Math. Control Signals Systems (1993) 6: 180–193

Mathematics of Control, Signals, and Systems

Dynamic Programming for Free-Time Problems with Endpoint Constraints*

J. Kotsiopoulos† and R. B. Vinter†

Abstract. If we are able to find a local verification function associated with an admissible trajectory $x(\cdot)$, then $x(\cdot)$ is a local minimizer. It is of interest therefore to know when such local verification functions exist. In this paper it is shown that the existence of a local verification function is necessary for $x(\cdot)$ to be a local minimizer, under a normality hypothesis. The novelty of these results is that they treat problems with a general endpoint constraint and where the endtime is a choice variable. Here the value function of the original problem does not serve as a local verification function; instead it must be constructed from some derived problem. The data are allowed to be measurable in the time variable, and the normality hypothesis is expressed in terms of recent free-endtime necessary conditions of optimality for problems with measurable time dependence.

Key words. Optimal control, Dynamic programming, Free endtime, Hamilton–Jacobi theory.

1. Introduction

We study the problem

(P) Minimize $g(T, x(T))$ over points $T \in [a, b]$ and $x(\cdot) \in AC([a, T]; \mathbb{R}^n)$ which satisfy the constraints

$$\dot{x}(t) \in F(t, x(t)) \qquad \text{a.e.} \quad t \in [a, T],$$

$$x(a) = x_0 \quad \text{and} \quad (T, x(T)) \in C,$$

for which the data are: an interval $[a, b] \subset \mathbb{R}$, a set $C \subset \mathbb{R}^{1+n}$, a function $g\colon [a, b] \times \mathbb{R}^n \to \mathbb{R}$, a multifunction $F\colon [a, b] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, and a point $x_0 \in \mathbb{R}^n$. We use the terminology "admissible trajectory" to mean an element $(T, x(\cdot))$, where T is a point in $[a, b]$ and $x(\cdot)$ is an arc in $AC([0, T]; \mathbb{R}^n)$ which satisfy the constraints of the above problem.

Our aim is to give necessary and sufficient conditions for local optimality, expressed in terms of a solution to the Hamilton–Jacobi inequality

$$\partial\varphi/\partial t + \operatorname*{Min}_{e \in F} (\partial\varphi/\partial x)e \geq 0. \tag{1.1}$$

* Date received: April 4, 1991. Date revised: August 25, 1992.

† Department of Electrical and Electronic Engineering and Centre for Process Systems Engineering, Imperial College of Science, Technology, and Medicine, London SW7 2BT, England.

Necessity will be proved under a normality hypothesis which we discuss shortly.

There is a common pattern to many of the proof techniques previously employed to derive optimality conditions of this sort (see, e.g., [B2], [FR], [F], [L], and [VW]). First we show that if we have a suitable solution to the Hamilton–Jacobi equation (call it a verification function) and a nominal state trajectory which satisfies a certain "compatibility" condition, then that arc must be a minimizer. This is the "sufficiency" part. Second, we show that if $x(\cdot)$ is a minimizer, then the value function V with the domain a subset of $[a, b] \times \mathbb{R}^n$, defined by

$$V(t, \xi) := \operatorname{Inf}\{g(T, x(T)) \colon \dot{x} \in F \text{ on } [a, T], x(t) = \xi \text{ and } (T, x(T)) \in C\}, \quad (1.2)$$

is a verification function associated with $x(\cdot)$. This is the necessity part.

This approach breaks down for the class of problems we consider, because of the presence of an endpoint constraint. Consider for example a fixed-time problem where the endpoint constraint set C takes the form

$$C = \{b\} \times (\mathbb{R} \times \{\xi\}) \quad (1.3)$$

for some point $\xi \in \mathbb{R}^{n-1}$. (This is the endpoint constraint we would need to consider in connection with a fixed endpoint integral cost problem, reformulated as a terminal cost problem.) Here the value function given by (1.2) is not a suitable candidate for a verification function. For one thing, the domain of V (points (t, ξ) for which the problem associated with $V(t, \xi)$ has admissible trajectories) is difficult to characterize, and may even have an empty interior. For another, V may have very poor analytic properties. It may for example be discontinuous. The major task then is to find a verification function φ to associate with a given minimizer for constrained problems, when we cannot choose V itself. This is accomplished as in [C] and [CV] by taking φ to be the value function for a related *unconstrained* problem (the "derived problem" we call it); in view of the absence of right endpoint constraints, the value function for the derived problem will have a suitable domain and regularity properties. A feature of verification functions for constrained problems is that they are not unique; indeed, various choices of derived problem can be made and they may generate distinct verification functions. The respects in which the results of this paper improve on those of [C] and [CV], which also treat problems with endpoint constraints, are that we now allow the terminal time T to be a choice variable (whereas previously it was fixed), and that we consider a differential inclusion which may depend measurably on the time variable (instead of continuously, as previously required.)

There is already substantial literature on the Hamilton–Jacobi equation and the value function for free-endtime problems. Earlier work focuses largely on interpreting solutions (appropriately defined) to the Hamilton–Jacobi equation as the value function for the original problem (P), in circumstances when the domain of the value function (regarded as a subset of x-space or (t, x)-space) is specified beforehand. Viscosity solution techniques have played a major role in establishing uniqueness of solutions to the Hamilton–Jacobi equation (see, e.g., [L], [B1], or [EJ]). Previous research, however, has little bearing on the main issues raised in this paper, which treats problems where the goal is to substitute in place of the value function a verification function on a larger, more convenient, domain, which still serves to provide optimality conditions. In this context there may be many such verification functions, and questions of uniqueness have little relevance. Interest centres instead on the interplay between the existence of verification functions and normality-type hypotheses, expressing in some way the nondegeneracy of first-order necessary conditions of optimality for free-endtime problems with data measurable in the time variable.

Normality is a sufficient condition for the existence of a verification function associated with a given minimizer. It may be replaced by a weaker hypothesis, "calmness", a condition on the sensitivity of the minimum cost to data perturbations, which is necessary and sufficient for the existence of a verification function (see [CV] and [K]). Normality can be regarded as a sufficient condition for calmness, which is easier to verify than calmness itself, for certain cases of interest.

"Solutions" to the Hamilton–Jacobi inequality (1.1) are defined in terms of the lower Dini derivative: given a function ψ: $[a, b] \times \mathbb{R}^n \to \mathbb{R}$, and points $e \in [a, b] \times \mathbb{R}^n$ and $y \in \{(y^0, y^1) \in \mathbb{R}^{1+n}: y^0 > 0\}$ the *lower Dini derivative* of ψ at e in the direction y, written $d^-\psi(e; y)$, is the point

$$d^-\psi(e; y) := \liminf_{\substack{y' \to y \\ h \downarrow 0}} h^{-1}\{\psi(e + hy') - \psi(e)\}.$$

Specifically, φ: $[a, b] \times \mathbb{R}^n \to \mathbb{R}$ is a *lower Dini solution to the Hamilton–Jacobi inequality* if a set $J \subseteq [a, b)$ of zero measure exists such that

$$\operatorname*{Min}_{e \in F(t,x)} \; d^-\varphi((t, x); (1, e)) \geq 0 \qquad \text{for all} \quad (t, x) \in (\{a, b)\backslash J) \times \mathbb{R}^n.$$

As previously mentioned we cannot expect verification functions associated with a minimizer to be unique for the constrained problems of interest here. This is a reason why we have not adopted "viscosity-type" definitions of solutions to the Hamilton–Jacobi equation, definitions which in other circumstances have the merit of yielding a unique verification function.

If we aim to establish the existence of verification functions for constrained problems, it is inevitable that some kind of nondegeneracy hypotheses be introduced. (See [C] which supplies examples of degenerate problems with endpoint constraints, even in the fixed-endtime case, where no verification function exists.) The nature of these nondegeneracy conditions is typically that the optimal control problem is *normal*, in the sense that first-order necessary conditions of optimality hold in a form where the cost multiplier is nonzero. The concept of "normality" employed in this paper relates to recent first-order optimality conditions (expressed in the form of a Hamiltonian inclusion) for free-time problems with data measurable in time.

2. Hypotheses

Let us recall that an admissible trajectory is an element $(T, x(\cdot))$ in which $T \in [a, b]$ and $x(\cdot) \in AC([a, T]; \mathbb{R}^n)$ satisfy the constraints of problem (P). It is often convenient to regard a state trajectory $(T, x(\cdot))$ as defining a function with the domain all of $[a, b]$; for this purpose we extend $x(\cdot)$ by constant extrapolation, i.e., we set $x(t) = T$ for $t > T$. Some concept of "neighbourhood" of an admissible state trajectory is required. Given an admissible trajectory $(\bar{T}, \bar{x}(\cdot))$ and $\varepsilon > 0$ we define

$$N_\varepsilon(\bar{T}, \bar{x}(\cdot)) := \{(t, \xi) \in [a, b] \times \mathbb{R}^n : t < \bar{T} + \varepsilon, |\xi - \bar{x}(t)| < \varepsilon\}.$$

An element $(T, x(\cdot) \in AC([a, T]; \mathbb{R}^n))$ is said to be in the ε neighborhood about $(\bar{T}, \bar{x}(\cdot))$ if $T \in (\bar{T} - \varepsilon, \bar{T} + \varepsilon) \cap [a, b]$, $|x(T) - \bar{x}(\bar{T})| < \varepsilon$, and graph $\{x(\cdot)\} \subset N_\varepsilon(\bar{T}, \bar{x}(\cdot))$. Notice that, in defining $N_\varepsilon(T, x(\cdot))$, we have employed the extrapolation of $x(\cdot)$ to $[a, b]$, described above, to incorporate points (t, ξ) with $t > T$ in $N_\varepsilon(T, x(\cdot))$.

Of primary interest in this paper is the characterization of "local" minimizers for (P): an admissible trajectory $(T, x(\cdot))$ is a *local minimizer* for (P) if $\varepsilon > 0$ exists such that it achieves the minimum value of $g(T, x(T))$ as $(T, x(\cdot))$ ranges over all admissible trajectories in the ε neighbourhood about $(\bar{T}, \bar{x}(\cdot))$. In these circumstances we also say "$(\bar{T}, \bar{x}(\cdot))$ is a minimizer over the ε neighbourhood of $(\bar{T}, \bar{x}(\cdot))$".

It is appropriate then that we should impose hypotheses of a local nature: they relate to a nominal trajectory $(\bar{T}, \bar{x}(\cdot))$ and a parameter $\varepsilon > 0$ defining a neighbourhood of $\bar{T}, \bar{x}(\cdot)$.

(H1) $F(\cdot, \xi)$ is measurable for all $\xi \in \mathbb{R}^n$.
(H2) For all $(t, \xi) \in N_\varepsilon(\bar{T}, \bar{x}(\cdot))$, $F(t, \xi)$ is compact and convex.
(H3) $\lambda(\cdot) \in L^1$ exists such that

$$\text{dist}_H(F(t, \xi), F(t, \xi')) \le \lambda(t)|\xi - \xi'| \qquad \text{for all} \quad (t, \xi), (t, \xi') \in N_\varepsilon(\bar{T}, \bar{x}(\cdot)).$$

(H4) $c > 0$ exists such that

$$F(t, \xi) \subset cB \qquad \text{for all} \quad (t, \xi) \in N_\varepsilon(\bar{T}, \bar{x}(\cdot)).$$

(H5) g is Lipschitz continuous on $\{(t, \xi): |t - \bar{T}| < \varepsilon, |\xi - \bar{x}(\bar{T})| < \varepsilon\}$.

3. Normality

In [CLV] first-order necessary conditions for a state trajectory to be a local minimizer for (P) are derived. To allow for the facts that (P) is a free-time problem and the F was permitted to be merely measurable in its time dependence, these conditions were expressed in terms of the convexified "set of essential values" of the maximized Hamiltonian.

Definition 3.1. Let $y: [a, b] \to \mathbb{R}$ be an essentially bounded function and take $T \in (a, b)$. Then the convexified set of essential values of y at T, written $\text{co ess}_{s \to T}\, y(s)$, is

$$\operatorname*{co\,ess}_{s \to T} y(s) := [\alpha^-, \alpha^+],$$

where

$$\alpha^- := \lim_{\varepsilon \downarrow 0} \operatorname*{ess\,inf}_{s \in [T-\varepsilon, T+\varepsilon]} y(s) \quad \text{and} \quad \alpha^+ := \lim_{\varepsilon \downarrow 0} \operatorname*{ess\,sup}_{s \in [T-\varepsilon, T+\varepsilon]} y(s).$$

Evidently in the case y is continuous at T we have

$$\operatorname*{co\,ess}_{s \to T} y(s) = \{y(T)\},$$

so taking the convexified essential value is a generalized notion of evaluating a function at a point.

Define the Hamiltonian function

$$H(t, x, p) := \sup_{e \in F(t,x)} p.e.$$

Let $(\bar{T}, \bar{x}(\cdot))$ be a local minimizer for (P) (with respect to a neighbourhood $N_\varepsilon(\bar{T}, \bar{x}(\cdot))$ about $(\bar{T}, \bar{x}(\cdot))$). Assume that hypotheses (H1)–(H5) are satisfied with respect to $N_\varepsilon(\bar{T}, \bar{x}(\cdot))$. It is shown in [CLV] that a function $p(\cdot) \in AC([a, \bar{T}], \mathbb{R}^n)$, a point $h \in \mathbb{R}$, and a parameter $\lambda \geq 0$ exist such that $\|p(\cdot)\|_{L_\infty} + \lambda > 0$ and the following conditions hold:

$$(-\dot{p}(t), \dot{x}(t)) \in \partial H(t, x(t), p(t)) \qquad \text{a.e.} \quad t \in [a, T],$$

$$h \in \operatorname*{co\,ess}_{s \to \bar{T}} H(s, \bar{x}(\bar{T}), p(T)),$$

and

$$(h, -p(\bar{T})) \in N_C(\bar{T}, \bar{x}(\bar{T})) + \lambda \partial g(\bar{T}\bar{z}(\bar{T})).$$

In these relationships $\partial H(t, x, p)$ denotes the generalized gradient of $H(t, \cdot, \cdot)$, ∂g is the generalized gradient of $g(\cdot, \cdot)$, and $N_C(\bar{z})$ is the Clarke normal core of C at $\bar{z}$. (See [C] for the definition of generalized gradients and normal cones.)

If it is possible to take $\lambda = 0$ these conditions are degenerate, since they then make no reference to the cost function, and we cannot expect them to convey useful information about minimizers. Of course in these circumstances $p(\cdot) \neq 0$ since we must have $\|p(\cdot)\|_{L_\infty} + \lambda \neq 0$. The normality hypothesis is that degeneracy of this type does not occur:

Definition 3.2. We say problem (P) is *normal* with reference to an admissible trajectory $(\bar{T}, \bar{x}(\cdot))$ if the only function $p(\cdot) \in AC([a, \bar{T}]; \mathbb{R}^n)$ which satisfies

$$(-\dot{p}(t), \dot{\bar{x}}(t)) \in \partial H(t, \bar{x}(t), p(t)) \qquad \text{a.e.} \quad t \in [a, \bar{T}]$$

and

$$(h, -p(\bar{T}) \in N_C(\bar{T}, \bar{x}(\bar{T}))$$

for some

$$h \in \operatorname*{co\,ess}_{s \to \bar{T}} H(s, \bar{x}(\bar{T}), p(\bar{T}))$$

is the function $p(\cdot) \equiv 0$.

4. The Main Results

We need to introduce the concept of a "local" verification function.

Definition 4.1. Given $\varepsilon > 0$ and an admissible trajectory $(\bar{T}, \bar{x}(\cdot))$, a verification function on the ε neighbourhood of $(\bar{T}, \bar{x}(\cdot))$ is a function φ: $[a, b] \times \mathbb{R}^n \to \mathbb{R}$ which satisfies the following conditions:

(a) φ is locally Lipschitz continuous.
(b) $\varphi(t, \xi) \leq g(t, \xi)$ whenever

$$(t, \xi) \in N_\varepsilon(\bar{T}, \bar{x}(\cdot)) \cap \{(t', \xi') \colon |t' - \bar{T}| < \varepsilon \text{ and } |\xi' - \bar{x}(\bar{T})| < \varepsilon\}.$$

(c) A null set $J \subset [a, b)$ exists such that

$$\min_{e \in F(t, \xi)} \{d^- \varphi((t, \xi); (1, e))\} \geq 0 \quad \text{for all} \quad (t, \xi) \in (([a, b)\backslash J) \times \mathbb{R}^n) \cap (N_\varepsilon(\bar{T}, \bar{x}(\cdot)).$$

(The "lower Dini derivative" $d^- \varphi$ referred to in (c) was defined in the introduction.)

The necessary and sufficient conditions for local optimality which we provide are the following:

Theorem 4.2. *Let $(\bar{T}, \bar{x}(\cdot))$ be an admissible trajectory for* (P) *such that $a < \bar{T} < b$. Assume that, for some $\varepsilon > 0$, hypotheses* (H1)–(H5) *are satisfied with reference to the ε neighbourhood of $(\bar{T}, \bar{x}(\cdot))$. We have*:

(i) *Suppose that a verification function φ on the ε neighbourhood of $(\bar{T}, \bar{x}(\cdot))$ exists such that*

$$\varphi(a, x) = g(\bar{T}, \bar{x}(\bar{T})). \tag{4.1}$$

Then $(\bar{T}, \bar{x}(\cdot))$ is a minimizer on $N_\varepsilon(\bar{T}, \bar{x}(\cdot))$.

(ii) *Suppose that $(\bar{T}, \bar{x}(\cdot))$ is a minimizer on $N_\varepsilon(\bar{T}, \bar{x}(\cdot))$, and problem* (P) *is normal with reference to $(\bar{T}, \bar{x}(\cdot))$. Then $\varepsilon' \in (0, \varepsilon)$ and a verification function φ on $N_{\varepsilon'}(\bar{T}, \bar{x}(\cdot))$ exist such that* (4.1) *is satisfied.*

If $(\bar{T}, \bar{x}(\cdot))$ is an admissible state trajectory and φ is a verification function on $N_\varepsilon(\bar{T}, \bar{x}(\cdot))$, for some $\varepsilon > 0$, which satisfies (4.1), *we say φ is a local verification function for $(\bar{T}, \bar{x}(\cdot))$.*

5. The Value Function for an Unconstrained Problem

An important step in the proof of Theorem 4.1, which is given in Section 6, is to examine properties of the value function of a simplified version of problem (P), from which the endpoont constraint "$(T, x(T)) \in C$" has been dropped. With this in mind we insert at this point a short section on the unconstrained problem:

(Q) Minimize $g(T, x(T)$ over points $T \in [a, b]$ and $x(\cdot) \in AC([a, T]; \mathbb{R}^n)$ which satisfy the constraints

$$\dot{x}(t) \in F(t, x(t)) \qquad \text{a.e.} \quad t \in [a, T]$$

and

$$x(a) = x_0.$$

The data is as before (we have of course taken $C = \mathbb{R} \times \mathbb{R}^n$), but we impose stronger "global" versions of hypotheses (H1)–(H5):

(C1) $F(\cdot, \xi)$ is measurable for all $\xi \in \mathbb{R}^n$.
(C2) For all (t, ξ), $F(t, \xi)$ is compact and convex.
(C3) $\lambda(\cdot) \in L^1$ exists such that

$$\operatorname{dist}_H(F(t, \xi), F(t, \xi')) \le \lambda(t)|\xi - \xi'| \qquad \text{for all} \quad (t, \xi), (t, \xi') \in [a, b] \times \mathbb{R}^n.$$

(C4) $c > 0$ exists such that

$$F(t, \xi) \subset cB \qquad \text{for all} \quad (t, \xi) \in [a, b] \times \mathbb{R}^n.$$

(C5) g is Lipschitz continuous on $\mathbb{R} \times \mathbb{R}^n$.

Denote by $V(\cdot, \cdot)$ the value function for (Q). In view of the fact that $C = \mathbb{R} \times \mathbb{R}^n$, it now takes the form

$$V(t, \xi) = \inf\{g(\overline{T}, \bar{x}(\overline{T}))\},$$

where the infimum is taken over admissible trajectories for a variant of (Q) in which t and ξ replace the initial time a and initial value of the state trajectories x_0, respectively. Under the hypotheses, the domain of $V(t, \xi)$ is all of $[a, b] \times \mathbb{R}^n$.

We show that $V(\cdot, \cdot)$ satisfies global versions of the conditions defining local verification functions (see Definition 4.1). Before doing so we need to assemble a few properties of reachable sets.

Definition 5.1. Take $[s, t] \subset [a, b]$ and $\xi \in \mathbb{R}^n$. The *reachable set on* $[s, t]$ *emanating from* ξ, written $R(t, s, \xi)$, is

$$R(t, s, \xi) := \{(t, x(t)) \colon x(\cdot) \in AC([s, t]; \mathbb{R}^n),\ x(s) = \xi \text{ and } \dot{x}(\sigma) \in F(\sigma, x(\sigma)), \text{ a.e. } [s, t]\}.$$

Minor variants on the arguments in the proof of [VW] yield:

Lemma 5.2. *Assume* (C1)–(C5). *Then* $R(\cdot, \cdot, \cdot)$ *takes values nonempty compact sets and a constant* K *exists such that*

$$\operatorname{dist}_H(R(t, s, \xi), R(t', s', \xi') \le K|(t, x, \xi) - (t', s', \xi')|$$

for all $(t, s, \xi), (t', s', \xi') \in ([a, b] \times [a, b] \times \mathbb{R}^n)^2$ *such that* $s \le t$, $s' \le t'$.

Lemma 5.3. *Assume* (C1)–(C5). *Then a null set* $J \subset [a, b)$ *exists such that*

$$F(t, \xi) = \lim_{h \downarrow 0} h^{-1} R(t + h, t, \xi)$$

for all $(t, \xi) \in ([a, b) \backslash J) \times \mathbb{R}^n$. (*The limit on the right is the Karatowski limit; part of the assertion of the lemma is that this limit exists.*)

Proof. See [VW]. ■

We note also

Lemma 5.4. *Suppose that* $\varphi: [a, b] \times \mathbb{R}^n \to \mathbb{R}$ *is a locally Lipschitz continuous function. Then*:

(i) *For every* $(t, \xi) \in [a, b) \times \mathbb{R}^n$ *and* $e \in \mathbb{R}^n$, *the lower Dini derivative reduces to*

$$d^- \varphi((t, \xi), (1, e)) = \liminf_{\varepsilon \downarrow 0} h^{-1}\{\varphi(t + h, \xi + he) - \varphi(t, \xi)\}.$$

(ii) *Given any sequence of numbers* $\{h_i\}$, $h_i \downarrow 0$, *and sequence of functions* $x_i(\cdot)$: $[t, t + h_i] \to \mathbb{R}^n$ *such that*

$$h_i^{-1}[x_i(t + h_i) - \xi] \to e \qquad \textit{as} \quad i \to \infty \quad \textit{for some} \quad e \in \mathbb{R}^n,$$

we have

$$\liminf_{i \to \infty} h_i^{-1}\{\varphi(t + h_i, x_i(t + h_i)) - \varphi(t, \xi)\}$$
$$= \liminf_{i \to \infty} h_i^{-1}\{\varphi(t + h_i, \xi + h_i e) - \varphi(t, \xi)\}.$$

Proof. See [VW]. ∎

We are now ready to list significant properties of V:

Proposition 5.5. *Assume* (C1)–(C5). *Then*:

(a) $V(\cdot, \cdot)$ *is Lipschitz continuous on* $[a, b] \times \mathbb{R}^n$.
(b) $V(t, \xi) \le g(t, \xi)$ *for all* $(t, \xi) \in [a, b] \times \mathbb{R}^n$.
(c) *A nullset* $J \subset [a, b)$ *exists such that*

$$\min_{e \in F(t, \xi)} d^- V((t, \xi), (1, e)) \ge 0 \qquad \textit{for all} \quad (t, \xi) \in ([a, b)\backslash J) \times \mathbb{R}^n.$$

Proof. (a) Take any $d > 0$ and choose (s, ξ), $(s', \xi') \in [a, b] \times \mathbb{R}^n$ such that $|(s, \xi) - (s', \xi')| \le d$. Since $R(\cdot, \cdot, \cdot)$ is continuous and compact valued, the set

$$S := \{(t, e) \in [a, b] \times \mathbb{R}^n : t \ge s, e \in R(t, s, \xi)\}$$

is compact. So we may find $(t, e) \in S$ such that $g(t, e)$ $(= \text{Min}_S\, g) = V(s, \xi)$. Set $t' = \max\{s', t\}$. Then $|(t, \xi) - (t', \xi)| \le |s - s'|$, so by Lipschitz continuity $e' \in R(t', s', \xi')$ exists such that $|e' - e| \le \sqrt{2Kd}$, where K is the constant from Lemma 5.2. We note that $|(t', e') - (t, e)| \le (1 + 2K^2)^{1/2} d$. However, then

$$V(s', \xi') \le g(t', e') \le g(t, e) + l(1 + 2K^2)^{1/2} d,$$

where l is the Lipschitz constant of g,

$$= V(s, \xi) + l(1 + 2K^2)^{1/2} d.$$

Since the roles of (s, ξ) and (s', ξ') in this argument are interchangeable we conclude

$$|V(s', \xi') - V(s, \xi)| \le l(1 + 2K^2)^{1/2} d.$$

(b) This is obvious, since the pay-off for terminating the state trajectory immediately is $g(t, \xi)$ and this then is an upper bound on the minimum cost $V(t, \xi)$.

(c) Take J to be the nullset of Lemma 5.3. Fix $t \in [a, b)\backslash J$, $\xi \in \mathbb{R}^n$, and $e \in F(t, \xi)$. Choose a sequence of numbers $\{h_i\}$ in $[0, b - t)$ such that

$$d^- V((t, \xi), (1, e)) = \lim_i h_i^{-1}[V(t + h_i, \xi + h_i e) - V(t, \xi)]. \tag{5.1}$$

(This is possible by Lemma 5.4.) Lemma 5.3 tell us that

$$e \in \liminf_{h \downarrow 0} h^{-1}(R(t + h_i t, \xi) - \xi)$$

which means that a sequence of absolutely continuous functions $\{x_i(\cdot): [t, t + h_i] \to \mathbb{R}^n\}$ exists such that $x_i(t) = \xi$, $\dot{x}_i \in F(t, x_i)$, for each i, and

$$h_i^{-1}(x_i(t + h_i) - \xi) \to e \tag{5.2}$$

as $i \to \infty$. By definition of the value function, however,

$$V(t + h_i, x_i(t + h_i)) - V(t, \xi) \geq 0.$$

It follows that

$$0 \leq \liminf_{i \to \infty} h_i^{-1}[V(t + h_i, x_i(t + h_i)) - V(t, \xi)].$$

However, the limit on the right

$$= \liminf_{i \to \infty} h_i^{-1}[V(t + h_i, \xi + h_i e) - V(t + \xi)]$$

by (5.2) and, in view of Lemma 5.4,

$$= d^- V((t, \xi); (1, e))$$

by (5.1). Since e was arbitrary, it follows that

$$\inf_{e \in F(t, \xi)} d^- V((t, \xi); (1, e)) \geq 0 \qquad \text{for all} \quad (t, \xi) \in ([a, b)\backslash J) \times \mathbb{R}^n. \qquad \blacksquare$$

6. Proof of Theorem 4.2

We suppose $\varepsilon > 0$ is given, and also an admissible trajectory $(\bar{T}, \bar{x}(\cdot))$ for (P) such that hypotheses (H1)–(H5) are satisfied on the ε neighbourhood of $(\bar{T}, \bar{x}(\cdot))$.

(i) (Sufficiency) Take any verification function φ for $(\bar{T}, \bar{x}(\cdot))$ on the ε neighbourhood of $(\bar{T}, \bar{x}(\cdot))$. Let J be the nullset of property (c) in Definition 4.1, and let $(T, x(\cdot))$ be any admissible state trajectory in the ε neighbourhood of $(\bar{T}, \bar{x}(\cdot))$. We must show that $g(T, x(T)) \geq g(\bar{T}, \bar{x}(\bar{T}))$.

The function $t \mapsto \varphi(t, x(t))$ is Lipschitz continuous, as the composition of two Lipschitz continuous functions. At all points $t \in (a, \bar{T})\backslash J$ where $t \mapsto \varphi(t, x(t))$ is differentiable, where

$$\lim_{h \downarrow 0} h^{-1}[x(t + h) - x(t)] = \dot{x}(t)$$

and where $\dot{x}(t) \in F(t, x(t))$ (such points have full measure) we have

$$\begin{aligned}\frac{d}{dt}\varphi(t, x(t)) &= \lim_{h \downarrow 0} h^{-1}[\varphi(t+h, x(t+h)) - \varphi(t, x(t))] \\ &= \lim_{h \downarrow 0} h^{-1}[\varphi(t+h, x(t) + h\dot{x}(t)) - \varphi(t, x(t))] \\ &= d^{-}\varphi((t, x(t)); (1, \dot{x}(t)) \geq 0\end{aligned}$$

by Lemma 5.4 and since φ is a verification function on the ε neighbourhood of $(\bar{T}, \bar{x}(\cdot))$. It follows that

$$\varphi(a, x_0) = \varphi(T, x(T)) - \int_a^T \frac{d}{dt}\varphi(t, x(t))\, dt \leq \varphi(T, x(T)).$$

However by property (b) of the verification functions (we refer again to Definition 4.1),

$$\varphi(T, x(T)) \leq g(T, x(T)),$$

since $(T, x(T)) \in C \cap \{(t, x): |t - \bar{T}| < \varepsilon \text{ and } |x - \bar{x}(\bar{T})| < \varepsilon\}$. The resulting inequality, combined with the condition "$\varphi(\bar{T}, \bar{x}(\bar{T})) = g(\bar{T}, \bar{x}(\bar{T}))$", confirms local optimality of $(\bar{T}, \bar{x}(\cdot))$.

(ii) (Necessity) Assume now that $(\bar{T}, \bar{x}(\cdot))$ is a minimizer on the ε neighbourhood about $(\bar{T}, \bar{x}(\cdot))$. We must construct a verification function on $N_{\varepsilon'}(\bar{T}, \bar{x}(\cdot))$, for some $\varepsilon' \in (0, \varepsilon)$, confirming local optimality of $(\bar{T}, \bar{x}(\cdot))$. The verification function is obtained from the value function for a derived problem, which we now introduce. The data for this problem is conveniently expressed in terms of the "truncation" function $\mathrm{tr}_{\varepsilon_1}$:

$$\mathrm{tr}_{\varepsilon_1}(\xi) := \begin{cases} \xi & \text{if } |\xi| \leq \varepsilon_1, \\ \varepsilon_1 \xi / |\xi| & \text{if } |\xi| > \varepsilon_1, \end{cases}$$

in which ε_1 is a parameter.

For $(t, \xi) \in [a, b] \times \mathbb{R}^n$ define

$$F_{\varepsilon_1}(t, \xi) := F(t, \bar{x}(t) + \mathrm{tr}_{\varepsilon_1}(\xi - \bar{x}(t)))$$

(remember the convention that $\bar{x}(\cdot)$ is extended to all of $[a, b]$ by constant extrapolation) and

$$g_{\varepsilon_1}(T, \xi) := g(\bar{T} + \mathrm{tr}_{\varepsilon_1}(t - \bar{T}), \bar{x}(\bar{T}) + \mathrm{tr}_{\varepsilon_1}(\xi - \bar{x}(\bar{T}))).$$

We also require the following functions, v_{ε_1} and w_{ε_1} (which also depend on the parameter ε_1):

$$v_{\varepsilon_1}(t, \xi) := \max\{|\xi - \bar{x}(\bar{T})| - \varepsilon_1/2, |T - \bar{T}| - \varepsilon_1/2, 0\}$$

and

$$w_{\varepsilon_1}(t, \xi) := \begin{cases} 0 & \text{if } |\xi - \bar{x}(t)| < \varepsilon_1/2, \\ |\xi - \bar{x}(t)| - \varepsilon_1/2 & \text{if } \varepsilon_1/2 \leq |\xi - \bar{x}(t)| \leq \varepsilon_1, \\ \varepsilon_1/2 & \text{if } \varepsilon_1 < |\xi - \bar{x}(t)| \end{cases}$$

for $(t, \xi) \in [a, b] \times \mathbb{R}^n$.

The derived problem, which involves parameters ε_1, k, and K, is as follows:

$(\mathrm{D}_{\varepsilon_1,k,K})$ Minimize $\eta_{\varepsilon_1,k,K}(T, (y(\cdot), x(\cdot)))$ over $T \in [a, b]$ and $(y(\cdot), x(\cdot)) \in AC([a, T]; \mathbb{R}^{1+n})$ which satisfy

$$(\dot{y}(t), \dot{x}(t)) \in \{kw_{\varepsilon_1}(t, x(t))\} \times F_{\varepsilon_1}(t, x(t)) \qquad \text{a.e.} \quad t \in [a, T]$$

and

$$(y(a), x(a)) = (0, x_0).$$

The cost function $\eta_{\varepsilon_1,k,K}$ in this problem is

$$\eta_{\varepsilon_1,k,K}(T, (y(\cdot), x(\cdot))) := g_{\varepsilon_1}(T, x(T)) + y(T) + kv_{\varepsilon_1}(T, x(T)) + K^2 d_C(T, x(T)).$$

We verify the following claim in the Appendix. The normality hypothesis has a central role in its proof.

Claim. *Numbers* $\varepsilon_1 \in (0, \varepsilon)$, $k > 0$, *and* $K > 0$ *exist such that* $(\overline{T}, y(\cdot) \equiv 0, \bar{x}(\cdot))$ *is a (global) minimizer for* $(\mathrm{D}_{\varepsilon_1,k,K})$.

Fix ε_1, k, and K as in the claim. $(\mathrm{D}_{\varepsilon_1,k,K})$ is an example of problem (Q) of Section 5, with data which satisfies hypotheses (C1)–(C5). Denote by $\tilde{V}$ the value function for $(\mathrm{D}_{\varepsilon_1,k,K})$. $\tilde{V}$ then has properties listed in Proposition 5.5. It is evident that $\tilde{V}$ has the form

$$\tilde{V}(t, (\eta, \xi)) = \eta + V(t, \xi) \qquad \text{for} \quad (t, (\eta, \xi)) \in [a, b] \times \mathbb{R}^{1+n}$$

for some Lipschitz continuous function V. By property (c) of the value function a nullset $J \subset [a, b)$ exists such that

$$\min_{v \in F_{\varepsilon_1}(t,\xi)} d^- V((T, \xi); (1, v)) + kw_{\varepsilon_1}(t, \xi) \geq 0 \tag{6.1}$$

for all $(t, \xi) \in ([a, b)\backslash J) \times \mathbb{R}^n$, while property (b) tells us

$$V(t, \xi) \leq \tilde{g}(t, \xi) \tag{6.2}$$

for all $(t, \xi) \in ([a, b)\backslash J) \times \mathbb{R}^n$ where

$$\tilde{g}(t, \xi) := g_{\varepsilon_1}(t, \xi) + kV_{\varepsilon_1}(t, \xi) + K^2 d_C(t, \xi).$$

We also note that, since $(\overline{T}, \bar{y}(\cdot) \equiv 0, \bar{x}(\cdot))$ is a minimizer for $(\mathrm{D}_{\varepsilon_1,k,K})$,

$$V(a, x_0) = g(\overline{T}, \bar{x}(\overline{T})). \tag{6.3}$$

Now set $\varepsilon' = \varepsilon_1/2$. Evidently $w_{\varepsilon_1}(t, \xi) = 0$ and $F_{\varepsilon_1}(t, \xi) = F(t, \xi)$ for $(t, \xi) \in N_{\varepsilon'}(\overline{T}, \bar{x}(\cdot))$. Also $\tilde{g}(t, \xi) = g(t, \xi)$ whenever $|\xi - \bar{x}(\overline{T})| \leq \varepsilon'$, $|T - \overline{T}| \leq \varepsilon$, and $(t, \xi) \in C$. The fact that $V(\cdot, \cdot)$ is a verification function for $(\overline{T}, \bar{x}(\cdot))$ on the ε' neighbourhood of $(\overline{T}, \bar{x}(\cdot))$ follows from these observations and (6.1), (6.2), and (6.3). ■

Appendix

This appendix is devoted to proving the claim in Section 6 that parameter values, ε_1, k, and K can be chosen in problem $(D_{\varepsilon_1,k,K})$ in such manner that $(\bar{T}, \bar{y}(\cdot) \equiv 0, \bar{x}(\cdot))$ is a minimizer for $(D_{\varepsilon_1,k,K})$. This will complete the proof of Theorem 4.2.

In accordance with the hypotheses of Theorem 4.2, part (ii), we assume that hypotheses (H1)–(H5) are satisfied with reference to the neighbourhood $N_\varepsilon(\bar{T}, \bar{x}(\cdot))$, and (P) is normal at $(\bar{T}, \bar{x}(\cdot))$.

The role of the following lemma is to show that, provided k is chosen appropriately, minimizers for the derived problem are in some sense close to $(\bar{T}, \bar{x}(\cdot))$.

Lemma A.1. *Take any $\varepsilon' \in (0, \varepsilon)$. Then k can be chosen such that*

$$\eta_{\varepsilon',k,K}(T, y(\cdot), x(\cdot)) > \eta_{\varepsilon',k,K}(\bar{T}, \bar{y}(\cdot) \equiv 0, \bar{x}(\cdot))$$

for any $K \geq 0$ and for any admissible trajectory $(T, y(\cdot), x(\cdot))$ for $(D_{\varepsilon',k,K})$ satisfying

$$\max\{\|x(\cdot) - \bar{x}(\cdot)\|_{L^\infty([a,T])}, |T - \bar{T}|, |x(T) - \bar{x}(\bar{T})|\} \geq \varepsilon'.$$

Proof. Let c be the constant of hypothesis (H4) corresponding to $N_\varepsilon(\bar{T}, \bar{x}(\cdot))$, and let κ be a bound on $|g(t, \xi)|$ over the set $\{(t, \xi): |t - \bar{T}| \leq \varepsilon, |\xi - \bar{x}(\bar{T})| \leq \varepsilon\}$. Notice that for any $\varepsilon' \in (0, \varepsilon)$ we have $|g_{\varepsilon'}(t, \xi)| \leq \kappa$ and $F_{\varepsilon'}(t, \xi) \subset cB$ for all $(t, \xi) \in [0, b] \times \mathbb{R}^n$.

Take any $\varepsilon' \in (0, \varepsilon)$ and $K \geq 0$. Select $\delta \in (0, \varepsilon'/9c)$ and $k > \max\{4\kappa/\varepsilon', 8\kappa/\delta\varepsilon'\}$. We show that k chosen in this way has the desired properties. Let $(T, (y(\cdot), x(\cdot)))$ be an admissible trajectory for $(D_{\varepsilon',k,K})$. Consider first the case

$$\max\{|T - \bar{T}|, |x(T) - \bar{x}(\bar{T})|\} \geq \varepsilon'.$$

We have

$$\eta_{\varepsilon',k,K}(T, (y(\cdot), x(\cdot))) \geq -\kappa + k\varepsilon'/2 > \kappa \geq \eta_{\varepsilon',k,K}(\bar{T}, (\bar{y}(\cdot) \equiv \bar{x}(\cdot))).$$

It remains then to consider the case when $|x(t) - \bar{x}(t)| \geq \varepsilon'$ for some $t \in [0, b]$. In view of the facts that $x(0) = \bar{x}(0)$ and $|\dot{x}(t)|, |\dot{\bar{x}}(t)| < c$ (almost everywhere) we have that $t - a > \varepsilon'/2c$. It follows that $\delta\ (< \varepsilon'/8c) < t - a$. Furthermore, $|x(s) - \bar{x}(s)| \geq 3\varepsilon'/4$ for $t - \delta \leq s \leq t$. However, then $\eta_{\varepsilon',k,K}(T, (y(\cdot), x(\cdot))) \geq -\kappa + (\delta\varepsilon'/4)k > \kappa \geq \eta_{\varepsilon',k,K}(\bar{T}, (\bar{y}(\cdot) \equiv 0, \bar{x}(\cdot)))$. The assertions of the lemma have been proved in this case too. ■

We now return to the proof of the claim. Take any sequence $\{\varepsilon_i\}$ of numbers in $(0, \varepsilon)$ such that $\varepsilon_i \downarrow 0$. Let $\{k_i\}$ be the sequence of numbers corresponding to $\varepsilon' = \varepsilon_2, \varepsilon_3, \ldots$ as in Lemma A.1. Now choose $\{K_i\}$to be a sequence of positive numbers such that

$$K_i \uparrow \infty \quad \text{and} \quad k_i/K_i \to 0. \tag{A.1}$$

Consider the sequence of problems $(D_{\varepsilon_i,k_i,K_i})$, $i = 1, 2, \ldots$. For each i, $(D_{\varepsilon_i,k_i,K_i})$ has a minimizer. Write it as $(T_i, (y_i(\cdot), x_i(\cdot))$. By Lemma A.1,

$$|T_i - \bar{T}| \leq \varepsilon_i, \qquad |x_i(T_i) - \bar{x}(\bar{T})| \leq \varepsilon_i, \quad \text{and} \quad \|x_i(\cdot) - \bar{x}(\cdot)\|_\infty \leq \varepsilon_i, \qquad i = 1, 2, \ldots. \tag{A.2}$$

Suppose that $(T_i, x_i(T_i) \in C$ for some i. Then, by local optimality of $(\bar{T}, \bar{x}(\cdot))$ for problem (P), we have

$$\eta_{\varepsilon_i,k_i,K_i}(T_i, y_i(\cdot), x_i(\cdot)) \geq g(T_i, x(T_i)) \geq g(\bar{T}, \bar{x}(\bar{T})) \geq \eta_{\varepsilon_i,k_i,K_i}(\bar{T}, (\bar{y}(\cdot) \equiv 0, \bar{x}(\cdot))).$$

In this case then the claim is true with $(\varepsilon_1, k, K) = (\varepsilon_i, k_i, K_i)$. The remaining case to be considered is when

$$(T_i, x_i, (T_i)) \notin C \qquad \text{for all } i. \tag{A.3}$$

Suppose this is true. We show that a contradiction of the normality hypothesis results, and thereby complete the proof of the claim.

The fact that $(T_i, y_i(\cdot), x_i(\cdot))$ is a minimizer for the derived problem can be expressed as follows. For each i, $(T_i, y_i(\cdot), x_i(\cdot))$ is a local minimizer for the problem:

(Q_i) Minimize $K_i^{-2}g(T, x(T)) + K_i^{-1}y(T) + k_iK_i^{-2}v_{\varepsilon_i}(T, x(T)) + d_C(T, x(T))$ over arcs $(T, y(\cdot), x(\cdot))$ which satisfy

$$(\dot{y}(t), \dot{x}(t)) \in \{0\} \times F(t, x(t)) + \{k_iK_i^{-1}w_{\varepsilon_i}(t, x(t))\} \times \{0\} \qquad \text{a.e.}$$

and

$$(y(0), x(0)) = (0, x_0).$$

Notice that we are justified in expressing (Q_i) in terms of g and F rather than g_{ε_i} and F_{ε_i} because, by (A.2), $(T_i, x_i(\cdot))$ lies in the ε_i neighbourhood of $(\bar{T}, \bar{x}(\cdot))$.

We now apply first-order necessary conditions in the form of a Hamiltonian inclusion, with reference to the local minimizer $(T_i, y_i(\cdot), x_i(\cdot))$ for (Q_i) [CLV]. In doing so we may set the cost multiplier to unity, since admissible arcs for the problem in question have free endpoints. As usual, H denotes the Hamiltonian function:

$$H(t, x, p) := \max_{e \in F(t,x)} p \cdot e.$$

Taking note of our assumption that $K_i \uparrow \infty$ and $k_i/K_i \to 0$, we deduce existence of a sequence of positive numbers $\{\delta_i\}$, $\delta_i \downarrow 0$, a sequence of numbers $\{h_i\}$, and a sequence $\{p_i\}$ of $AC([a, T_i]; \mathbb{R}^n)\}$ such that

$$(-\dot{p}_i(t), \dot{x}_i(t)) \in \partial H(t, x_i(t), p_i(t)) + \delta_i\bar{B} \qquad \text{a.e. on } [a, T_i], \tag{A.4}$$

$$(h_i, -p_i(T_i)) \in \partial d_C(T_i, x_i(T_i)) + \delta_i\bar{B}, \tag{A.5}$$

and

$$h_i \in \operatorname*{co\,ess}_{s \to T_i} H(s, x_i(T_i), p_i(T)) + \delta_i\bar{B}. \tag{A.6}$$

In view of (A.3), all points in $\partial d_C(T_i, x_i(T_i))$ have unit norm. It follows that

$$|(h_i, -p_i(T_i))| \to 1 \qquad \text{as} \quad i \to \infty. \tag{A.7}$$

Now regard the functions $x_i(\cdot)$ and $p_i(\cdot)$ as having the domain the fixed interval $[a, b]$, obtained by constant extrapolation. By (A.2) we have $T_i \to T$ and $\|x_i(\cdot) - \bar{x}(\cdot)\|_{L^\infty\{[a,b]\}} \to 0$ as $i \to \infty$. The $p_i(\cdot)$'s are uniformly bounded and equicontinuous. Following extraction of subsequences then we are assured that $h_i \to h$ for some number h, and $\|p_i(\cdot) - p(\cdot)\|_{L^\infty([a,b])} \to 0$ for some $p(\cdot) \in AC([a, b]; \mathbb{R}^n)$.

Relationships (A.4), (A.5), and (A.6) are robust under limit taking and yield

$$(-\dot{p}(t), \dot{\bar{x}}(t)) \in \partial H(t, \bar{x}(t), p(t)) \qquad \text{a.e. on } [a, \bar{T}],$$

$$(h, -p(\bar{T})) \in \partial d_C(\bar{T}, \bar{x}(\bar{T})),$$

$$h \in \operatorname*{co\,ess}_{s \to \bar{T}} H(s, \bar{x}(\bar{T}), p(\bar{T})).$$

By (A.6) and (A.7), however, $p(\bar{T}) \neq 0$. Since $\partial d_C \subset N_C$ we have arrived at a contradiction of the normality assumption.

References

[B1] M. Bardi, A boundary value problem for the minimum time function, *SIAM J. Control Optim.*, **27** (1989), 778–785.

[B2] V. G. Boltyanskii, Sufficient conditions for optimality and justification of the dynamic programming method, *SIAM J. Control*, **4** (1966), 326–361.

[C] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley–Interscience, New York, 1983.

[CLV] F. H. Clarke, P. D. Loewen, and R. B. Vinter, Differential inclusions with free time, *Ann. Inst. Henri Poincaré Anal. Non Linéaire*, **5** (1988), 573–593.

[CV] 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.

[EJ] L. C. Evans and M. R. James, The Hamilton–Jacobi–Bellman equation for time optimal control, *SIAM J. Control Optim.*, **27** (1989), 1477–1489.

[FR] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, 1975.

[F] H. Frankowska, Optimal trajectories associated to a solution of the contingent Hamilton–Jacobi equation, *Appl. Math. Optim.*, **19** (1989), 291–311.

[K] J. Kotsiopoulos, Characterization of Deterministic Optimal Processes with Discontinuous Trajectories, Ph.D. thesis, Imperial College, University of London, 1990.

[L] P. L. Lions, *Generalized Solutions of Hamilton–Jacobi Equations*, Pitman, Boston, 1982.

[VW] R. B. Vinter and P. R. Wolenski, Hamilton–Jacobi theory for optimal control problems with data measurable in time, *SIAM J. Control Optim.*, **28** (1990), 1404–1419.