This paper was presented at the 38th IEEE Conference on Decision and Control, held in Phoenix, Arizona, December 7-10, 1999, and has appeared in the conference proceedings. The date of this version is August 12, 1999.

The Euler-Lagrange equation under weak regularity conditions [1]

Héctor J. Sussmann
Department of Mathematics; Rutgers, the State University of New Jersey
Hill Center—Busch Campus, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA
E-mail: sussmann@hamilton.rutgers.edu Web site: http://www.math.rutgers.edu/~sussmann

1. Introduction

This note is a continuation of our recent work (cf. Sussmann [4, 5, 6, 7, 8, 9]), on a general nonsmooth version of the finite-dimensional Pontryagin Maximum Principle (cf. Pontryagin *et al.* [3], Berkovitz [1]). Our purpose here is to announce a further extension of this result for the special case of finite-dimensional classical calculus of variations problems in $\mathbb{R}^n$, with constraints $\dot{\xi}(t) \in U$, where U is a given subset of $\mathbb{R}^n$. For simplicity, we only consider problems with fixed endpoints.

Naturally, our result is a generalization of the classical Euler-Lagrange equations with the Weierstrass side conditions, stated in the Hamiltonian language of optimal control theory. As in our earlier work on the Maximum Principle, the approach used here is classical, in the sense that it involves "needle variations" and uses a fixed-point argument.

2. The Euler-Lagrange equation

Let us assume that

(A1) *Q is an open subset of $\mathbb{R}^n$.*
(A2) *U is a subset of $\mathbb{R}^n$.*
(A3) *L is a function from $Q \times U \times [a, b]$ to $\mathbb{R}$.*

We define $\mathcal{A}(Q, U, L, a, b)$—the set of "admissible arcs"—to be the set of all absolutely continuous maps $\xi : [a, b] \to \mathbb{R}^n$ such that $\xi(t) \in Q$ for all $t \in [a, b]$, $\dot{\xi}(t) \in U$ for almost all $t \in [a, b]$, and the function $t \to L(\xi(t), \dot{\xi}(t), t)$ is integrable on $[a, b]$. For fixed $\bar{q}, \hat{q} \in Q$, we let $\mathcal{A}_{\bar{q},\hat{q}}(Q, U, L, a, b)$ denote the set of all $\xi \in \mathcal{A}$ such that $\xi(a) = \bar{q}$ and $\xi(b) = \hat{q}$.

We then assume

(A4) *$\mathcal{N}$ is a subset of $\mathcal{A}(Q, U, L, a, b)$,*

and write $\mathcal{N}_{\bar{q},\hat{q}}(\delta) \stackrel{\text{def}}{=} \mathcal{N} \cap \mathcal{A}_{\bar{q},\hat{q}}(Q, U, L, a, b)$.

We then consider a standard fixed-endpoints calculus of variations problem of the form

$$\boxed{\begin{array}{l} \text{minimize } \hat{I}(\xi, \eta) = \int_a^b L(\xi(t), \eta(t), t)dt\,, \\ \text{subject to } \ \xi \in \mathcal{N}_{\bar{q},\hat{q}}(\delta) \ \text{ and} \\ \dot{\xi}(t) = \eta(t) \text{ for almost all } t \in [a, b]\,. \end{array}} \tag{1}$$

We define a function $H : Q \times U \times \mathbb{R}^n \times [a, b] \to \mathbb{R}$—the "normal Hamiltonian"—by letting

$$H(q, u, p, t) = p.u - L(q, u, t)\,. \tag{2}$$

Suppose that (ξ_*, η_*) is a solution of (1), so $\eta_* = \dot{\xi}_*$. Let us compare $I(\xi_*)$ with the cost $I(\xi)$ of another trajectory $\xi \in \mathcal{A}(Q, U, L, a, b)$ such that $\xi(a) = \xi_*(a)$. We find

$$I(\xi) - I(\xi_*) = \int_a^b \psi(t).\Big(\dot{\xi}(t) - \eta_*(t)\Big)\, dt$$
$$-\int_a^b\Big(H(\xi(t), \dot{\xi}(t), \psi(t), t) - H(\xi_*(t), \eta_*(t), \psi(t), t)\Big)dt,$$

where $[a, b] \ni t \to \psi(t) \in \mathbb{R}^n$ is, for the time being, an arbitrary absolutely continuous function of t, and we have expressed L terms of H using the identity $L(q, u, t) = p \cdot u - H(q, u, p, t)$. Integration by parts—which is perfectly justified because ψ, ξ, and ξ_* are absolutely continuous—then yields

$$I(\xi) - I(\xi_*) =$$
$$-\int_a^b \dot{\psi}(t).\Big(\xi(t) - \xi_*(t)\Big)\, dt + \psi(b) \cdot (\xi(b) - \xi_*(b))$$
$$-\int_a^b\Big(H(\xi(t), \dot{\xi}(t), \psi(t), t) - H(\xi_*(t), \eta_*(t), \psi(t), t)\Big)dt,$$

where we have used the fact that $\xi(a) - \xi_*(a) = 0$.

For $\delta > 0$, use $\mathcal{T}(\xi_*, \delta)$ to denote the tube

$$\mathcal{T}(\xi_*, \delta) \stackrel{\text{def}}{=} \{(q, t) : a \le t \le b, \|q - \xi_*(t)\| \le \delta\}\,. \tag{3}$$

Since $\xi_* \in \mathcal{A}(Q, U, L, a, b)$, it is clear that there exists a $\bar{\delta}_1$ such that

(C1) *$\bar{\delta}_1 > 0$ and $\mathcal{T}(\xi_*, \bar{\delta}_1)$ is a subset of $Q \times [a, b]$.*

We then use $\mathcal{A}^\delta(Q, U, L, a, b)$, $\mathcal{A}^\delta_{\bar{q},\hat{q}}(Q, U, L, a, b)$, to denote, respectively, the set of all $\xi \in \mathcal{A}(Q, U, L, a, b)$ such that the graph of ξ is contained in $\mathcal{T}(\xi_*, \delta)$, i.e. such that $\|\xi(t) - \xi_*(t)\| \le \delta$ for all $t \in [a, b]$ and the set $\mathcal{A}_{\bar{q},\hat{q}}(Q, U, L, a, b) \cap \mathcal{A}^\delta(Q, U, L, a, b)$. We define

$$\begin{array}{rcl} \mathcal{N}(\delta) & = & \mathcal{N} \cap \mathcal{A}^\delta(Q, U, L, a, b)\,, \\ \mathcal{N}_{\bar{q},\hat{q}}(\delta) & \stackrel{\text{def}}{=} & \mathcal{N} \cap \mathcal{A}^\delta_{\bar{q},\hat{q}}(Q, U, L, a, b)\,. \end{array}$$

We then assume:

(A5) *$L(q, u, t)$ is continuous as a function of q for each fixed $u \in U$, $t \in [a, b]$.*
(A6) *the function $t \to L(q, u, t)$ is Lebesgue measurable for each fixed $q \in Q$, $u \in U$.*
(A7) *the function $t \mapsto L(q, \dot{\xi}_*(t), t)$ is Lebesgue measurable for each fixed $q \in Q$.*
(A8) *For some $\bar{\delta}_2 \in]\,0, \bar{\delta}_1]$, there exist*
(A8.1) *a measurable set-valued map $W : [a, b] \to 2^{\mathbb{R}^n}$ with nonempty compact convex values,*
(A8.2) *measurable functions $k_\delta : [a, b] \to [0, +\infty]$, for $0 < \delta \le \bar{\delta}_2$,*
(A8.3) *an integrable function $k : [a, b] \to [0, +\infty]$,*
such that
(A8.a) *the inequality*

$$\begin{array}{rcl} L(q, \eta_*(t), t) & \le & L(\xi_*(t), \eta_*(t), t) \\ & & +F_{W(t),\xi_*(t)}(q) + \delta k_\delta(t) \end{array} \tag{4}$$

holds whenever $(q, t) \in \mathcal{T}(\xi_*, \delta)$, where

$$F_{W,x}(q) = \max\{w.(q-x) : w \in W(t)\}\,.$$

(A8.b) $\sup\Big\{\|w\| : w \in W(t)\Big\} \le k(t)$ *whenever $t \in [a, b]$,*

[1] Research supported in part by NSF Grant DMS98-03411-00798 and AFOSR Grant 0923.

(A8.c) $\lim_{\delta\downarrow 0}\int_a^b k_\delta(t)dt = 0$.

Assumption (A8) clearly implies that there exist an integrable function $\hat{k} : [a,b] \to \mathbb{R}\cup\{+\infty\}$ and a $\bar{\delta}_3$ in $]0,\bar{\delta}_2]$ such that the *Carathéodory-type upper bound*

$$L(q,\eta_*(t),t) \le \hat{k}(t) \text{ for } (q,t)\in\mathcal{T}(\xi_*,\bar{\delta}_3) \qquad (5)$$

holds. We then define $W_\delta(q,t)$, for $\delta\in]0,\bar{\delta}_3]$ and (q,t) in $\mathcal{T}(\xi_*,\delta)$, to be the set of those $w\in W(t)$ such that

$$L(q,\eta_*(t),t) \le L(\xi_*(t),\eta_*(t),t) + w.(q-\xi_*(t)) + \delta k_\delta(t)\,.$$

Clearly, $W_\delta(q,t)$ is compact, convex and nonempty for every $q,t\in\mathcal{T}(\xi_*,\delta)$, as long as $\delta\le\bar{\delta}_2$.

If $\xi\in\mathcal{A}(Q,U,L,a,b)$, we define $W_\delta^\xi(t)\stackrel{\rm def}{=}W_\delta(\xi(t),t)$. Then $[a,b]\ni t\to W_\delta^\xi(t)\subseteq\mathbb{R}^n$ is a measurable set-valued function with nonempty compact convex values, provided that $\delta\le\bar{\delta}_2$, and $\xi\in\mathcal{A}^\delta(Q,U,L,a,b)$. (The fact that the values are compact, convex and nonempty is trivial. To prove measurability of W_δ^ξ, we have to show that for each compact subset K of $\mathbb{R}^n$, if we define

$$E_K = \{t\in[a,b] : W_\delta^\xi(t)\cap K\ne\emptyset\}\,,$$

then E_K is Lebesgue measurable. Let $\mathbb{Q}^n$ be the set of all points of $\mathbb{R}^n$ with rational coordinates. Define

$$\begin{aligned}\rho(t) &= L(\xi(t),\eta_*(t),t) - L(\xi_*(t),\eta_*(t),t)\,,\\ \sigma_q(t) &= \operatorname{dist}(q,W(t)) \text{ for } q\in\mathbb{R}^n\,.\end{aligned}$$

Then k_δ, ρ and the σ_q are measurable functions. By Lusin's theorem, there are pairwise disjoint subsets J_i, $i=0,1,2,\ldots$, of $[a,b]$ such that

(i) for $i>0$, the set J_i is compact and the restrictions $k_\delta\lceil J_i$, $\rho\lceil J_i$, $\sigma_q\lceil J_i$ of k_δ, ρ, and the σ_q for $q\in\mathbb{Q}^n$, to J_i are continuous;

(ii) J_0 has measure zero;

(iii) $[a,b]=\cup_{i=0}^\infty J_i$.

To prove that E_K is measurable, it suffices to show that $E_K\cap J_i$ is measurable for $i>0$. Let $i>0$. We will show that $E_K\cap J_i$ is closed. Let $\{t_j\}$ be a sequence of members of $E_K\cap J_i$ that converges to a limit $t\in J_i$. Then $W_\delta^\xi(t_j)\cap K\ne\emptyset$, so we can pick $w_j\in W_\delta^\xi(t_j)\cap K$. Since K is compact we may assume, after passing to a subsequence, that $w=\lim_{j\to\infty}w_j$ exists. Clearly, w belongs to K. To prove that w also belongs to $W_\delta^\xi(t)$, we must show that $\rho(t)\le w\cdot(\xi(t)-\xi_*(t))+\delta k_\delta(t)$ and $w\in W(t)$. The first assertion follows from the fact that $\rho(t_j)\le w\cdot(\xi(t_j)-\xi_*(t_j))+\delta k_\delta(t_j)$—since $w_j\in W_\delta(\xi(t_j),t_j)$—and ρ, ξ, ξ_* and k_δ are continuous on J_i. To prove that $w\in W(t)$, it suffices to show that $\operatorname{dist}(w,W(t))=0$, since $W(t)$ is closed. Assume that $\operatorname{dist}(w,W(t))=\alpha>0$. Pick $q\in\mathbb{Q}^n$ such that $\|w-q\|<\frac{\alpha}{3}$. Then $\sigma_q(t)>\frac{2\alpha}{3}$. On the other hand, $\sigma_q(t_j)\le\|q-w_j\|$, since $w_j\in W(t_j)$. Since σ_q is continuous on J_i, and $w_j\to w$, we have $\sigma_q(t)\le\|q-w\|<\frac{\alpha}{3}$. This contradiction proves that $w\in W(t)$, as desired.)

We let $\Gamma(W)$ denote the set of all measurable selections of W, so the members of $\Gamma(W)$ are the measurable maps $\omega:[a,b]\to\mathbb{R}^n$ such that $\omega(t)$ belongs to $W(t)$ for almost all $t\in[a,b]$. Then $\Gamma(W)$ is a nonempty, convex, weakly compact subset of $L^1([a,b],\mathbb{R}^n)$. Similarly, we use $\Gamma(W_\delta^\xi)$ to denote the set of all measurable selections of W_δ^ξ, so $\Gamma(W_\delta^\xi)$ is also a nonempty, convex, weakly compact subset of $L^1([a,b],\mathbb{R}^n)$.

We now define

$$\begin{aligned}R_{\xi,\omega}(t) &= L(\xi(t),\eta_*(t),t)-L(\xi_*(t),\eta_*(t),t)\\ &\quad -\omega(t)\cdot(\xi(t)-\xi_*(t))\,.\end{aligned}$$

Then the estimate

$$R_{\xi,\omega}(t)\le\delta k_\delta(t) \qquad (6)$$

holds for $\xi\in\mathcal{A}(Q,U,L,a,b)$, $t\in[a,b]$, $\omega\in\Gamma(W_\delta^\xi)$.

We then have, writing $\xi,\xi_*,\dot{\xi},\eta_*,\omega,R_{\xi,\omega},S_\xi$, rather than $\xi(t)$, $\xi_*(t)$, $\dot{\xi}(t)$, $\eta_*(t)$, $\omega(t)$, $R_{\xi,\omega}(t)$, $S_\xi(t)$,

$$\begin{aligned}H(\xi,\dot{\xi},\psi,t)-H(\xi_*,\eta_*,\psi,t) &= H(\xi_*,\dot{\xi},\psi,t)\\ -H(\xi_*,\eta_*,\psi,t)-\omega.(\xi-\xi_*)-R_{\xi,\omega}-S_\xi\,,\end{aligned}$$

where

$$\begin{aligned}S_\xi(t) &\stackrel{\rm def}{=} \Big(L(\xi(t),\dot{\xi}(t),t)-L(\xi_*(t),\dot{\xi}(t),t)\Big)\\ &\quad -\Big(L(\xi(t),\eta_*(t),t)-L(\xi_*(t),\eta_*(t),t)\Big)\,.\end{aligned}$$

Now let $\psi_\omega(t)=-\int_t^b\omega(s)ds$. Then, if ξ belongs to $\mathcal{A}(Q,U,L,a,b)$,

$$\begin{aligned}&I(\xi)-I(\xi_*)=\mathcal{R}_{\xi,\omega}\\ &+\int_a^b\Big(H(\xi_*,\eta_*,\psi_\omega,t)-H(\xi_*,\eta,\psi_\omega,t)\Big)dt\,,\end{aligned} \qquad (7)$$

where we have used the fact that $\psi_\omega(b)=0$, written $\eta(t)=\dot{\xi}(t)$, and defined the remainder term $\mathcal{R}_{\xi,\omega}$ to be

$$\mathcal{R}_{\xi,\omega}\stackrel{\rm def}{=}\int_a^b R_{\xi,\omega}(t)dt+\int_a^b S_\xi(t)dt\,. \qquad (8)$$

Now $t\to\eta(t)$ is essentially an arbitrary U-valued function, except only for the fact that it has to satisfy the integral constraint $\int_a^b\eta(t)dt=\hat{q}-\bar{q}$. Ignoring this fact temporarily—and ignoring also the presence of the remainder term $\mathcal{R}_{\xi,\omega}$, and the fact that ψ_ω depends on ω—the above formal calculation has achieved the "decoupling" of the individual minimization problems at various times t. If we let $h(t,u)=H(\xi_*(t),u,\psi(t),t)$, then (7) essentially implies that

$$\int_a^b h(t,\eta(t))dt\le\int_a^b h(t,\eta_*(t))dt$$

for all U-valued functions $\eta(\cdot)$, from which it should follow that $h(t,\eta_*(t))=\max\{h(t,u):u\in U\}$ for a. e. t, which is essentially the Hamiltonian maximization condition.

We now make the above argument rigorous, for which it is necessary that we take into account the constraint $\int_a^b\eta(t)dt=\hat{q}-\bar{q}$, and deal with the remainder term $\mathcal{R}_{\xi,\omega}$ and the dependence of ψ on ω. For this purpose, we will make combined needle variations, as we now explain.

We take pairs $(\mathbf{u},\mathbf{t})$ of m-tuples $\mathbf{u}=(u_1,\ldots,u_m)$, $\mathbf{t}=(t_1,\ldots,t_m)$ such that the u_i belong to U and the t_i are points in $[a,b[$. We let $v_i=u_i-\eta_*(t_i)$, and use $\mathcal{V}$ to denote the convex hull of the vectors v_i. We call the pair $(\mathbf{u},\mathbf{t})$ *normal* if 0 is an interior point of $\mathcal{V}$, and we call $(\mathbf{u},\mathbf{t})$ *abnormal* if it is not normal. We consider the normal and abnormal cases separately.

Suppose first that $(\mathbf{u},\mathbf{t})$ is abnormal. This implies that there exists a nonzero vector $\bar{p}\in\mathbb{R}^n$ such that $\bar{p}\cdot(u_i-\eta_*(t_i))\le 0$ for $i=1,\ldots,m$. Define a function

$\psi : [a,b] \to \mathbb{R}^n$ by letting $\psi(t) \equiv \bar{p}$. Let $\psi_0 = 0$. Then ψ trivially satisfies the differential inclusion

$$\dot{\psi}(t) \in \psi_0 W(t) \text{ for almost all } t \in [a,b]\,. \tag{9}$$

Moreover, if we define $H^s(q,u,p,t) = p.u - sL(q,u,t)$, we see that

$$H^{\psi_0}(\xi_*(t_i), u_i, \psi(t_i), t_i) \le H^{\psi_0}(\xi_*(t_i), \eta_*(t_i), \psi(t_i), t_i) \quad \text{for } i = 1,\ldots,m\,. \tag{10}$$

We now assume that $(\mathbf{u},\mathbf{t})$ is normal, and seek to prove the existence of a nontrivial pair (ψ,ψ_0) such that (9) and (10) hold. In this case, we set $\psi_0 = 1$, so H^{ψ_0} is the normal Hamiltonian H defined earlier in (2).

We define

$$\Sigma^m \stackrel{\mathrm{def}}{=} \Big\{ (r_1,\ldots,r_m) \in [0,1]^m : r_1 + \ldots + r_m = 1 \Big\}\,,$$

and use Σ_0^m to denote the subset of Σ^m consisting of all those m-tuples $\mathbf{r} \in \Sigma^m$ such that $\mathbf{r}\cdot\mathbf{v} = 0$, where, for $\mathbf{r} = (r_1,\ldots,r_m)$, we write $\mathbf{r}\cdot\mathbf{v} \stackrel{\mathrm{def}}{=} \sum_{i=1}^m r_i v_i$. Then Σ_0^m is a—possibly empty—compact convex subset of Σ^m.

We now let

$$\begin{aligned} z_i^\omega &= H(\xi_*(t_i), \eta_*(t_i), \psi_\omega(t_i), t_i) \\ &\quad - H(\xi_*(t_i), u_i, \psi_\omega(t_i), t_i)\,, \\ \mathbf{r}.\mathbf{z}^\omega &= \sum_{i=1}^m r_i z_i^\omega\,, \end{aligned}$$

and distinguish two cases, namely,

(A) *There exists $\omega \in \Gamma(W)$ such that $\mathbf{r}.\mathbf{z}^\omega \ge 0$ for all $\mathbf{r} \in \Sigma_0^m$.*

(B) *For every $\omega \in \Gamma(W)$ there exists an $\mathbf{r} \in \Sigma_0^m$ such that $\mathbf{r}.\mathbf{z}^\omega < 0$.*

Suppose (A) holds, and fix an ω that satisfies the conditions of (A). Let C_0 denote the set of vectors $\{(v_i, z_i^\omega) : i = 1,\ldots,m\}$, so C_0 is a finite subset of $\mathbb{R}^{n+1}$. Let C be the convex hull of C_0, so C is a compact convex subset of $\mathbb{R}^{n+1}$. Regarding $\mathbb{R}^{n+1}$ as the product $\mathbb{R}^n \times \mathbb{R}$, let $D = \{0\}\times\,]-\infty,0\,[$. Then $C \cap D = \emptyset$. So there exists a nonzero vector $(\bar{p}, p_0) \in \mathbb{R}^{n+1}$ such that $\bar{p}\cdot v + p_0 z \ge 0$ for all $(v,z) \in C$, and $\bar{p}\cdot v + p_0 z \le 0$ for all $(v,z) \in D$. Since $(0,-1) \in D$, p_0 must be nonnegative. If $p_0 = 0$, it would follow that $\bar{p} \ne 0$, and $\bar{p}\cdot v_i \ge 0$ for $i = 1,\ldots,m$, implying that $(\mathbf{u},\mathbf{t})$ is abnormal, which is a contradiction. So $p_0 \ne 0$, and then $p_0 > 0$, so we can clearly assume that $p_0 = 1$. We then have

$$\bar{p}\cdot(u_i - \eta_*(t_i)) = \bar{p}\cdot v_i \ge -z_i^\omega\,, \tag{11}$$

and

$$\begin{aligned} z_i^\omega &= \psi_\omega(t_i)\cdot(\eta_*(t_i) - u_i) \\ &\quad -(L(\xi_*(t_i), \eta_*(t_i), t_i) - L(\xi_*(t_i), u_i, t_i))\,, \end{aligned}$$

so

$$\begin{aligned} &\Big(\psi_\omega(t_i) - \bar{p}\Big)\cdot(u_i - \eta_*(t_i)) \\ &\le L(\xi_*(t_i), u_i, t_i) - L(\xi_*(t_i), \eta_*(t_i), t_i)\,, \end{aligned} \tag{12}$$

and then, if we let $\psi(t) = \psi_\omega(t) - \bar{p}$, we find

$$\psi(t_i)\cdot(u_i - \eta_*(t_i)) \le L(\xi_*(t_i), u_i, t_i) - L(\xi_*(t_i), \eta_*(t_i), t_i)\,,$$

that is,

$$\begin{aligned} &\psi(t_i)\cdot u_i - L(\xi_*(t_i), u_i, t_i) \\ &\le \psi(t_i)\cdot\eta_*(t_i) - L(\xi_*(t_i), \eta_*(t_i), t_i)\,, \end{aligned} \tag{13}$$

so $H(\xi_*(t_i), u_i, \psi(t_i), t_i) \le H(\xi_*(t_i), \eta_*(t_i), \psi(t_i), t_i)$.

Since $\psi_0 = 1$, the map ψ satisfies (10). Moreover, since $\dot{\psi}(t) = \omega(t)$ a. e., it is clear that ψ satisfies (9) as well.

So in case (A) we have proved the existence of a nontrivial (ψ,ψ_0) for which $\psi_0 \ge 0$ and (9) and (10) hold.

We must now consider the remaining possibility, namely, case (B). In this case, we will construct a "variation" of ξ_* that will lead to a trajectory ξ for which $I(\xi) < I(\xi_*)$, thereby contradicting the assumption that ξ_* solves our minimization problem.

Assumptions (A5), (A6), and (A7), together with the Scorza-Dragoni Theorem, imply that there exist pairwise disjoint subsets $J_\nu(\mathbf{u})$ of $[a,b]$, for $\nu = 0,1,2,\ldots$, such that

(i) for $\nu > 0$, the set $J_\nu(\mathbf{u})$ is compact and the function

$$\mathcal{T}(\xi_*, \bar{\delta}_3) \cap (Q \times J_\nu(\mathbf{u})) \ni (q,t) \to \Lambda(q,t) \in \mathbb{R}^{1+m+n}$$

is continuous, where we use $\Lambda(q,t)$ to denote the vector

$$(L(q, \eta_*(t), t), L(q, u_1, t), \ldots, L(q, u_m, t), \eta_*(t));$$

(ii) $J_0(\mathbf{u})$ has measure zero;

(iii) $[a,b] = \cup_{\nu=0}^\infty J_\nu(\mathbf{u})$.

For $\nu = 1,2,\ldots$, we let $\tilde{J}_\nu(\mathbf{u}) = \cup_{k=1}^\nu J_k(\mathbf{u})$, and observe that Λ is continuous on $\mathcal{T}(\xi_*, \bar{\delta}_3) \cap (Q \times \tilde{J}_\nu(\mathbf{u}))$, because the $J_k(\mathbf{u})$, for $k = 1,\ldots,\nu$, are compact and pairwise disjoint. We let $G_\nu(\mathbf{u})$ denote the set of all $t \in [a,b]$ such that t is a point of density of $\tilde{J}_\nu(\mathbf{u})$ and a Lebesgue point of η_*.

Remark 1 Recall that a *Lebesgue point* of a scalar or vector-valued function μ on an interval $[a,b]$ is a point $t \in\,]\,a,b\,[$ such that

$$\lim_{h\downarrow 0} h^{-1} \int_{t-h}^{t+h} \|\mu(s) - \mu(t)\| ds = 0\,.$$

A *point of density* of a measurable subset E of $[a,b]$ is a point $t \in E$ which is a Lebesgue point of the indicator function χ_E. Equivalently, t is a point of density of E iff $t \in E$, $a < t < b$, and $\lim_{h\downarrow 0} h^{-1}\mathrm{meas}(E \cap [t-h, t+h]) = 0$. It is well known that if μ is integrable then almost every point of $[a,b]$ is a Lebesgue point of μ. This implies, in particular, that if E is measurable then almost all points of E are points of density. $\diamondsuit$

We let $G(\mathbf{u}) = \cup_{\nu=1}^\infty G_\nu(\mathbf{u})$, and assume that

(C2) $t_1,\ldots,t_m$ *belong to* $G(\mathbf{u})$.

It then follows that there exists a ν such that

(C3) $t_1,\ldots,t_m$ *belong to* $G_\nu(\mathbf{u})$.

From now on, we fix a ν for which (C3) holds, and a positive $\bar{\varepsilon}_1$ such that $t_i + \bar{\varepsilon}_1 < b$ whenever $i = 1,\ldots,m$, and $[t_i, t_i + \bar{\varepsilon}_1] \cap [t_j, t_j + \bar{\varepsilon}_1] = \emptyset$ whenever $t_i \ne t_j$.

For $\varepsilon \in\,]\,0, \bar{\varepsilon}_1]$, and $\mathbf{r} = (r_1,\ldots,r_m) \in \Sigma^m$, we define disjoint measurable subsets $I_1(\mathbf{r},\varepsilon)$, …, $I_m(\mathbf{r},\varepsilon)$ of the interval $[a,b]$ by letting $I_i(\mathbf{r},\varepsilon)$ be the intersection of the interval $[t_i + (r_1 + \ldots + r_{i-1})\varepsilon, t_i + (r_1 + \ldots + r_i)\varepsilon\,[$ with the set $G_\nu(\mathbf{u})$. We emphasize that these sets are disjoint even if the t_i are not distinct.

We let

$$\begin{aligned}
\theta_i(r,\varepsilon) &= r_i - \frac{\operatorname{meas}(I_i(\mathbf{r},\varepsilon))}{\varepsilon}\,,\\
\theta(\varepsilon) &= \max\Big\{\theta_i(\mathbf{r},\varepsilon) : \mathbf{r}\in\Sigma^m, i=1,\ldots,m\Big\}\,,\\
\tilde\theta(\varepsilon) &= \frac{1}{\varepsilon}\sum_i \int_{t_i}^{t_i+\varepsilon} \|\eta_*(t)-\eta_*(t_i))\|dt\,,\\
\kappa_1 &= \max\{\|v_i\| : i=1,\ldots,m\}\,,\\
\Theta(\varepsilon) &= \theta(\varepsilon)+\kappa_1\tilde\theta(\varepsilon)\,.
\end{aligned}$$

and observe that $\Theta(\varepsilon)\to 0$ as $\varepsilon\downarrow 0$.

We then define, for every pair $(\mathbf{r},\varepsilon)\in\Sigma^m\times\,]\,0,\bar\varepsilon_1]$, a function $\eta^{\mathbf{r},\varepsilon} : [a,b]\mapsto U$, by letting

$$\eta^{\mathbf{r},\varepsilon}(t) = \begin{cases} \eta_*(t) & \text{if} \quad t\notin\cup_{i=1}^m I_i(\mathbf{r}.\varepsilon),\\ u_i & \text{if} \quad t\in I_i(\mathbf{r},\varepsilon),\ i=1,\ldots,m\,. \end{cases}$$

It is clear that each $\eta^{\mathbf{r},\varepsilon}$ is integrable. We then define $\xi^{\mathbf{r},\varepsilon}(t) = \bar q + \int_a^t \eta^{\mathbf{r},\varepsilon}(s)ds$, so the $\xi^{\mathbf{r},\varepsilon}$ are absolutely continuous. Clearly,

$$\begin{aligned}
\int_a^b \eta^{\mathbf{r},\varepsilon}(t)dt &= \int_a^b \eta_*(t)dt + \sum_i \int_{I_i(\mathbf{r},\varepsilon)} (u_i-\eta_*(t))dt\\
&= \hat q - \bar q + \varepsilon\sum_i r_i(1-\theta_i(\mathbf{r},\varepsilon))\cdot v_i\\
&\quad -\sum_i \int_{I_i(\mathbf{r},\varepsilon)} (\eta_*(t)-\eta_*(t_i))dt\\
&= \hat q - \bar q + \varepsilon\mathbf{r}\cdot\mathbf{v} + \zeta^{\mathbf{r},\varepsilon}\,,
\end{aligned}$$

where $\|\zeta^{\mathbf{r},\varepsilon}\|\le\varepsilon\Theta(\varepsilon)$. Naturally, $\xi^{\mathbf{r},\varepsilon}(b)$ is equal to $\hat q+\varepsilon\mathbf{r}\cdot\mathbf{v}+\zeta^{\mathbf{r},\varepsilon}$. So, if we choose $\mathbf{r}\in\Sigma_0^m$, then the map $\varepsilon\to\xi^{\mathbf{r},\varepsilon}$ is a "variation of ξ_* with fixed endpoints to first order." We will see below how to make a small correction to obtain an $\mathbf{r}(\varepsilon)$ such that $\xi^{\mathbf{r}(\varepsilon),\varepsilon}(b)$ is exactly equal to $\hat q$.

If $t\in[a,b]$, and we write $\rho_i = r_1+\ldots+r_i$, then the integral $\int_a^t \|\eta^{\mathbf{r},\varepsilon}(s)-\eta_*(s)\|\,ds$ is equal to $\sum_i\int_{I_i(\mathbf{r},\varepsilon)\cap[a,t]}\|u_i-\eta_*(s)\|\,ds$, which is bounded above by $\sum_i\int_{t_i+\rho_{i-1}\varepsilon}^{t_i+\rho_i\varepsilon}\|u_i-\eta_*(s)\|\,ds$. Since $\|u_i-\eta_*(s)\|$ is bounded by $\|v_i\|+\|\eta_*(t_i)-\eta_*(s)\|$, we get

$$\int_a^t \|\eta^{\mathbf{r},\varepsilon}(s)-\eta_*(s)\|\,ds \le \varepsilon(\kappa_1+\tilde\theta(\varepsilon))\,. \tag{14}$$

We now choose $\bar\varepsilon_2\in\,]\,0,\bar\varepsilon_1]$ such that $\tilde\theta(\varepsilon)\le 1$ for all $\varepsilon\in\,]\,0,\bar\varepsilon_2]$, and $(\kappa_1+1)\bar\varepsilon_2\le\bar\delta_3$. It then follows that

$$||\xi^{\mathbf{r},\varepsilon}(t)-\xi_*(t)||\le(\kappa_1+1)\varepsilon \ \text{ for all } \ t\in[a,b]\,, \tag{15}$$

whenever $\varepsilon\in\,]\,0,\bar\varepsilon_2]$. From now on, we choose $\delta=\delta(\varepsilon)=(\kappa_1+1)\varepsilon$, so that $\xi^{\mathbf{r},\varepsilon}\in\mathcal{A}^{\delta}(Q,U,L,a,b)$—and, in particular, $\xi^{\mathbf{r},\varepsilon}$ belongs to $\mathcal{A}^{\bar\delta_3}(Q,U,L,a,b)$—whenever $\mathbf{r}\in\Sigma^m$ and $0<\varepsilon\le\bar\varepsilon_2$.

We now estimate the difference $I(\xi^{\mathbf{r},\varepsilon})-I(\xi_*)$ using (7), with $\xi^{\mathbf{r},\varepsilon}$ in the role of ξ. We fix an $\omega\in\Gamma(W)$. Write $\alpha(u,t) = H(\xi_*(t),u,\psi_\omega(t),t)$, and then define $\beta_i(t) = \alpha(u_i,t)-\alpha(\eta_*(t),t)$. Then $\int_a^b\Big(\alpha(\eta^{\mathbf{r},\varepsilon}(t),t)-\alpha(\eta_*(t),t)\Big)dt$ is equal to $\sum_i\int_{I_i(\mathbf{r},\varepsilon)}\Big(\alpha(u_i,t)-\alpha(\eta_*(t),t)\Big)dt$, i. e., to $-\varepsilon\sum_i r_i(1-\theta_i(\mathbf{r},\varepsilon))z_i^\omega+\mathcal{E}_1(\varepsilon)$, which equals $-\varepsilon\mathbf{r}\cdot\mathbf{z}^\omega+\mathcal{E}_2(\varepsilon)$, where the error terms $\mathcal{E}_\ell(\varepsilon)$ are given by $\mathcal{E}_2(\varepsilon) = \mathcal{E}_1(\varepsilon)+\varepsilon\sum_i r_i\theta_i(\mathbf{r},\varepsilon))z_i^\omega$ and $\mathcal{E}_1(\varepsilon)=\sum_i\int_{I_i(\mathbf{r},\varepsilon)}\Big(\beta_i(t)-\beta_i(t_i)\Big)dt$.

It then follows from (7) that

$$I(\xi^{\mathbf{r},\varepsilon})-I(\xi_*) = \varepsilon\mathbf{r}\cdot\mathbf{z}^\omega-\mathcal{E}_2(\varepsilon)+\mathcal{R}_{\xi^{\mathbf{r},\varepsilon},\omega}\,. \tag{16}$$

We now estimate the remainder terms of (16). Clearly, $\alpha(u,t) = \psi_\omega(t).u-L(\xi_*(t),u,t)$. Let $\Lambda_i(t)=L(\xi_*(t),u_i,t)$ for $i>0$, $\Lambda_0(t)=L(\xi_*(t),\eta_*(t),t)$. For $\varepsilon>0$, define $\sigma_1(\varepsilon)$ to be the supremum of the numbers $|\Lambda_i(t)-\Lambda_i(s)|$, ranging over all $i\in\{0,\ldots,m\}$ and all $t,s\in\tilde J_\nu(\mathbf{u})$ such that $|t-s|\le\varepsilon$. Also, let $\sigma_2(\varepsilon)$ be the supremum of the integrals $\int_t^{t+\varepsilon}k(s)\,ds$, for $a\le t\le b-\varepsilon$. Let

$$\sigma_3(\varepsilon) = \sup\{\|\eta_*(t)-\eta_*(s)\| : t,s\in\tilde J_\nu(\mathbf{u}), |t-s|\le\varepsilon\}\,.$$

Then, for $\ell=1,2,3$, $\sigma_\ell(\varepsilon)\to 0$ as $\varepsilon\downarrow 0$, because the Λ_i and η_* are continuous on $\tilde J_\nu(\mathbf{u})$, and k is integrable. Moreover, if $t\in I_i(\mathbf{r},\varepsilon)$, then

$$\begin{aligned}
\beta_i(t)-\beta_i(t_i) &= \Lambda_0(t)-\Lambda_0(t_i)-(\Lambda_i(t)-\Lambda_i(t_i))\\
&\quad +(\psi_\omega(t)-\psi_\omega(t_i))\cdot v_i+\psi_\omega(t)\cdot(\eta_*(t)-\eta_*(t_i))\,.
\end{aligned}$$

so $|\beta_i(t)-\beta_i(t_i)|\le\sigma_4(\varepsilon)$, where $\kappa_2=\int_a^b k(t)\,dt$, and $\sigma_4(\varepsilon)=2\sigma_1(\varepsilon)+\kappa_1\sigma_2(\varepsilon)+\kappa_2\sigma_3(\varepsilon)$.

It then follows that $\|\mathcal{E}_1(\varepsilon)\|\le\varepsilon\sigma_4(\varepsilon)$.

On the other hand, $z_i^\omega=-\psi_\omega(t_i)\cdot v_i+\Lambda_i(t_i)-\Lambda_0(t_i)$. So, if we let $\kappa_3=\max\{\Lambda_i(t_i)-\Lambda_0(t_i) : i=1,\ldots,m\}$, we get the bound $|z_i^\omega|\le\kappa_4$, where $\kappa_4=\kappa_1\kappa_2+\kappa_3$. This implies the estimate $\|\mathcal{E}_2(\varepsilon)\|\le\varepsilon\sigma_5(\varepsilon)$, where we let $\sigma_5(\varepsilon)=\sigma_4(\varepsilon)+\kappa_4\theta(\varepsilon)$.

We now get a bound for $\mathcal{R}_{\xi^{\mathbf{r},\varepsilon},\omega}$, using (8). First, it is clear that $S_{\xi^{\mathbf{r},\varepsilon}}(t)=0$ unless t belongs to $\cup_{i=1}^m I_i(\mathbf{r},\varepsilon)$. Let $\sigma_6(\varepsilon)$ be the supremum of the numbers $|L(q,\eta_i(t),t)-L(\xi_*(t),\eta_i(t),t)|$, for $i=0,1,\ldots,m$, $t\in\tilde J_\nu(\mathbf{u})$, and $\|q-\xi_*(t)\|\le(\kappa_1+1)\varepsilon$, where $\eta_i(t)=u_i$ for $i=1,\ldots,m$, and $\eta_0(t)=\eta_*(t)$. Then the bound $\int_a^b\|S_{\xi^{\mathbf{r},\varepsilon}}(t)\|dt\le\varepsilon\sigma_6(\varepsilon)$ holds and, moreover, $\sigma_6(\varepsilon)$ goes to 0 as $\varepsilon\downarrow 0$, because the $m+1$ functions $(q,t)\to L(q,\eta_i(t),t)$ are continuous on $\mathcal{T}(\xi_*,\bar\delta_3)$. So

$$I(\xi^{\mathbf{r},\varepsilon})-I(\xi_*) = \varepsilon\mathbf{r}\cdot\mathbf{z}^\omega+\int_a^b R_{\xi^{\mathbf{r},\varepsilon},\omega}(t)dt+\mathcal{E}_3(\varepsilon)\,, \tag{17}$$

where $\mathcal{E}_3(\varepsilon)=-\mathcal{E}_2(\varepsilon)+\int_a^b S_{\xi^{\mathbf{r},\varepsilon}}(t)dt$, and then $\|\mathcal{E}_3(\varepsilon)\|\le\varepsilon\sigma_7(\varepsilon)$, where $\sigma_7(\varepsilon)=\sigma_5(\varepsilon)+\sigma_6(\varepsilon)$, so $\sigma_7(\varepsilon)\to 0$ as $\varepsilon\downarrow 0$.

Now, if we knew that $\omega\in\Gamma(W_\delta^{\xi^{\mathbf{r},\varepsilon}})$, we could apply (6) and get the estimate

$$\int_a^b \|R_{\xi^{\mathbf{r},\varepsilon},\omega}(t)\|dt\le\varepsilon\sigma_8(\varepsilon)\,, \tag{18}$$

where $\sigma_8(\varepsilon)=(\kappa+1)\int_a^b k_{(\kappa+1)\varepsilon}(t)\,dt$, so that $\sigma_8(\varepsilon)\to 0$ as $\varepsilon\downarrow 0$, thanks to (A8.c). If, in addition, we could guarantee that $\mathbf{r}\cdot\mathbf{z}^\omega<0$ then we would know, for small ε, that $I(\xi^{\mathbf{r},\varepsilon})<I(\xi_*)$. If we could achieve this for an $\mathbf{r}\in\Sigma_0^m$, then $\xi^{\mathbf{r},\varepsilon}$ would satisfy the same endpoint conditions as ξ_*, up to a small error. Then, if we could correct the error, we would obtain a $\xi\in\mathcal{A}_{\bar q,\hat q}(Q,U,L,a,b)$

for which $I(\xi) < I(\xi_*)$. Finally, if we could make sure that this ξ is in $\mathcal{N}$, we would have obtained the desired contradiction.

There is, however, no obvious reason for ω to belong to $\Gamma(W_\delta^{\xi^{\mathbf{r},\varepsilon}})$. Indeed, our hypotheses guarantee that for every ω there is an $\mathbf{r} \in \Sigma_0^m$ for which $\mathbf{r} \cdot \mathbf{z}^\omega < 0$. Then, for such an $\mathbf{r}$, and a small enough ε, $\Gamma(W_\delta^{\xi^{\mathbf{r},\varepsilon}})$ is nonempty, but this does not imply that $\omega \in \Gamma(W_\delta^{\xi^{\mathbf{r},\varepsilon}})$. To obtain an ω and an $\mathbf{r}$ with the desired properties, and to correct the error on the terminal state, we need a fixed-point argument.

First, we fix a positive radius R and a continuous map $v \to \boldsymbol{\sigma}(x) = (\sigma_1(x), \ldots, \sigma_m(x))$ from the closed ball $\mathbb{B}(R) \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : \|x\| \le R\}$ to Σ^m, such that $\sum_{i=1}^m \sigma_i(x) v_i = x$ for every $x \in \mathbb{B}(R)$. (The existence of R and $\boldsymbol{\sigma}$ follows easily from the normality assumption.)

Next, for each $\omega \in \Gamma(W)$ we choose an $r(\omega) \in \Sigma_0^m$ such that $\mathbf{r}(\omega) \cdot \mathbf{z}^\omega < 0$. Since $\Gamma(W)$ is a compact metric space—with the distance d defined by
$d(\omega_1, \omega_2) = \sup\{\|\psi_{\omega_1}(t) - \psi_{\omega_2}(t)\| : a \le t \le b\}$,
which gives rise to the topology of weak convergence—and the map $\omega \to \mathbf{z}^\omega$ is continuous, a simple partition of unity argument enables us to choose a continuous function $\Gamma(W) \ni \omega \to \mathbf{r}(\omega) \in \Sigma_0^m$ such that $\mathbf{r}(\omega) \cdot \mathbf{z}^\omega < 0$ for all $\omega \in \Gamma(W)$. Then the compactness of $\Gamma(W)$ implies that the scalar function $\Gamma(W) \ni \omega \to \mathbf{r}(\omega) \cdot \mathbf{z}^\omega$ attains its maximum value $-\bar{\mu}$, which is of course negative, so $\bar{\mu} > 0$ and $\mathbf{r}(\omega) \cdot \mathbf{z}^\omega \le -\bar{\mu}$ for all $\omega \in \Gamma(W)$. We let $\kappa_3 = \sup\{\|\mathbf{z}^\omega\| : \omega \in \Gamma(W)\}$, write $\bar{\mu} = 3\hat{\mu}$, and choose a number s such that $0 < 3s < 1$ and $\kappa_3 s \le \hat{\mu}$. We then define

$$\hat{\mathbf{r}}(\omega, x) = (1-s)\mathbf{r}(\omega) + s\boldsymbol{\sigma}(x)\,, \text{ for } \omega \in \Gamma(W),\ x \in \mathbb{B}(R)\,.$$

Then $\hat{\mathbf{r}}(\omega, x) \cdot \mathbf{z}^\omega \le -\hat{\mu}$ for all $(\omega, x) \in \Gamma(W) \times \mathbb{B}(R)$.

We now fix an $\varepsilon \in]\,0, \bar{\varepsilon}_2]$, and let $\delta = (\kappa_1 + 1)\varepsilon$. It is then clear that the map $\Sigma^m \ni \mathbf{r} \to \eta^{\mathbf{r},\varepsilon} \in L^1([a, b], \mathbb{R}^n)$ is continuous. So the map

$$\Gamma(W) \times \mathbb{B}(R) \ni (\omega, x) \to \eta^{\hat{\mathbf{r}}(\omega,x),\varepsilon} \in L^1([a, b], \mathbb{R}^n)$$

is continuous. Since $\delta \le \bar{\delta}_3$, and the arc $\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}$ belongs to $\mathcal{A}^\delta(Q, U, L, a, b)$ for every (ω, x), the set $\mathcal{U}_\varepsilon(\omega, x) \stackrel{\text{def}}{=} \Gamma\Big(W_\delta^{\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}}\Big)$ is a weakly compact, convex, nonempty subset of $L^1([a, b], \mathbb{R}^n)$. Moreover, it is easy to show that the set-valued map $\mathcal{U}_\varepsilon$ is upper semicontinuous. (Indeed, suppose $\omega_j \to \omega$, $x_j \to x$, $\zeta_j \in \mathcal{U}_\varepsilon(\omega_j, x_j)$, and $\zeta_j \to \zeta$. Then, for each j, the inequality

$$\begin{aligned} L(\xi^{\hat{\mathbf{r}}(\omega_j,x_j),\varepsilon}(t), \eta_*(t), t) &\le L(\xi_*(t), \eta_*(t), t) \\ &+\zeta_j(t) \cdot (\xi^{\hat{\mathbf{r}}(\omega_j,x),\varepsilon}(t) - \xi_*(t)) + \delta k_\delta(t) \end{aligned}$$

holds for almost every t, so we can pick a subset E_0 of $[a, b]$ of measure $b - a$ such that the inequality holds for all $t \in E_0$. Since $q \to L(q, \eta_*(t), t)$ is continuous, and the arcs $\xi^{\hat{\mathbf{r}}(\omega_j,x_j),\varepsilon}$ converge uniformly to $\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}$, we find

$$L(\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(t), \eta_*(t), t) \le L(\xi_*(t), \eta_*(t), t) + \lambda(t) \quad (19)$$

for all $t \in E_0$, where

$$\lambda(t) = \delta k_\delta(t) + \liminf_{j\to\infty} \Big(\zeta_j(t) \cdot (\xi^{\hat{\mathbf{r}}(\omega_j,x_j),\varepsilon}(t) - \xi_*(t))\Big)\,.$$

Now, for each $t \in E_0$, the sequence $\{\zeta_j(t)\}_{j=1}^\infty$ is bounded, and the vectors $\xi^{\hat{\mathbf{r}}(\omega_j,x_j),\varepsilon}(t) - \xi_*(t)$ converge to $\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(t) - \xi_*(t)$. Therefore,

$$\lambda(t) = \delta k_\delta(t) + \liminf_{j\to\infty} \Big(\zeta_j(t) \cdot (\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(t) - \xi_*(t))\Big)\,.$$

Let C_j denote the closed convex hull in L^1 of the functions ζ_k for $k \ge j$. Then C_j is weakly closed, so $\zeta \in C_j$. So, for each j, ζ is a limit in L^1, as $k \to \infty$, of convex combinations $\tilde{\zeta}_{j,k} = \sum_{\ell=j}^{j+A(j,k)} \alpha_{j,k,\ell}\zeta_\ell$. By passing to a subsequence of $\{\tilde{\zeta}_{j,k}\}_{k=1}^\infty$, we may assume that $\tilde{\zeta}_{j,k}(t) \to \zeta(t)$ as $k \to \infty$ for almost all t, i.e. that there is a subset E_j of $[a, b]$ of measure $b - a$ such that $\tilde{\zeta}_{j,k}(t) \to \zeta(t)$ as $k \to \infty$ for all $t \in E_j$. Let $E = \bigcap_{j=0}^\infty E_j$. If $t \in E$, and $\alpha > 0$, then there is a j such that $\zeta_\ell(t) \cdot (\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(t) - \xi_*(t)) \ge \lambda(t) - \delta k_\delta(t) - \alpha$ whenever $\ell \ge j$. It then follows that for all k the inequality $\tilde{\zeta}_{j,k}(t) \cdot (\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(t) - \xi_*(t)) \ge \lambda(t) - \delta k_\delta(t) - \alpha$ holds. Hence $\zeta(t) \cdot (\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(t) - \xi_*(t)) \ge \lambda(t) - \delta k_\delta(t) - \alpha$. Since α was arbitrary, we have shown that

$$\zeta(t) \cdot (\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(t) - \xi_*(t)) \ge \lambda(t) - \delta k_\delta(t)\,.$$

It then follows from (19) that

$$\begin{aligned} L(\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(t), \eta_*(t), t) &\le L(\xi_*(t), \eta_*(t), t) \\ &+\zeta(t) \cdot (\xi^{\mathbf{r}(\omega),\varepsilon}(t) - \xi_*(t)) + \delta k_\delta(t) \end{aligned}$$

for all $t \in E$. Since E has full measure, and ζ clearly belongs to $\Gamma(W)$, we conclude that $\zeta \in \mathcal{U}_\varepsilon(\omega, x)$, as desired.)

We now impose the extra condition that $\Theta(\varepsilon) \le Rs$, and observe that

$$\begin{aligned} \xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(b) &= \hat{q} + \varepsilon\hat{\mathbf{r}}(\omega, x) \cdot \mathbf{v} + \zeta^{\hat{\mathbf{r}}(\omega,x),\varepsilon} \\ &= \hat{q} + (1-s)\varepsilon\mathbf{r}(\omega) + \varepsilon s\boldsymbol{\sigma}(x) \cdot \mathbf{v} + \zeta^{\hat{\mathbf{r}}(\omega,x),\varepsilon} \\ &= \hat{q} + \varepsilon s\boldsymbol{\sigma}(x) \cdot \mathbf{v} + \zeta^{\hat{\mathbf{r}}(\omega,x),\varepsilon} \\ &= \hat{q} + \varepsilon s x + \zeta^{\hat{\mathbf{r}}(\omega,x),\varepsilon}\,, \end{aligned}$$

so

$$\Big\|\frac{1}{\varepsilon s}\Big(\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(b) - \hat{q}\Big) - x\Big\| \le \frac{\|\zeta^{\hat{\mathbf{r}}(\omega,x),\varepsilon}\|}{\varepsilon s} \le \frac{\Theta(\varepsilon)}{s} \le R\,.$$

Therefore, if we define

$$\mathcal{V}_\varepsilon(\omega, x) = \mathcal{U}_\varepsilon(\omega, x) \times \Big\{x - \frac{1}{\varepsilon s}\Big(\xi^{\hat{\mathbf{r}}(\omega,x),\varepsilon}(b) - \hat{q}\Big)\Big\}\,,$$

we see that $\mathcal{V}_\varepsilon$ is an upper semicontinuous set-valued map on $\Gamma(W) \times \mathbb{B}(R)$ whose values are nonempty compact convex subsets of $\Gamma(W) \times \mathbb{B}(R)$. So we can apply the infinite-dimensional version of the Kakutani fixed-point theorem and conclude that there exists a pair $(\omega_\varepsilon, x_\varepsilon)$ such that $\omega_\varepsilon \in \mathcal{U}_\varepsilon(\omega_\varepsilon, x_\varepsilon)$ and $x_\varepsilon - \frac{1}{\varepsilon s}\Big(\xi^{\hat{\mathbf{r}}(\omega_\varepsilon,x_\varepsilon),\varepsilon}(b) - \hat{q}\Big) = x_\varepsilon$. It then follows that $\xi^{\hat{\mathbf{r}}(\omega_\varepsilon,x_\varepsilon),\varepsilon}(b) = \hat{q}$.

Since $\omega_\varepsilon \in \Gamma\Big(W_\delta^{\xi^{\hat{\mathbf{r}}(\omega_\varepsilon,x_\varepsilon),\varepsilon}}\Big)$, we can apply (17) and (18) with $\mathbf{r} = \hat{\mathbf{r}}(\omega_\varepsilon, x_\varepsilon)$ and $\omega = \omega_\varepsilon$, and obtain

$$\begin{aligned} I(\xi^{\hat{\mathbf{r}}(\omega_\varepsilon,x_\varepsilon),\varepsilon}) - I(\xi_*) &\le \varepsilon\hat{\mathbf{r}}(\omega_\varepsilon, x_\varepsilon) \cdot \mathbf{z}^{\omega_\varepsilon} + \varepsilon\sigma_9(\varepsilon) \\ &\le \varepsilon(\sigma_9(\varepsilon) - \hat{\mu})\,, \end{aligned} \quad (20)$$

where $\sigma_9 = \sigma_7 + \sigma_8$.

It now suffices to take ε such that $\sigma_9(\varepsilon) < \hat{\mu}$, and we get $I(\xi^{\hat{\mathbf{r}}(\omega_\varepsilon,x_\varepsilon),\varepsilon}) < I(\xi_*)$. Since $I(\xi^{\hat{\mathbf{r}}(\omega_\varepsilon,x_\varepsilon),\varepsilon})$ belongs to $\mathcal{A}_{\bar{q},\hat{q}}(Q,U,L,a,b)$, we have contradicted the assumption that ξ_* was a solution of (1), provided that $I(\xi^{\hat{\mathbf{r}}(\omega_\varepsilon,x_\varepsilon),\varepsilon}) \in \mathcal{N}$. To guarantee this, we assume

(A9) *$\mathcal{N}$ is a variational neighborhood of η_*.*

Condition (A9) means that for every finite subset U_0 of U there exists an $\varepsilon > 0$ such that, if $\eta : [a,b] \to U$ is such that $\eta(t) \in \{\eta_*(t)\} \cup U_0$ for every $t \in [a,b]$, $\{t : \eta(t) = u\}$ is measurable for every $u \in U_0$, and $\text{meas}\Big(\{t : \eta(t) \neq \eta_*(t)\}\Big) < \varepsilon$, then $\eta \in \mathcal{N}$.

Under assumptions (A1)-(A9), we have shown that, if ξ_* is a solution of our minimization problem, then for every m and every $\mathbf{u} = (u_1, \ldots, u_m) \in U^m$ there exists a subset $G(\mathbf{u})$ of $[a,b]$ of full measure, such that for every m-tuple $\mathbf{t} = (t_1, \ldots, t_m)$ of members of $G(\mathbf{u})$ there is a nontrivial pair (ψ, ψ_0) that satisfies (9) and (10). Moreover, it is clear from the definition of $G(\mathbf{u})$ that $G(\mathbf{u})$ only depends on the set $V = \{u_1, \ldots, u_m\}$, so we can define $G(V)$ to be equal to $G(\mathbf{u})$ is V is any finite subset of U and $V = \{u_1, \ldots, u_m\}$. Then what we have shown implies that, if V is any finite subset of U, and T is a finite subset of $G(V)$, there is a nontrivial pair (ψ, ψ_0) that satisfies (9), $\psi_0 \geq 0$, and

$$\mathcal{H}^{\psi_0}_{\xi_*,\psi}(u,t) \leq \mathcal{H}^{\psi_0}_{\xi_*,\psi}(\eta_*(t),t) \quad \text{for} \quad t \in T\,. \tag{21}$$

where we let $\mathcal{H}^{\psi_0}_{\xi_*,\psi}(u,t) \stackrel{\text{def}}{=} H^{\psi_0}(\xi_*(t), u, \psi(t), t)$. Let $\Psi(V,T)$ be the set of all pairs (ψ, ψ_0) that satisfy (9), (21), $\psi_0 \geq 0$, and $\|\psi(b)\| + \psi_0 = 1$. Then, if $V \subseteq U$ is finite, and $T \subseteq G(V)$ is finite, the set $\Psi(V,T)$ is compact in $C^0([a,b], \mathbb{R}^n) \times \mathbb{R}$, and we have shown that it is nonempty. Now, let $\mathbf{T}$ be the set of all finite subsets of $G(V)$. Then if $T_1 \in \mathbf{T}, \ldots, T_k \in \mathbf{T}$,

$$\Psi(V,T_1) \cap \ldots \cap \Psi(V,T_k) = \Psi(V, T_1 \cup \ldots \cup T_k) \neq \emptyset\,.$$

So the family $\{\Psi(V,T)\}_{T \in \mathbf{T}}$ has the finite intersection property. Hence $\bigcap_{T\in\mathbf{T}} \Psi(V,T) \neq \emptyset$. That is, there exists, for each finite subset V of U, a pair (ψ, ψ_0) that satisfies (9), $\psi_0 \geq 0$, the normalization condition $\|\psi(b)\| + \psi_0 = 1$, and

$$\mathcal{H}^{\psi_0}_{\xi_*,\psi}(u,t) \leq \mathcal{H}^{\psi_0}_{\xi_*,\psi}(\eta_*(t),t) \quad \text{for } t \in G(V)\,,\ u \in V. \tag{22}$$

Now, let $\mathbf{V}$ be the set of all finite subsets of U. For each $V \in \mathbf{V}$, let $\hat{\Psi}(V)$ be the set of all pairs (ψ, ψ_0) that satisfy (9), $\psi_0 \geq 0$, the normalization condition $\|\psi(b)\| + \psi_0 = 1$, and

$$\mathcal{H}^{\psi_0}_{\xi_*,\psi}(u,t) \leq \mathcal{H}^{\psi_0}_{\xi_*,\psi}(\eta_*(t),t) \text{ for a.e. } t \in [a,b]\,,\text{ if } u \in V.$$

We have shown that $\hat{\Psi}(V) \neq \emptyset$ for every $V \in \mathbf{V}$. Moreover, $\hat{\Psi}(V)$ is clearly compact in $C^0([a,b], \mathbb{R}^n) \times \mathbb{R}$. If $V_1 \in \mathbf{V}, \ldots, V_k \in \mathbf{V}$, then

$$\hat{\Psi}(V_1) \cap \ldots \cap \hat{\Psi}(V_k) = \hat{\Psi}(V_1 \cup \ldots \cup V_k) \neq \emptyset\,.$$

So the family $\{\hat{\Psi}(V)\}_{V \in \mathbf{V}}$ has the finite intersection property. Hence $\bigcap_{V\in\mathbf{V}} \hat{\Psi}(V) \neq \emptyset$. That is, there exists a pair (ψ, ψ_0) that satisfies (9), $\psi_0 \geq 0$, and the normalization condition $\|\psi(b)\| + \psi_0 = 1$, and is such that for every $u \in U$ the inequality

$$H^{\psi_0}(\xi_*(t), u, \psi(t), t) \leq H^{\psi_0}(\xi_*(t), \eta_*(t), \psi(t), t) \tag{23}$$

holds for almost all $t \in [a,b]$.

Finally, if we assume

(A10) *the map $U \ni u \to L(\xi_*(t), u, t)$ is continuous for every $t \in [a,b]$,*

then we can use the fact that U is separable, and conclude that, if (ψ, ψ_0) satisfies the above conclusions, then

$$\max\Big\{H^{\psi_0}(\xi_*(t), u, \psi(t), t) : u \in U\Big\} = H^{\psi_0}(\xi_*(t), \eta_*(t), \psi(t), t) \tag{24}$$

for almost all $t \in [a,b]$. So we have proved:

Theorem 1 *Assume that conditions (A1)-(A9) hold, and (ξ_*, η_*) is a solution of (1). Then there exists a nontrivial pair (ψ, ψ_0) that satisfies (9), $\psi_0 \geq 0$, and is such that for every $u \in U$ (23) holds for almost all $t \in [a,b]$. If, in addition, (A10) holds, then (24) is satisfied for almost all $t \in [a,b]$.* ♢

References

[1] Berkovitz, L. D., *Optimal Control Theory.* Springer-Verlag, New York, 1974.

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

[3] Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko, *The Mathematical Theory of Optimal Processes.* Wiley, New York, 1962.

[4] Sussmann, H. J., *A strong version of the Maximum Principle under weak hypotheses.* In *Proc. 33rd IEEE Conference on Decision and Control, Orlando, FL, 1994*, IEEE publications, New York, 1994, pp. 1950–1956.

[5] Sussmann, H. J., *A strong version of the Lojasiewicz Maximum Principle.* In *Optimal Control of Differential Equations,* N. H. Pavel Ed., Marcel Dekker, New York, 1994, pp. 293–309.

[6] Sussmann, H. J., *Some recent results on the maximum principle of optimal control theory.* In *Systems and Control in the Twenty-First Century,* C. I. Byrnes, B. N. Datta, D. S. Gilliam & C. F. Martin Eds., Birkhäuser, Boston, 1997, pp. 351–372.

[7] Sussmann, H. J., *A strong maximum principle for systems of differential inclusions.* In *Proc. 35th IEEE Conference on Decision and Control, Kobe, Japan, Dec. 1996*, IEEE publications, New York, 1996, pp. 1809–1814.

[8] Sussmann, H. J., *Multidifferential calculus: chain rule, open mapping and transversal intersection theorems.* In *Optimal Control: Theory, Algorithms, and Applications.* W. W. Hager & P. M. Pardalos Eds., Kluwer Acad. Publishers, 1998, pp. 436–487.

[9] Sussmann, H. J., *Transversality conditions and a strong maximum principle for systems of differential inclusions.* In *Proc. 37th IEEE Conference on Decision and Control, Tampa, FL, Dec. 1998,* IEEE publications, New York, 1998, pp. 1–6.