|$ for all x ;
- (ii) there exist $\alpha \in \mathbb{R}$ and a coercive, proper, nondecreasing function θ on $[0, \infty)$ with $\theta(0) = 0$ such that $L(x, v) \geq \theta(\max\{0, |v| - \alpha|x|\})$ for all (x, v) .

Moreover, the same statement can be made about the dual Lagrangian M , thanks to (1.8) and the symmetry in Assumption 1.1.

Lemma 2.2. *Consider a sequence of absolutely continuous arcs $x_i : [0, \tau_i] \mapsto \mathbb{R}^n$, where $\{\tau_i\}_{i=1}^\infty$ is an increasing sequence diverging to infinity. Assume that $\lim_{i \rightarrow \infty} x_i(0) = \xi$ and that there exists a constant K such that, for all $i = 1, 2, \dots$, $\int_0^{\tau_i} L(x_i(t), \dot{x}_i(t)) dt \leq K$. Then there exists a locally absolutely continuous arc $x : [0, +\infty) \rightarrow \mathbb{R}^n$ with $x(0) = \xi$ such that $\int_0^{+\infty} L(x(t), \dot{x}(t)) dt \leq K$.*

Proof. Fix a sufficiently large i_0 . As $L \geq 0$, for all $i \geq i_0$, we have $\delta_{\xi+\mathbb{B}}(x_i(0)) + \int_0^{\tau_i} L(x_i(t), \dot{x}_i(t)) dt \leq K$. Theorem 1 of [21] states that the necessarily convex set of all absolutely continuous arcs on $[0, \tau_i]$ satisfying this inequality is weakly compact. (That the assumptions required by this theorem are satisfied when $\delta_{\xi+\mathbb{B}}$ above is replaced by δ_ξ was demonstrated in the proof of Theorem 4.5 in [25]; the same arguments are valid in the current setting.) Consequently, one can select a subsequence of x_i 's, weakly converging on $[0, \tau_i]$ to some arc \bar{x}_i . Of course, $\bar{x}_i(0) = \xi$, and $\int_0^{\tau_i} L(\bar{x}_i(t), \dot{\bar{x}}_i(t)) dt \leq K$. Similar arguments applied to the interval $[0, \tau_{i+1}]$ show that from the previously chosen subsequence, we can select a further subsequence weakly converging on $[0, \tau_{i+1}]$ to \bar{x}_{i+1} . Note that \bar{x}_{i+1} agrees with \bar{x}_i on $[0, \tau_i]$ and satisfies a similar integral bound on $[0, \tau_{i+1}]$.

Now, the standard diagonalization argument leads us to a subsequence of arcs x_i , weakly converging on every bounded time interval to some arc x defined on $[0, \infty)$. This arc satisfies the conditions required by the lemma. □

This immediately leads to existence of optimal arcs, as stated below. We will also call upon the lemma in the study of epi-convergence of finite-horizon value

functions in Section 3. We note that various existence results for infinite horizon problems do exist (see Carlson et al. [5] and the references therein); they do not apply here due to our weak coercivity assumptions.

Corollary 2.3. *The value function V is proper, lsc, and convex. If $V(\xi) < +\infty$, then optimal arcs in (1.1) exist. Symmetrically, the value function W is proper, lsc, convex, and if $W(\eta) < +\infty$, then optimal arcs in (1.5) exist.*

The use of the established Hamilton-Jacobi theory for convex value functions for finite time horizon problems will be made possible through the Principle of Optimality: for any $\tau > 0$ and $\xi \in \mathbb{R}^n$, the value function V satisfies

$$(2.4) \quad V(\xi) = \inf \left\{ \int_0^\tau L(x(t), \dot{x}(t)) dt + V(x(\tau)) \mid x(0) = \xi \right\}.$$

This is standard, we do not include the proof.

Lemma 2.4. *For any proper, lsc, and convex function $f : \mathbb{R}^n \mapsto \overline{\mathbb{R}}$, the following statements are equivalent:*

- (a) $f(\xi) = \inf \left\{ \int_0^\tau L(x(t), \dot{x}(t)) dt + f(x(\tau)) \mid x(0) = \xi \right\} \quad \forall \xi \in \mathbb{R}^n, \tau \geq 0;$
- (b) $H(x, -\partial f(x)) = 0$ for all $x \in \mathbb{R}^n$.

Proof. Theorem 2.5 of [25], after an appropriate change of variables, states that the value function $U : [0, +\infty) \times \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ defined by

$$U(\tau, \xi) = \inf \left\{ \int_0^\tau L(x(t), \dot{x}(t)) dt + f(x(\tau)) \mid x(0) = \xi \right\}$$

satisfies, for all $\tau > 0$, $\xi \in \mathbb{R}^n$, the following nonstationary Hamilton-Jacobi equation: $\rho + H(x, -\eta) = 0$ for any $(\rho, \eta) \in \partial^g U(\tau, \xi)$. Here, $\partial^g U$ is the generalized subdifferential of U ; see [24], Chapter 6, for details. If (a) holds, then $U(\tau, \xi) = f(\xi)$ for all $\tau \geq 0$, all $\xi \in \mathbb{R}^n$, and $\partial^g U(\tau, \xi) = 0 \times \partial f(\xi)$. Consequently, we obtain (b).

Now assume (b). Consider $u : \mathbb{R} \times \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ defined as $u(t, x) = f(x)$ if $t \geq 0$, $u(t, x) = +\infty$ if $t < 0$. Then u is proper, lsc, meets the boundary condition $u(0, x) = f(x)$, and satisfies $\rho + H(x, -\eta) = 0$ for all $(\rho, \eta) \in \partial^g u(t, x) = 0 \times \partial f(x)$ if $t > 0$ while $\rho + H(x, -\eta) \leq 0$ for all $(\rho, \eta) \in \partial^g u(0, x) = (-\infty, 0] \times \partial f(x)$. The last equality follows from the fact that $u(t, x) = f(x) + \delta_{(-\infty, 0]}(t)$ and from the calculus rules of the generalized subdifferential (10.9 in [24]). Note that in fact $(t, x) \mapsto f(x)$ and $(t, x) \mapsto \delta_{(-\infty, 0]}(t)$ are convex functions of (t, x) . A uniqueness result of Galbraith [8] states that u must consequently be the value function defined by the right-hand side of the equation in (a). Thus, (b) implies (a). \square

Corollary 2.5 (Hamilton-Jacobi equation). *The value function V satisfies the equation $H(x, -\partial V(x)) = 0$ for all $x \in \mathbb{R}^n$. Dually, the value function W satisfies $H(-\partial W(y), y) = 0$ for all $y \in \mathbb{R}^n$.*

Lemma 2.6.

- (a) *Suppose that a proper, lsc, and convex function f solves (2.1). Then $f \geq V$. Dually, if a proper, lsc, and convex function g solves (2.2), then $g \geq W$.*
- (b) *A proper, lsc, and convex function f solves (2.1) if and only if g defined by $g(y) = f^*(-y)$ solves (2.2). Dually, a proper, lsc, and convex function g solves (2.2) if and only if f defined by $f(x) = g^*(-x)$ solves (2.1).*

Proof. We show (a) first. If $f(\xi) = +\infty$, then $f(\xi) \geq V(\xi)$. Suppose $f(\xi) < +\infty$. As f solves (2.1), Lemma 2.4 implies, in light of $f \geq 0$, that for any $\tau \geq 0$,

$$f(\xi) \geq \inf \left\{ \int_0^\tau L(x(t), \dot{x}(t)) dt \mid x(0) = \xi \right\}.$$

Pick a sequence $\tau_i \rightarrow \infty$, and let $x_i(\cdot)$ be an arc minimizing $\int_0^{\tau_i} L(x(t), \dot{x}(t)) dt$ subject to $x_i(0) = \xi$ (it exists by Theorem 5.2 of [25]). Lemma 2.2 can now be used to conclude the existence of an arc \bar{x} on $[0, +\infty)$ such that $x(0) = \xi$ and $\int_0^{+\infty} L(\bar{x}(t), \dot{\bar{x}}(t)) dt \leq f(\xi)$. By the definition of the value function, $V(\xi) \leq f(\xi)$.

We now turn to (b). Suppose f satisfies (2.1). The equation $H(x, -\partial f(x)) = 0$ for all $x \in \mathbb{R}^n$ means that $H(x, y) = 0$ for all (x, y) satisfying $-y \in \partial f(x)$. The inclusion is equivalent to $x \in \partial f^*(-y)$ which turns to $-x \in \partial g(y)$ with $g(y) = f^*(-y)$. Consequently, we have $H(x, y) = 0$ for all (x, y) with $-x \in \partial g(y)$ which is exactly $H(-\partial g(y), y) = 0$ for all $y \in \mathbb{R}^n$. It was explained in Section 1.2 that $f \geq 0, f(0) = 0$ is equivalent to $f^* \geq 0, f^*(0) = 0$; this implies that $g \geq 0, g(0) = 0$. Consequently, g satisfies (2.2). The arguments can be easily reversed. \square

We can now prove Theorem 2.1.

Proof.

- (a) implies (c): By Lemma 2.5, the value function W solves (2.2). Lemma 2.6(b) implies that $W^*(-\cdot)$ solves (2.1). But (a) states that V is the unique such function, and thus $V(\cdot) = W^*(-\cdot)$.
- (b) implies (c): Proof is symmetric to the one above.
- (c) implies (a): Suppose that f solves (2.1). Then by Lemma 2.6(a), $f \geq V$. On the other hand, by Lemma 2.6(b), the function g given by $g(y) = f^*(-y)$ solves (2.2). Using Lemma 2.6(a) again implies that $g \geq W$. The last inequality is equivalent to $g^* \leq W^*$, which, thanks to (c), means that $f \leq V$. Thus $f = V$, and so the value function is the unique solution to (2.1).
- (c) implies (a): Proof is symmetric to the one above.
- (d) implies (c): As $e\text{-}\lim_{\tau \rightarrow +\infty} V_\phi(\tau, \cdot) = V$, then by continuity of conjugacy with respect to epi-convergence as described in (1.13), $e\text{-}\lim_{\tau \rightarrow +\infty} V_\phi^*(\tau, \cdot) = V^*$. Here, for a fixed $\tau \geq 0$, $V_\phi^*(\tau, \cdot)$ denotes the convex conjugate of $V_\phi(\tau, \cdot)$. Conjugacy of finite time horizon value functions (1.11) implies that $V_\phi^*(\tau, \cdot) = W_{\phi^*}(\tau, -\cdot)$. Thus $e\text{-}\lim_{\tau \rightarrow +\infty} W_{\phi^*}(\tau, -\cdot) = V^*$. But by the second part of (d), the limit above is $W(-\cdot)$. Consequently, $V^*(\cdot) = W(-\cdot)$.
- (c) implies (d): Set $\phi = V$. By the definition of V_ϕ and the Principle of Optimality (2.4), $V_\phi(\tau, \cdot) = V$ for all $\tau > 0$. Moreover, by the assumption that $\phi^* = W(-\cdot)$, the same reasoning in the dual setting implies $W_{\phi^*}(\tau, \cdot) = W$. Statements in (d) now trivially follow. \square

3. FINITE-HORIZON APPROXIMATION AND SUFFICIENT CONDITIONS

This section presents sufficient conditions, under which the equivalent statements in Theorem 2.1 hold. In contrast to Rockafellar [22] and Goebel [10], we do not approach this through the saddle-point behavior of Hamiltonian trajectories, but rather by approximating the infinite-horizon problems by finite-horizon ones.

We begin with a necessary condition. For any proper, lsc, and convex function k , it is true that $k(a) + k^*(b) \geq a \cdot b$, for all a, b . In particular, if $k(a) = 0$ and

$k^*(b) = 0$, then $a \cdot b \leq 0$. Applying this to the value functions V and W yields the necessary condition below. The change in sign in the inequality is due to the extra minus sign in the conjugacy relationship between V and W .

Lemma 3.1 (necessary condition). *Suppose that the equivalent statements in Theorem 2.1 hold. Consider the sets*

$$(3.1) \quad C_0 = \{x \in \mathbb{R}^n \mid V(x) = 0\}, \quad D_0 = \{y \in \mathbb{R}^n \mid W(y) = 0\}.$$

Then

$$(3.2) \quad c \cdot d \geq 0 \quad \text{for all } c \in C_0, d \in D_0.$$

In particular, if $C \times D$ is the set of saddle points of H , then $c \cdot d \geq 0$ for all $c \in C$, $d \in D$.

The last statement above calls for an explanation. A saddle point of H is any point $(c, d) \in \mathbb{R}^{2n}$ such that $H(x, d) \leq H(c, d) \leq H(c, y)$ for any $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$. The set of all saddle points is necessarily a product of two closed and convex sets; see Rockafellar [20]. Moreover, as $(0, 0)$ is a saddle point and $H(0, 0) = 0$, we have $H(c, d) = 0$ for any $(c, d) \in C \times D$. Also, if (c, d) is a saddle point of H , so are $(c, 0)$, $(0, d)$, and thus, by the conjugacy formulas of Section 1.2, $L(c, 0) = 0$ and $M(d, 0) = 0$. (For a further discussion, in terms of the subdifferentials of L , M , and H , see Rockafellar [22].) Consequently, $V(c) = 0$, $W(d) = 0$.

Lemma 3.2. *Suppose that the sets C_0 , D_0 defined in (3.1) are compact. For any ξ , any optimal trajectory x for $V(\xi)$, $\lim_{t \rightarrow \infty} \text{dist}_{C_0}(x(t)) = 0$. Similarly, for any η , any optimal trajectory y for $W(\eta)$, $\lim_{t \rightarrow \infty} \text{dist}_{D_0}(y(t)) = 0$.*

Proof. Let x be an optimal trajectory for $V(\xi)$. Thus $V(\xi) = \int_0^{+\infty} L(x(t), \dot{x}(t)) dt$ and also, by Principle of Optimality, $V(x(s)) = \int_s^{+\infty} L(x(t), \dot{x}(t)) dt$ for all $s \geq 0$. As $V(\xi)$ is finite (else x cannot be optimal), the latter integral converges to 0 as $s \rightarrow \infty$, and thus $V(x(s)) \searrow 0$.

As C_0 is compact and nonempty ($V(0) = 0$), any other level set $C_r = \{x \mid V(x) \leq r\}$ of the convex function V must be compact (3.23 in [24]) and consequently the maximum of $\text{dist}_{C_0}(x)$ over $x \in C_r$ must converge to 0 as $r \searrow 0$. This, and $V(x(s)) \searrow 0$, yields the desired conclusion. \square

Lemma 3.3 (finite horizon approximation). *Let C be any closed and convex set such that if $V(\xi) < +\infty$, then $\lim_{t \rightarrow \infty} \text{dist}_C(x(t)) = 0$ for some optimal trajectory $x(\cdot)$ for $V(\xi)$. Let ϕ be any proper, lsc, and convex function with $\phi \geq 0$ and such that if a sequence x_i satisfies $\lim_{i \rightarrow \infty} \text{dist}_C(x_i) = 0$, then $\lim_{i \rightarrow \infty} \phi(x_i) = 0$. Then the finite horizon value function V_ϕ in (1.10) satisfies $e\text{-}\lim_{\tau \rightarrow \infty} V_\phi(\tau, \cdot) = V$.*

Proof. We start by showing the first inequality in the definition of epi-convergence (1.12). Fix $\xi \in \mathbb{R}^n$, and pick any sequences $\tau_i \rightarrow \infty$, $\xi_i \rightarrow \xi$. Denote $\liminf_i V_\phi(\tau_i, \xi_i)$ by K . If $K = +\infty$, the inequality holds. In the other case, passing to a subsequence if necessary, we have $\lim_i V_\phi(\tau_i, \xi_i) = K$ for some real number K , and in particular, for an $\epsilon > 0$, $V_\phi(\tau_i, \xi_i) < K + \epsilon$ for all large enough i . Optimal arcs x_i in the definition (1.10) of $V_\phi(\tau_i, \xi_i)$ exist, and as $\phi \geq 0$, $\int_0^{\tau_i} L(x_i(t), \dot{x}_i(t)) dt < K + \epsilon$. Lemma 2.2 implies now that there exists an arc x_ϵ on $[0, +\infty)$ with $x_\epsilon(0) = \xi$ and $\int_0^{+\infty} L(x_\epsilon(t), \dot{x}_\epsilon(t)) dt \leq K + \epsilon$. Consequently, $V(\xi) \leq K + \epsilon$. The desired inequality easily follows. Note that so far we only used the nonnegativity of ϕ .

We now turn to the second inequality in (1.12). Fix $\xi \in \mathbb{R}^n$, and pick any sequence $\tau_i \rightarrow \infty$. If $V(\xi) = +\infty$, the inequality trivially holds. In the other case, let an arc x on $[0, +\infty)$ be optimal in the definition (1.1) of $V(\xi)$, with $\text{dist}_C(x(\tau_i))$ converging to 0. By the assumed property of ϕ , $\phi(x(\tau_i))$ also converges to 0. Pick $\epsilon > 0$. For all large enough i , by the definition of V_ϕ , we have

$$V(\xi) + \epsilon \geq \int_0^{\tau_i} L(x_i(t), \dot{x}(t)) dt + \phi(x(\tau_i)) \geq V_\phi(\tau_i, \xi).$$

Thus $\limsup_i V_\phi(\tau_i, \xi) \leq V(\xi)$, and the constant sequence $\xi_i = \xi$ fulfills the necessary for epi-convergence inequality. \square

We remark that the set C satisfying the assumptions of Lemma 3.3 may be smaller than C_0 , the zero level set of V defined in (3.1) (in Lemma 3.2 we concluded that a compact C_0 always satisfies these assumptions). For example, consider a Lagrangian $L(x, v) = (x + v)^2$. For any initial point $\xi \in \mathbb{R}^n$, the optimal trajectory is $x_\xi(t) = \xi e^{-t}$. Consequently, the set $C = \{0\}$ satisfies the conditions of Lemma 3.3 (in fact, any closed and convex set containing the origin does), while $C_0 = \mathbb{R}^n$. For a Lagrangian $L(x, v) = \delta_{[-1,1]}(v)$, while again $C_0 = \mathbb{R}^n$, any closed and convex set C in \mathbb{R}^n has the property that for any $\xi \in \mathbb{R}^n$ and any $c \in C$, there exists an optimal trajectory for $V(\xi)$ converging (in finite time) to c ; in particular, any such C fits the conditions of Lemma 3.3. Related discussion, and conditions guaranteeing that all trajectories of zero cost converge to the origin, can be found in Leizarowitz [17]. For our purposes, we will need compact sets C with the discussed property, and corresponding compact sets D for the dual Lagrangian M . (Such compact sets certainly exist for the two examples just described, and for the examples, the conclusions of Theorem 3.4 hold.)

Theorem 3.4 (sufficient conditions). *Assume that:*

- (a) *there exists a compact and convex set C such that, for any $\xi \in \mathbb{R}^n$, if $V(\xi) < +\infty$, then there exists an optimal trajectory $x(\cdot)$ for $V(\xi)$ satisfying $\lim_{t \rightarrow \infty} \text{dist}_C(x(t)) = 0$;*
- (b) *there exists a compact and convex set D such that, for any $\eta \in \mathbb{R}^n$, if $W(\eta) < +\infty$, then there exists an optimal trajectory $y(\cdot)$ for $W(\eta)$ satisfying $\lim_{t \rightarrow \infty} \text{dist}_D(y(t)) = 0$;*
- (c) *for all $c \in C$, $d \in D$, $c \cdot d \geq 0$.*

Then there exists a proper, lsc, and convex function for which (2.3) holds, and consequently, all of the equivalent statements in Theorem 2.1 hold.

Proof. In light of Lemma 3.3, it is sufficient to demonstrate the existence of a proper, lsc, and convex function ϕ , with $\phi \geq 0$, $\phi(0) = 0$, such that:

- if a sequence x_i satisfies $\lim_{i \rightarrow \infty} \text{dist}_C(x_i) = 0$, then $\lim_{i \rightarrow \infty} \phi(x_i) = 0$,
- if a sequence y_i satisfies $\lim_{i \rightarrow \infty} \text{dist}_D(y_i) = 0$, then $\lim_{i \rightarrow \infty} \phi^*(-y_i) = 0$.

As convex functions are continuous in the interior of their effective domains, it is sufficient that ϕ be nonnegative, zero on C , and finite on a neighborhood of C , while ϕ^* be zero on $-D$ and finite on a neighborhood of that set. Pick a constant K such that $C \cup D \subset \text{int } K\mathbb{B}$, with \mathbb{B} being the closed unit ball in \mathbb{R}^n . Define

$$\phi(x) = \sigma_{-D}(x) + \delta_{K\mathbb{B}}(x).$$

Here, σ_{-D} denotes the support function of $-D$, defined as $\sigma_{-D}(x) = \sup\{x \cdot d \mid d \in -D\}$, and $\delta_{K\mathbb{B}}$ is the indicator of $K\mathbb{B}$. By compactness of D , σ_{-D} is finite on \mathbb{R}^n ,

while by choice of K , $\delta_{K\mathbb{B}}$ is continuous on a neighborhood of C . For any $x \in C$, $\sigma_{-D}(c) = 0$, since by assumption $c \cdot d \leq 0$ for all $d \in -D$, while $0 \in D$. Obviously, for such x , $\delta_{K\mathbb{B}}(x) = 0$. Thus ϕ has the desired properties with respect to C .

Now consider ϕ^* , which turns out to be the inf-convolution of conjugates of the functions involved in the definition of ϕ (the conjugate of the support function is the indicator, and vice-versa, 11.23 in [24]):

$$\phi^*(y) = \inf_{z \in \mathbb{R}^n} \{ \delta_{-D}(y - z) + K\|z\| \}.$$

Clearly, $\phi^*(y) = 0$ for all $y \in -D$, moreover, ϕ^* inherits continuity (on the whole space) from that of the norm; see 9.11 in [24]. Consequently, ϕ^* has the requested properties with respect to $-D$. This finishes the proof. \square

Corollary 3.5. *Suppose that the sets C_0, D_0 defined in (3.1) are compact. Then (3.2) is a necessary and sufficient condition for all of the equivalent statements in Theorem 2.1 to hold.*

Recall that Lemma 3.1 concluded that if $C \times D$ is the set of saddle points of H , then $C \cdot D \geq 0$ is necessary for the equivalent statements of Theorem 2.1 to be true. We point out that it is not sufficient even if C and D are compact. In fact, even the uniqueness of a saddle point (corresponding to $C = D = \{0\}$) is not sufficient. Consider a Hamiltonian given by $H(x, y) = x(y + 1) + .5y^2$ for $x < 0$, $H(x, y) = xy + .5y^2$ for $x \geq 0$. Then $(0, 0)$ is the unique saddle point of H , while there are two convex functions f_1, f_2 satisfying (2.1): for $x \leq 0$, $\nabla f_1(x) = \nabla f_2(x) = x - \sqrt{x^2 - x}$, for $x > 0$, $\nabla f_1(x) = 0$, $\nabla f_2(x) = .5x$. Note that here $H(x, 0) = 0$ for all $x \geq 0$, excluding such “lack of growth”, as described in (a) of the corollary below, will lead to sufficiency.

Corollary 3.6. *Each of the following conditions is sufficient for all of the equivalent statements in Theorem 2.1 to hold:*

- (a) *There exist compact sets C, D with $C \cdot D \geq 0$ such that $H(x, 0) < 0 < H(0, y)$ for all $x \notin C$, all $y \notin D$.*
- (b) *On some neighborhood of $(0, 0)$, the Hamiltonian H is strictly convex in x , strictly concave in y .*
- (c) *The Lagrangian L is continuous at $(0, 0)$ and $L(x, v) > 0$ when $x \neq 0$.*
- (d) *The Lagrangian L is differentiable at $(0, 0)$ and there exists K such that $L(x, v) = 0$ implies $\|x\| \leq K$.*

Proof. As $H(x, 0) = -\inf_v L(x, v)$, the first inequality in (a) implies that $L(x, v) > 0$ for all v , all $x \notin C$. Thus $V(\xi) > 0$ for all $\xi \notin C$, which leads to $C_0 \subset C$ being compact. The second inequality in (a), through symmetric arguments, leads to $D_0 \subset D$ being compact. As $C \cdot D \geq 0$, (3.2) is in place, and the conclusion holds. Condition (b) is a special case of (a), with $C = D = \{0\}$: if (b) holds, $H(x, 0) < 0 < H(0, y)$ for all $x \neq 0$, all $y \neq 0$.

If (c) holds, clearly $C_0 = \{0\}$. As $(0, 0)$ is in the interior of the set where L is finite, $\partial L(0, 0)$ is a compact set. But $\partial L(0, 0)$ is exactly the set of all (y, w) with $L(y, w) = 0$; see 11.8 in [24]. Consequently, $L(y, w) > 0$ for all y outside some compact set, and D_0 is compact. As $C_0 \cdot D_0 = 0$, the conclusion follows.

Condition (d) is (almost) dual to (c). Differentiability of L at $(0, 0)$ (then necessarily $\nabla L(0, 0) = (0, 0)$ and $M(y, w) = 0$ only if $(y, w) = (0, 0)$) implies that $D_0 = \{0\}$, while C_0 is contained in a ball of radius K . \square

Lagrangians or Hamiltonians corresponding to optimal control problems often do not meet the conditions of Corollary 3.6. However, a combination of growth and observability or detectability properties can be used to verify the sufficient conditions of Theorem 3.4; see our discussion in Section 4.

We conclude this section with some consequences of duality of value functions. The proposition below was shown for strictly convex, strictly concave Hamiltonians in Rockafellar [22]; for planar Hamiltonians a more general result, not requiring duality, was given in Goebel [10].

A Hamiltonian trajectory on $[0, +\infty)$ is a pair of locally absolutely continuous arcs $x(\cdot), y(\cdot) : [0, +\infty) \rightarrow \mathbb{R}^n$ such that

$$(3.3) \quad -\dot{y}(t) \in \tilde{\partial}_x H(x(t), y(t)), \quad \dot{x}(t) \in \partial_y H(x(t), y(t)), \quad \text{for a. a. } t \in [0, +\infty).$$

Above, $\tilde{\partial}_x H(x, y) = -\partial_x(-H)(x, y)$ is the subdifferential (in the concave sense) of the concave function $H(\cdot, y)$, while $\partial_y H(x, y)$ is the subdifferential of the convex function $H(x, \cdot)$. For finite horizon problems, Theorem 2.4 of [25] describes how subdifferentials of the value function evolve from those of the terminal cost according to (3.3). This result was used in the analysis of regularity of the value functions in [9]; we give a corresponding application of Proposition 3.7 in Section 4.2.

Proposition 3.7. *Suppose that the equivalent statements in Theorem 2.1 hold. Then $\eta \in -\partial V(\xi)$ (which by duality is equivalent to $\xi \in -\partial W(\eta)$) if and only if there exists a Hamiltonian trajectory $(x(\cdot), y(\cdot))$ on $[0, +\infty)$ such that*

$$(3.4) \quad (x(0), y(0)) = (\xi, \eta) \quad \text{and} \quad \lim_{t \rightarrow +\infty} x(t) \cdot y(t) \leq 0.$$

Proof. By the very definition of M in (1.6), for any arcs $x(\cdot), y(\cdot)$ one has

$$(3.5) \quad \int_0^T L(x(t), \dot{x}(t)) dt + \int_0^T M(y(t), \dot{y}(t)) dt \geq x(T) \cdot y(T) - x(0) \cdot y(0),$$

and the inequality turns into an equation if and only if $(x(\cdot), y(\cdot))$ is a Hamiltonian trajectory on $[0, T]$ (see equation (8.3) in [19]). In particular, if $x(\cdot)$ and $y(\cdot)$ form a Hamiltonian trajectory on $[0, +\infty)$ and (3.4) holds, then

$$(3.6) \quad -\xi \cdot \eta \leq V(\xi) + W(\eta) \leq \lim_{t \rightarrow \infty} x(t) \cdot y(t) - \xi \cdot \eta \leq -\xi \cdot \eta,$$

where the first inequality holds by a general property of convex functions ($k(a) + k^*(b) \geq a \cdot b$), the second by (3.5) and the definitions of $V(\xi)$ and $W(\eta)$, and the last one by (3.4). Thus (3.6) turns into an equation, and this immediately leads to $\eta \in -\partial V(\xi)$ (for any convex function, $k(a) + k^*(b) = a \cdot b$ if and only if $b \in \partial k(a)$). Note also that: (i) $\lim_{t \rightarrow +\infty} x(t) \cdot y(t) = 0$, (ii) $x(\cdot)$ yields the minimum in the definition of $V(\xi)$, while $y(\cdot)$ yields the minimum in the definition of $W(\eta)$.

Now suppose $\eta \in -\partial V(\xi)$, equivalently $\xi \in -\partial W(\eta)$. Since both subdifferentials are nonempty, both $V(\xi)$ and $W(\eta)$ are finite. Let $x(\cdot)$ and $y(\cdot)$ be optimal arcs in the definitions of $V(\xi)$ and $W(\eta)$. We first argue that $x(\cdot)$ and $y(\cdot)$ form a Hamiltonian trajectory. This can be seen by writing $V(\xi)$ as a value function for a finite horizon problem as in (2.4), similarly for $W(\eta)$. By duality of V and W , the finite horizon problems fit the framework of [25]. That $x(\cdot)$ and $y(\cdot)$ form a Hamiltonian trajectory (on each subinterval $[0, \tau]$ of $[0, +\infty)$) was concluded in the proof of Theorem 2.4 in [25]. Without relying on $(x(\cdot), y(\cdot))$ being a Hamiltonian trajectory, we have, by (3.5) and $\eta \in -\partial V(\xi)$, that $-\xi \cdot \eta = V(\xi) + W(\eta) \geq \lim_{t \rightarrow +\infty} x(t) \cdot y(t) - \xi \cdot \eta$ and thus the desired inequality in (3.4) holds. \square

The first part of the proof above is similar to the proof of Proposition 2.2 in [22]. In the second part, in contrast to [22] we could not rely on any a priori convergence properties of optimal arcs. Open-loop optimality conditions, as stated below, are an immediate consequence of the arguments just given. By characterizing the adjoint arcs as optimal for the dual problem, they shed new light on the existing results, for example by Benveniste and Scheinkman [2] (where dual problems were not discussed); see also Carlson et al. [5]. Here we do not pursue the topic of optimality conditions without duality and only through an example in Section 4.2 illustrate applications of the theory to closed-loop (feedback) optimality conditions.

Corollary 3.8 (optimality conditions). *Suppose that the equivalent statements in Theorem 2.1 hold.*

- (a) *If $(x(\cdot), y(\cdot))$ is a Hamiltonian trajectory on $[0, +\infty)$ such that (3.4) holds, then $x(\cdot)$ is an optimal arc in the definition (1.1) of $V(\xi)$ and $y(\cdot)$ is an optimal arc in the definition (1.5) of $W(\eta)$.*
- (b) *If $x(\cdot)$ is an optimal arc in the definition (1.1) of $V(\xi)$, then for any $\eta \in -\partial V(\xi)$ there exists an arc $y(\cdot)$ such that $y(0) = \eta$, $(x(\cdot), y(\cdot))$ is a Hamiltonian trajectory on $[0, +\infty)$, and $\lim_{t \rightarrow +\infty} x(t) \cdot y(t) = 0$.*
- (b') *If $y(\cdot)$ is an optimal arc in the definition (1.5) of $W(\eta)$, then for any $\xi \in -\partial W(\eta)$ there exists an arc $x(\cdot)$ such that $x(0) = \xi$, $(x(\cdot), y(\cdot))$ is a Hamiltonian trajectory on $[0, +\infty)$, and $\lim_{t \rightarrow +\infty} x(t) \cdot y(t) = 0$.*

4. APPLICATIONS TO OPTIMAL CONTROL

We now briefly illustrate how our results apply to problems in the classical optimal control format. Details on how control problems can be modeled in the Bolza format are in Clarke [6], conditions on the control problem guaranteeing that the resulting Bolza problem fits our framework can be found in Goebel and Rockafellar [13] and Goebel [9], and for constructions of dual control problems see Rockafellar [23]. Consult Kwakernaak and Sivan [16] for the linear systems material we use below.

4.1. Linear-quadratic regulator. For a pair of symmetric and positive definite matrices Q and R , consider the following problem:

$$(4.1) \quad \min \frac{1}{2} \int_0^{+\infty} z(t) \cdot Qz(t) + u(t) \cdot Ru(t) dt$$

s.t. $\dot{x}(t) = Ax(t) + Bu(t)$, $z(t) = Cx(t)$, $x(0) = \xi$.

If (A, B) is stabilizable and (A, C) is detectable, the value function for (4.1) is given by $V(\xi) = \frac{1}{2}\xi \cdot P\xi$, with P being the unique positive semidefinite and symmetric solution to the Riccati equation $A^T P + PA + C^T Q C - P B R^{-1} B^T P = 0$. However, even under such assumptions, this value function need not be the unique solution to the Hamilton-Jacobi equation (2.1), with the Hamiltonian associated to (4.1) being

$$(4.2) \quad H(x, y) = y \cdot Ax - \frac{1}{2}x \cdot C^T Q C x + \frac{1}{2}y B R^{-1} B^T y.$$

Indeed, consider (4.1) with $Q = R = 1$, $A = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $C = \begin{bmatrix} 0 & 1 \end{bmatrix}$. Then $H(x, y) = -x_1 y_1 - \frac{1}{2}x_2^2 + \frac{1}{2}y_2^2$, and concentrating on separable solutions

$f(x) = f_1(x_1) + f_2(x_2)$ yields a unique convex f_2 but four different f_1 's (only one of them is quadratic) for which (2.1) holds; see our discussion in Section 1.1.

Application of Theorem 3.4 gives conditions under which the solution to the Hamilton-Jacobi equation with the Hamiltonian (4.2) is unique, without requiring that it be quadratic. First we describe the control problem dual to (4.1):

$$(4.3) \quad \begin{aligned} \min & \frac{1}{2} \int_0^{+\infty} w(t) \cdot R^{-1}w(t) + v(t) \cdot Q^{-1}v(t) dt \\ \text{s.t. } & \dot{y}(t) = -A^T y(t) - C^T v(t), \quad w(t) = B^T x(t), \quad y(0) = \eta. \end{aligned}$$

Detectability of (A, C) implies that optimal trajectories in (4.1) (in fact all trajectories with a finite cost) converge to 0. This also holds for the dual (4.3) if $(-A^T, B^T)$ is detectable, equivalently, if $(-A, B)$ is stabilizable. Theorem 3.4 yields:

Corollary 4.1. *Suppose that (A, C) is detectable and $(-A, B)$ is stabilizable. Then the equivalent statements in Theorem 2.1 hold for the pair of problems (4.1), (4.3). In particular the value function for the control problem (4.1) is the unique proper, lsc, and convex solution to (2.1) with the Hamiltonian (4.2).*

Similar arguments, and the same conclusion (with an updated Hamiltonian) is valid if a control constraint $u(t) \in U$ is introduced in (4.1), with U a convex set with $0 \in \text{int } U$. Then, the convex function $r(w) = \sup_{u \in U} \{w \cdot u - \frac{1}{2}u \cdot Ru\}$ replaces $\frac{1}{2}w \cdot R^{-1}w$ in (4.3). As $0 \in \text{int } U$, r has good growth properties which guarantee, together with detectability of $(-A^T, B^T)$, that optimal trajectories approach 0.

4.2. Stabilization of linear systems with saturating actuators. We conclude by outlining an application of our results to stabilization of the most common nonlinear systems in control engineering, a linear systems with saturating actuators:

$$(4.4) \quad \dot{x}(t) = Ax(t) + B\sigma(u(t)).$$

The question is whether (4.4) can be stabilized with continuous feedback on the set of all states that can be driven to 0 with an open-loop control. Details and a discussion can be found in Goebel [12].

The properties of the saturation function $\sigma : \mathbb{R}^k \rightarrow \mathbb{R}^k$ are almost always such that $\sigma = \nabla s$ for some convex function which is nonnegative, strictly convex on some neighborhood of 0, and $s(0) = 0$. For example, consider $\sigma(u) = \text{sign}(u) \min\{|u|, 1\}$ (such saturation effectively corresponds to a constraint $|u| \leq 1$), or $\sigma(u) = \arctan u$.

Consider the following optimal control problem (with linear dynamics!):

$$(4.5) \quad \text{minimize } \int_0^\infty \frac{1}{2}x(t) \cdot Qx(t) + s^*(w(t)) dt \quad \text{s.t. } \begin{cases} \dot{x}(t) = Ax(t) + Bw(t), \\ x(0) = \xi \end{cases}$$

and let V be the corresponding value function. If Q is symmetric and positive definite and (A, B) is controllable, then optimal trajectories for V , as well as those for the associated dual problem, converge to 0. Thus, by Theorem 3.4, the equivalent conditions of Theorem 2.1 hold. One can then show that V is continuously differentiable on $\text{dom } V$, the set of points where it is finite (which is open). This is done either via Proposition 3.7 and certain monotonicity properties of the Hamiltonian trajectories, or by showing that the dual value function is strictly convex and invoking conjugacy. Then, standard Hamilton-Jacobi theory arguments (here, $H(x, y) = y \cdot Ax - \frac{1}{2}x \cdot Qx + s(B^T y)$) show that the solutions to

$$\dot{x}(t) = Ax(t) + B\nabla s(-B^T \nabla V(x(t)))$$

are optimal for (4.5) (and in particular, the system above is asymptotically stable on $\text{dom } V$). But as $\nabla s = \sigma$, setting $u(t) = -B^T \nabla V(x(t))$ in (4.4) renders the system asymptotically stable, furthermore, V is a Lyapunov function verifying this. This generalizes a result by Sontag et al. [26], and more importantly, shows relevance of convex optimal control in stabilization of often encountered nonlinear systems.

REFERENCES

- [1] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Birkhauser, 1997. MR1484411 (99e:49001)
- [2] L.M. Benveniste and J.A. Scheinkman. Duality theory for dynamic optimization models of economics: the continuous time case. *J. Economic Theory*, 27:1–19, 1982. MR0662928 (84j:90022)
- [3] G. Di Blasio. Optimal control with infinite horizon for distributed parameter systems with constrained controls. *SIAM J. Control Optim.*, 29(4):909–925, 1991. MR1111667 (92c:49014)
- [4] P. Cannarsa and G. Da Prato. Nonlinear optimal control with infinite horizon for distributed parameter systems and stationary Hamilton-Jacobi equations. *SIAM J. Control Optim.*, 27(4):861–875, 1989. MR1001924 (90j:49019)
- [5] D.A. Carlson, A. Haurie, and A. Leizarowitz. *Infinite Horizon Optimal Control: Deterministic and Stochastic Systems*. Springer-Verlag, 1991.
- [6] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, 1983. MR0709590 (85m:49002)
- [7] H. Frankowska. Optimal trajectories associated with a solution of the contingent Hamilton-Jacobi equation. *Appl. Math. Optim.*, 19:291–311, 1989. MR0974188 (90b:49043)
- [8] G.N. Galbraith. Extended Hamilton-Jacobi characterization of value functions in optimal control. *SIAM J. Control Optim.*, 39(1):281–305, 2001. MR1780920 (2001j:49042)
- [9] R. Goebel. Convex optimal control problems with smooth Hamiltonians. *SIAM J. Control Optim.* in press.
- [10] R. Goebel. Planar generalized Hamiltonian systems with large saddle sets. *J. Nonlinear Convex Anal.*, 3(3):365–380, 2002. MR1947105 (2003k:49041)
- [11] R. Goebel. Stationary Hamilton-Jacobi equations for convex control problems – uniqueness and duality of solutions. In *Optimal Control, Stabilization, and Nonsmooth Analysis, Lecture Notes in Control and Information Sciences*, vol. 301, Springer, Berlin, 2004, pp. 313–322. MR2079692
- [12] R. Goebel. Stabilizing a linear systems with saturation through optimal control. In *Proceedings of the 43rd IEEE Conference on Decision and Control, Bahamas*, 2004.
- [13] R. Goebel and R.T. Rockafellar. Generalized conjugacy in Hamilton-Jacobi theory for fully convex Lagrangians. *Journal of Convex Analysis*, 9(1):463–473, 2002. MR1970567 (2004b:49063)
- [14] R. Goebel, A. Teel, T. Hu, and Z. Lin. Dissipativity for dual linear differential inclusions through conjugate storage functions. In *Proceedings of the 43rd IEEE Conference on Decision and Control, Bahamas*, 2004.
- [15] V. Iftode. Variational solutions of stationary Hamilton-Jacobi equations. *Italian J. Pure Appl. Math.*, 5:117–123, 1999. MR1736005 (2001d:49046)
- [16] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems*. Wiley-Interscience, 1972. MR0406607 (53:10394)
- [17] A. Leizarowitz. Existence of overtaking optimal trajectories for problems with convex integrands. *Mathematics of Operations Research*, 10(3):450–461, 1985. MR0798390 (87f:49002)
- [18] F. Da Lio. On the Bellman equation for infinite horizon problems with unbounded cost functional. *Appl. Math. Optim.*, 41:171–197, 2000. MR1731417 (2001i:49050)
- [19] R.T. Rockafellar. Conjugate convex functions in optimal control and the calculus of variations. *J. Math. Anal. Appl.*, 32:174–222, 1970. MR0266020 (42:929)
- [20] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970. MR0274683 (43:445)
- [21] R.T. Rockafellar. Existence and duality theorems for convex problems of Bolza. *Trans. Amer. Math. Soc.*, 159:1–40, 1971. MR0282283 (43:7995)
- [22] R.T. Rockafellar. Saddle points of Hamiltonian systems in convex problems of Lagrange. *Journal of Optimization Theory and Applications*, 12(4), 1973. MR0358516 (50:10977)

- [23] R.T. Rockafellar. Hamiltonian trajectories and duality in the optimal control of linear systems with convex costs. *SIAM J. Control Optim.*, 27(5):1007–1025, 1989. MR1009335 (90i:49017)
- [24] R.T. Rockafellar and R. J-B Wets. *Variational Analysis*. Springer, 1998. MR1491362 (98m:49001)
- [25] R.T. Rockafellar and P.R. Wolenski. Convexity in Hamilton-Jacobi theory, 1: Dynamics and duality. *SIAM J. Control Optim.*, 39(5):1323–1350, 2000. MR1825581 (2002b:49053)
- [26] E.D. Sontag, H.J. Sussmann, and Y.D. Yang. A general result on the stabilization of linear systems using bounded controls. *IEEE Trans. Automat. Control*, 39(12):2411–2424, 1994. MR1337566 (96g:93060)

CENTER FOR CONTROL ENGINEERING AND COMPUTATION, UNIVERSITY OF CALIFORNIA, SANTA BARBARA, CALIFORNIA 93106

Current address: 3518 NE 42 St., Seattle, Washington 98105

E-mail address: rafal@ece.ucsb.edu