

A Local Second- and Third-order Maximum Principle

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@math.rutgers.edu

Web site: <http://www.math.rutgers.edu/~sussmann>

1. Introduction

In this note we announce preliminary results obtained by applying the methodology of Sussmann [4, 5, 6, 7, 8] to the problem studied by Ledzewicz and Schättler (cf. [1, 2]), of deriving high-order necessary conditions for a minimum in optimal control theory, extending the classical Pontryagin Maximum Principle (abbr. MP) of [3]. The work of [1, 2] uses high-order generalizations of the theorems of Lyusternik and Avakov. We pursue a different approach, constructing needle variations and, at the crucial point where a “topological argument” is needed, applying the Brouwer fixed point theorem. The result is a high-order version of the MP that contains and extends the results of [1, 2]. In particular, our version makes it clear that new multipliers occur for the first time for the third-order principle. It turns out that what is called a “second-order MP” in [1, 2] is naturally a “third-order MP” in our setting. Our third-order MP contains new multipliers exactly as the result of [2] does, but in addition it also contains the classical multipliers, to which the new multipliers are coupled in a precise way, described in Theorem 6.1.

We present an abstract necessary condition for set separation (Theorem 4.1), a necessary condition for a minimum in an abstract setting (Theorem 5.1), and, finally, the optimal control result (Theorem 6.1). We also present an example showing that our results apply in cases where those of [1, 2] do not.

For lack of space, we omit the proofs, we limit our discussion to problems with fixed endpoints, and we only consider the second and third-order cases.

2. Preliminary definitions

We assume in this section that \mathcal{X}, \mathcal{Y} are normed spaces, $\mathcal{S} \subseteq \mathcal{X}$, $x_* \in \mathcal{S}$, and \mathcal{F} is a map from \mathcal{S} to \mathcal{Y} .

2.1. High-order differentiability. Let $k \in \mathbb{N}$. We say that \mathcal{F} is of class D^k at x_* in the direction of \mathcal{S} if

1. \mathcal{F} is continuous on $N \cap \mathcal{S}$ for some neighborhood N of x_* in \mathcal{X} ,
2. there exists a continuous polynomial map $\mathcal{P} : \mathcal{X} \rightarrow \mathcal{Y}$ of degree $\leq k$ without constant term such that

$$\lim_{x \rightarrow x_*, x \in \mathcal{S}, x \neq x_*} \frac{\mathcal{F}(x) - \mathcal{F}(x_*) - \mathcal{P}(x - x_*)}{\|x - x_*\|^k} = 0. \quad \diamond$$

2.2. Variations and variational directions.

A *variation of x_* in \mathcal{S}* is a continuous curve $[0, \bar{\varepsilon}_\eta] \ni \varepsilon \mapsto \eta(\varepsilon) \in \mathcal{S}$, defined on the interval $[0, \bar{\varepsilon}_\eta]$ for some positive $\bar{\varepsilon}_\eta$, such that $\eta(0) = x_*$. We use $\text{Var}(x_*, \mathcal{S})$ to denote the set of all variations of x_* in \mathcal{S} .

Let $k \in \mathbb{N}$. A *k th-order variational direction of \mathcal{S} at x_** is a k -tuple $(v_1, \dots, v_k) \in \mathcal{X}^k$ for which there exists

a variation $\eta \in \text{Var}(x_*, \mathcal{S})$ such that

$$\eta(\varepsilon) = x_* + \varepsilon v_1 + \dots + \varepsilon^k v_k + o(\varepsilon^k)$$

as $\varepsilon \downarrow 0$. We use $\text{var}_k(x_*, \mathcal{S})$ to denote the set of all k th-order variational directions of \mathcal{S} at x_* .

A *limiting first-order variational direction of \mathcal{S} at x_** is a $v \in \mathcal{X}$ which is the limit of a sequence $\{v_j\}_{j \in \mathbb{N}}$ of first-order variational directions of \mathcal{S} at x_* . We use $\bar{\text{var}}_1(x_*, \mathcal{S})$ to denote the set of all limiting first-order variational directions of \mathcal{S} at x_* .

If $v_* \in \mathcal{X}$, then $\text{var}_{2,1}(x_*, v_*, \mathcal{S})$, $\text{var}_{3,1}(x_*, v_*, \mathcal{S})$ will denote, respectively, the set of all $w \in \mathcal{X}$ such that $(v_*, w) \in \text{var}_2(x_*, \mathcal{S})$, and the set of all pairs $(w, z) \in \mathcal{X} \times \mathcal{X}$ such that $(v_*, w, z) \in \text{var}_3(x_*, \mathcal{S})$. If $(v_*, w_*) \in \mathcal{X} \times \mathcal{X}$, we use $\text{var}_{3,2}(x_*, v_*, w_*, \mathcal{S})$ to denote the set of all $z \in \mathcal{X}$ such that $(v_*, w_*, z) \in \text{var}_3(x_*, \mathcal{S})$.

2.3. Abundant subsets. We say that a subset \mathcal{S}_0 of \mathcal{S} is $(\mathcal{S}, \mathcal{F})$ -*abundant* if there exists a sequence $\{\sigma_j\}_{j \in \mathbb{N}}$ of maps $\sigma_j : \mathcal{S} \rightarrow \mathcal{S}_0$ such that

1. the map $\mathcal{F} \circ \sigma_j$ is continuous for each j ,
2. $\mathcal{F} \circ \sigma_j$ converges to \mathcal{F} uniformly on compact subsets of \mathcal{S} as $j \rightarrow \infty$.

2.4. Feasible sets and 2-regularity. Following [1], we define the notion of a “second-order feasible set,” and of “2-regularity,” as follows.

Definition 2.1 The *second-order feasible set of \mathcal{S} in \mathcal{X} at x_* in the direction of v_** is the set $FS_{\mathcal{X}}^{(2)}(\mathcal{S}; x_*, v_*)$ of all $w \in \mathcal{X}$ for which there exist a neighborhood W of w in \mathcal{X} and an $\bar{\varepsilon} \in \mathbb{R}$ such that $\bar{\varepsilon} > 0$ and $x_* + \varepsilon v_* + \varepsilon^2 W \subseteq \mathcal{S}$ for all $\varepsilon \in [0, \bar{\varepsilon}]$. \diamond

Definition 2.2 Assume that \mathcal{F} is of class D^2 at x_* in the direction of \mathcal{S} . Let P_1, P_2 be, respectively, a continuous linear map from \mathcal{X} to \mathcal{Y} and a continuous bilinear symmetric map from \mathcal{X}^2 to \mathcal{Y} , such that $\mathcal{F}(x_* + v) = \mathcal{F}(x_*) + P_1(v) + \frac{1}{2}P_2(v, v) + o(\|v\|^2)$ as $v \rightarrow 0$, $x_* + v \in \mathcal{S}$. Let π be the canonical projection from \mathcal{Y} to $\mathcal{Y}/P_1(\mathcal{X})$. We say that \mathcal{F} is *2-regular at $(x_*, v_*, \mathcal{X}, \mathcal{S})$* if (a) $FS_{\mathcal{X}}^{(2)}(\mathcal{S}; x_*, v_*) \neq \emptyset$, (b) the linear map $\mathcal{X} \ni u \mapsto (P_1 \cdot u, \pi_v(P_2(v_*, u))) \in P_1(\mathcal{X}) \times (\mathcal{Y}/P_1(\mathcal{X}))$ is onto, and (c) $P_2(v_*, v_*) \in P_1(\mathcal{X})$. \diamond

3. The basic assumptions

We let ν be a positive integer, which in the results to be stated below will be equal to 1, 2, or 3. If k is a positive integer, and v is any object, then we use $v^{<k>}$ to denote the k -tuple consisting of v repeated k times.

For the separation theorem, we will assume that

*Supported in part by NSF Grant DMS98-03411-00798.

- (A1) \mathcal{U}^a (the “ambient source space”) and Y (the “target space”) are normed spaces, Y is finite-dimensional, \mathcal{U} is a convex subset of \mathcal{U}^a , $u_* \in \mathcal{U}$, $\mathcal{F} : \mathcal{U} \mapsto Y$ is a map, $\omega \in Y$ and $\omega \neq 0$.
- (A2) \mathcal{A} is an $(\mathcal{U}, \mathcal{F})$ -abundant subset of \mathcal{U} such that $u_* \in \mathcal{A}$.
- (A3) \mathcal{F} is of class D^ν at u_* ,
- (A4) $\{P_k\}_{k=1}^\nu$ is a family such that, (a) for each k , P_k is a continuous k -linear symmetric map from $(\mathcal{U}^a)^k$ to Y , and (b) as $v \rightarrow 0$, $u_* + v \in \mathcal{U}$, we have
- $$\mathcal{F}(u_* + v) = \mathcal{F}(u_*) + \sum_{k=1}^\nu \frac{1}{k!} P_k(v^{<k>}) + o(\|v\|^\nu).$$

We define

$$y_* \stackrel{\text{def}}{=} \mathcal{F}(u_*), \quad (1)$$

$$\Omega \stackrel{\text{def}}{=} \{y_* + t\omega : t \in \mathbb{R}, t \geq 0\}, \quad (2)$$

$$\Omega_0 \stackrel{\text{def}}{=} \{t\omega : t \geq 0\}, \quad (3)$$

so that $y_* \in \mathcal{F}(\mathcal{A}) \cap \Omega$ and $\Omega = y_* + \Omega_0 = \{y_* + z : z \in \Omega_0\}$.

We will look for necessary conditions for the following separation property:

$$(\text{SEP}) \quad \mathcal{F}(\mathcal{A}) \cap \Omega = \{y_*\}.$$

We write

$$\mathcal{V} = \text{var}_1(u_*, \mathcal{U}). \quad (4)$$

If $\nu > 1$, we will consider, in addition to \mathcal{V} , a finite set $\hat{\mathcal{V}}$ of “independent first-order variational directions with higher-order effects.” This means, to begin with, that

$$(A5) \quad \hat{\mathcal{V}} \text{ is a finite subset of } \mathcal{V}.$$

We will need additional assumptions on $\hat{\mathcal{V}}$. The first one is the “independence condition,” whose purpose is to enable us to add variations

$$\varepsilon \mapsto \eta_v(\varepsilon) = u_* + \tilde{\eta}_v(\varepsilon) = u_* + \varepsilon v + o(\varepsilon)$$

of u_* in \mathcal{U} corresponding to different members v of $\hat{\mathcal{V}}$, and still obtain a variation of u_* in \mathcal{U} . We will guarantee this by requiring that the values $\tilde{\eta}_v(\varepsilon)$ belong to sets \mathcal{U}_v^0 such that $u_* + \sum_\ell \tilde{\eta}_v(\varepsilon) \in \mathcal{U}$, from which it will follow that $u_* + \sum_v \tilde{\eta}_v(\varepsilon) \in \mathcal{U}$. The sets \mathcal{U}_v^0 will actually play a role in the statement of our main separation theorems. Therefore, rather than merely assume that they exist, we will actually regard them as additional data for our separation problem. If we write $\mathcal{U}_v = \mathcal{U}_v^0 + u_*$, the “independence condition” becomes:

- (A6) $\{\mathcal{U}_v\}_{v \in \hat{\mathcal{V}}}$ is a family of convex subsets of \mathcal{U}^a such that
1. $u_* \in \mathcal{U}_v$ for every $v \in \hat{\mathcal{V}}$,
 2. $u_* + \sum_{v \in \hat{\mathcal{V}}} (\mathcal{U}_v - u_*) \subseteq \mathcal{U}$,
 3. $v \in \text{var}_1(u_*, \mathcal{U}_v)$ for each $v \in \hat{\mathcal{V}}$.

In addition to (A5) and (A6), we will impose—if $\nu > 1$ —the “cancellation and cross-term conditions”

$$(A7) \quad \text{the vector } P_1 \cdot v \text{ belongs to } \Omega_0 \text{ whenever } v \in \hat{\mathcal{V}},$$

$$(A8) \quad \text{whenever } v, v' \in \hat{\mathcal{V}} \text{ are such that } v \neq v', \text{ the vector } P_2(v, v') \text{ belongs to } \Omega_0.$$

When $\nu = 3$, we will need the following conditions:

- (A3.9.a) $\hat{\mathcal{V}}^2$ and $\hat{\mathcal{V}}^3$ are subsets of \mathcal{V} , $\hat{\mathcal{V}}^2 \cup \hat{\mathcal{V}}^3 = \mathcal{V}$ and $\hat{\mathcal{V}}^2 \cap \hat{\mathcal{V}}^3 = \emptyset$;
- (A3.9.b) whenever $v, v', v'' \in \hat{\mathcal{V}}^3$ and $v \neq v' \neq v'' \neq v$, it follows that $P_3(v, v', v'') \in \Omega_0$;
- (A3.9.c) whenever $v, v' \in \hat{\mathcal{V}}^3$, $v \neq v'$, w belongs to $\text{var}_{2,1}(u_*, v, \mathcal{U}_v)$, and $\frac{1}{2}P_2(v, v) + P_1 \cdot w \in \Omega_0$, it follows that $\frac{1}{2}P_3(v, v, v') + P_2(w, v') \in \Omega_0$.

If $\nu = 3$, then for each $v \in \hat{\mathcal{V}}^3$ and each subset S of $\text{var}_{3,1}(u_*, v, \mathcal{U}_v)$ we let $Y_{1,v,S}$ be the linear span of ω and all the vectors $\frac{1}{2}P_2(v, v) + P_1 \cdot w$, for all $(w, z) \in S$, and define $Y_{1,v} \stackrel{\text{def}}{=} Y_{1,v, \text{var}_{3,1}(u_*, v, \mathcal{U}_v)}$.

For the abstract minimization theorem, we will assume that

- (M1) \mathcal{U}^a and X are normed spaces, X is finite-dimensional, \mathcal{U} is a convex subset of \mathcal{U}^a , $u_* \in \mathcal{U}$, $\mathcal{E} : \mathcal{U} \mapsto X$ is a map, and $\mathcal{I} : \mathcal{U} \mapsto \mathbb{R}$ is a function,
- (M2) \mathcal{A} is an $(\mathcal{U}, \mathcal{F})$ -abundant subset of \mathcal{U} such that $u_* \in \mathcal{A}$, where \mathcal{F} is the map

$$\mathcal{U} \ni u \mapsto \mathcal{F}(u) \stackrel{\text{def}}{=} (\mathcal{E}(u), \mathcal{I}(u)) \in Y \stackrel{\text{def}}{=} X \times \mathbb{R}, \quad (5)$$

- (M3) \mathcal{E} and \mathcal{I} are of class D^ν at u_* ,

- (M4) for each $k \in \{1, \dots, \nu\}$, P_k and p_k are continuous k -linear symmetric maps from $(\mathcal{U}^a)^k$ to X , \mathbb{R} , respectively, and the formulas

$$\mathcal{E}(u_* + v) = \mathcal{E}(u_*) + \sum_{k=1}^\nu \frac{1}{k!} P_k(v^{<k>}) + o(\|v\|^\nu),$$

$$\mathcal{I}(u_* + v) = \mathcal{I}(u_*) + \sum_{k=1}^\nu \frac{1}{k!} p_k(v^{<k>}) + o(\|v\|^\nu)$$

$$\text{hold as } v \rightarrow 0, u_* + v \in \mathcal{U}.$$

If $\nu = 2$ or $\nu = 3$, we will also assume that conditions (A5) and (A6) hold, as well as the following *second-order cancellation and cross-term conditions*:

- (M7) $P_1 \cdot v = 0$ and $p_1 \cdot v \leq 0$ whenever $v \in \hat{\mathcal{V}}$,
- (M8) $P_2(v_1, v_2) = 0$ and $p_2(v_1, v_2) \leq 0$ whenever $v_1, v_2 \in \hat{\mathcal{V}}$ are such that $v_1 \neq v_2$,

If $\nu = 3$, we will assume the following *third-order cancellation and cross-term conditions*:

- (M3.9.a) $\hat{\mathcal{V}}^2 \cap \hat{\mathcal{V}}^3 = \emptyset$, and $\hat{\mathcal{V}} = \hat{\mathcal{V}}^2 \cup \hat{\mathcal{V}}^3$,
- (M3.9.b) Whenever $v, v', v'' \in \hat{\mathcal{V}}^3$ and $v \neq v' \neq v'' \neq v$, it follows that $P_3(v, v', v'') = 0$ and $p_3(v, v', v'') \leq 0$.
- (M3.9.c) Whenever $v, v' \in \hat{\mathcal{V}}^3$, $v \neq v'$, and w is such that $w \in \text{var}_{2,1}(u_*, v, \mathcal{U}_v)$, $\frac{1}{2}P_2(v, v) + P_1 \cdot w = 0$ and $\frac{1}{2}p_2(v, v) + p_1 \cdot w \leq 0$, it follows that $\frac{1}{2}P_3(v, v, v') + P_2(w, v') = 0$ and also that $\frac{1}{2}p_3(v, v, v') + p_2(w, v') = 0$.

We let

$$x_* = \mathcal{E}(u_*). \quad (6)$$

We will seek necessary conditions for the *minimization condition*:

- (MIN) $\mathcal{I}(u_*) \leq \mathcal{I}(u)$ for all $u \in \mathcal{A}$ such that $\mathcal{E}(u) = x_*$.

For the optimal control result, we will consider a control system

$$\dot{x} = f(x, u, t), \quad x \in \mathbb{R}^n, \quad u \in U, \quad a \leq t \leq b, \quad (7)$$

and a Lagrangian

$$\mathbb{R}^n \times U \times [a, b] \ni (x, u, t) \mapsto L(x, u, t) \in \mathbb{R}. \quad (8)$$

We let $F(x, u, t) \stackrel{\text{def}}{=} (f(x, u, t), L(x, u, t))$, and assume:

- (C1) $m \in \mathbb{N}$ and U is a closed convex subset of \mathbb{R}^m ,
- (C2) the partial map $\mathbb{R}^n \times U \ni (x, u) \mapsto F(x, u, t) \in \mathbb{R}^{n+1}$ is of class C^ν for every $t \in [a, b]$,
- (C3) the partial map $[a, b] \ni t \mapsto F(x, u, t) \in \mathbb{R}^{n+1}$ is measurable for every $(x, u) \in \mathbb{R}^n \times U$,
- (C4) the maps of condition (C2) are bounded, together with all their partial derivatives with respect to (x, u) of order $\leq \nu$, on every compact subset of the product $\mathbb{R}^n \times U \times [a, b]$.

We let $\mathcal{U}^a = L^\infty([a, b], \mathbb{R}^m)$, and let \mathcal{U} be the set of those controls $u(\cdot) \in \mathcal{U}^a$ that are U -valued.

A *trajectory-control pair* is a pair $(x(\cdot), u(\cdot))$ consisting of an absolutely continuous curve $x(\cdot) : [a, b] \mapsto \mathbb{R}^n$ and a control $u(\cdot) : [a, b] \mapsto U$ with the property that $\dot{x}(t) = f(x(t), u(t), t)$ for almost all $t \in [a, b]$.

We fix an initial state \bar{x} and a terminal state \hat{x} , and let $TCP(\bar{x}, \hat{x})$ denote the set of all trajectory-control pairs $(x(\cdot), u(\cdot))$ such that $x(a) = \bar{x}$ and $x(b) = \hat{x}$.

We assume that we are given a pair $(x_*(\cdot), u_*(\cdot))$ (the “reference trajectory-control pair”) such that

- (C5) $(x_*(\cdot), u_*(\cdot)) \in TCP(\bar{x}, \hat{x})$, and $(x_*(\cdot), u_*(\cdot))$ minimizes the functional

$$TCP(\bar{x}, \hat{x}) \ni (x(\cdot), u(\cdot)) \mapsto J(x(\cdot), u(\cdot)) \in \mathbb{R},$$

$$\text{where } J(x(\cdot), u(\cdot)) \stackrel{\text{def}}{=} \int_a^b L(x(t), u(t), t) dt.$$

We write f_x, f_u, L_x, L_u for $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial u}, \frac{\partial L}{\partial x}, \frac{\partial L}{\partial u}$, and similarly for higher order derivatives. Also, we write

$$f_x^*(t) \stackrel{\text{def}}{=} f_x(x_*(t), u_*(t), t), \quad f_u^*(t) \stackrel{\text{def}}{=} f_u(x_*(t), u_*(t), t),$$

$$L_x^*(t) \stackrel{\text{def}}{=} L_x(x_*(t), u_*(t), t), \quad L_u^*(t) \stackrel{\text{def}}{=} L_u(x_*(t), u_*(t), t),$$

and use a similar notation for higher order derivatives, so that, for example, $f_{xxu}^*(t)$ is, for each time $t \in [a, b]$, an \mathbb{R}^n -valued trilinear form on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$ and, if $\varphi_1, \varphi_2, \theta$ are bounded measurable functions on $[a, b]$ with values in $\mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^m$, respectively, then $f_{xxu}^*(\varphi_1, \varphi_2, \theta)$ denotes the function $[a, b] \ni t \mapsto f_{xxu}^*(t)(\varphi_1(t), \varphi_2(t), \theta(t))$.

We also assume that

- (C6) \mathcal{H} is a finite set of pairs (h, v) such that $v \in \mathcal{U}^a$, h is an absolutely continuous map from $[a, b]$ to \mathbb{R}^n , $h(a) = h(b) = 0$,

$$\dot{h}(t) = \frac{\partial f}{\partial x}(x_*(t), u_*(t), t) \cdot h(t) + \frac{\partial f}{\partial u}(x_*(t), u_*(t), t) \cdot v(t)$$

for a.e. $t \in [a, b]$. and

$$\int_a^b \left(L_x^*(t) \cdot h(t) + L_u^*(t) \cdot v(t) \right) dt \leq 0. \quad (9)$$

Furthermore, we assume that

- (C7) $\{I_{h,v}\}_{(h,v) \in \mathcal{H}}$ is a family of pairwise disjoint subintervals of $[a, b]$ such that

$$\{t \in [a, b] : v(t) \neq 0\} \subseteq I_{h,v} \text{ for all } (h, v) \in \mathcal{H}. \quad (10)$$

We remark that conditions (C6) and (10) imply that

$$\{t \in [a, b] : h(t) \neq 0\} \subseteq I_{h,v} \text{ for all } (h, v) \in \mathcal{H}. \quad (11)$$

We define the *input-to-trajectory map* $\mathcal{T} : \mathcal{U} \mapsto C^0([a, b], \mathbb{R}^n)$ by first defining \mathcal{T} as a map from $\tilde{\mathcal{U}}$ to $C^0([a, b], \mathbb{R}^n)$, where $\tilde{\mathcal{U}}$ is the set of all controls $u \in \mathcal{U}$ such that the initial value problem $\dot{x} = f(x, u(t), t)$, $x(a) = \bar{x}$, has a solution (which, of course, is unique) defined on $[a, b]$. We then extend \mathcal{T} to all of \mathcal{U} by defining $\mathcal{T}(u)$ to be the curve x_* if $u \in \mathcal{U} \setminus \tilde{\mathcal{U}}$. Naturally, with this definition \mathcal{T} need not be continuous on \mathcal{U} , but it is continuous (and, in fact, ν times continuously Fréchet differentiable) on $\tilde{\mathcal{U}}$, which is a relative neighborhood of u_* in \mathcal{U} , and this more than suffices to make our abstract theorems applicable.

We then define the *terminal point map* \mathcal{E} by letting $\mathcal{E}(u) = \mathcal{T}(u)(b)$ for $u \in \mathcal{U}$.

For each $(h, v) \in \mathcal{H}$, we define \mathcal{U}_v^a to be the space $L^\infty(I_{h,v}, \mathbb{R}^m)$, so \mathcal{U}_v^a can be identified with a subspace of \mathcal{U}^a in an obvious way. Also, we let \mathcal{U}_v be the subset of \mathcal{U} consisting of all functions $u \in \mathcal{U}$ such that $u(t) = u_*(t)$ for $t \notin I_{h,v}$. Finally, we let $\mathcal{U}_v^0 = \{u - u_* : u \in \mathcal{U}_v\}$, so \mathcal{U}_v^0 is a convex subset of \mathcal{U}_v^a and $\mathcal{U}_v = \mathcal{U}_v^0 + u_*$.

We define the *restricted terminal map* $\mathcal{E}_v^0 : \mathcal{U}_v^0 \mapsto \mathbb{R}^n$ to be the map given by

$$\mathcal{E}_v^0(u) = \mathcal{E}(u + u_*) \quad \text{for } u \in \mathcal{U}_v.$$

4. The separation theorem

We define C_1 to be the smallest closed convex cone in Y that contains all the vectors $P_1 \cdot v$, $v \in \mathcal{V}$.

If $\nu > 1$, we let C_2 be the smallest closed convex cone in Y that contains $C_1(\mathcal{V})$ as well as all the vectors $\frac{1}{2}P_2(v, v) + P_1 \cdot w$, for all (v, w) such that $v \in \hat{\mathcal{V}}$ and $(v, w) \in \text{var}_2(u_*, \mathcal{U}_v)$.

If $\nu = 3$, we define

$$\tilde{P}(v, w) \stackrel{\text{def}}{=} \frac{1}{2}P_2(v, v) + P_1 \cdot w, \quad (12)$$

$$\hat{P}(v, w, z) \stackrel{\text{def}}{=} \frac{1}{6}P_3(v, v, v) + P_2(v, w) + P_1 \cdot z. \quad (13)$$

and let C_3 be the smallest closed convex cone in Y that contains C_2 as well as all the vectors y of the form $y = \hat{P}(v, w, z)$ for some (v, w, z) such that $v \in \hat{\mathcal{V}}^3$, $(v, w, z) \in \text{var}_3(u_*, \mathcal{U}_v)$, and $\hat{P}(v, w) \in \Omega_0$.

For $\nu = 1, 2, 3$, the ν -th-order surjectivity condition says:

(SUR- ν) ω is an interior point of the cone C_ν .

The separation theorem can now be stated.

Theorem 4.1 Assume that $\nu \in \{1, 2, 3\}$, and the data $\mathcal{U}^a, Y, \mathcal{U}, u_*, \mathcal{F}, \omega, \mathcal{A}, P_1, \dots, P_\nu$ satisfy Assumptions (A1), (A2), (A3), (A4). Let $y_*, \Omega, \mathcal{V}, \Omega_0$ be defined by (1), (2), (4), (3). If $\nu = 2$ or $\nu = 3$, assume in addition that $\hat{\mathcal{V}}, \{\mathcal{U}_v\}_{v \in \hat{\mathcal{V}}}$ are such that (A5), (A6), (A7) and (A8) hold. If $\nu = 3$, assume that (A3.9.a), (A3.9.b) and (A3.9.c) are satisfied. Then the ν th-order surjectivity condition (SUR- ν) implies that the separation property (SEP) does not hold. \diamond

Theorem 4.1 implies the following “multiplier” result:

Theorem 4.2 Assume that the hypotheses of Theorem 4.1 are satisfied and (SEP) holds. Then there exists a nonzero covector $\bar{\lambda} \in Y^*$, such that

- (1) $\langle \bar{\lambda}, \omega \rangle \leq 0$,
- (2) $\langle \bar{\lambda}, P_1 \cdot v \rangle \geq 0$ for all $v \in \overline{\text{var}}_1(u_*, \mathcal{U})$,
- (3) if $\nu > 1$ then $\langle \bar{\lambda}, \tilde{P}(v, w) \rangle \geq 0$ whenever $v \in \hat{\mathcal{V}}$ and $w \in \text{var}_{2,1}(u_*, v, \mathcal{U}_v)$,
- (4) if $\nu = 3$ then for every $v \in \hat{\mathcal{V}}^3$

$$(4.\&) \quad \langle \bar{\lambda}, \hat{P}(v, w, z) \rangle \geq 0 \text{ whenever } w, z \text{ are such that } (w, z) \in \text{var}_{3,1}(u_*, \mathcal{U}_v) \text{ and } \tilde{P}(v, w) \in \Omega_0.$$

In addition, if $\nu > 1$, then Condition (3) implies, in particular, that

$$(3.\#) \quad \langle \bar{\lambda}, P_1 \cdot w \rangle \geq 0 \text{ whenever } v \in \hat{\mathcal{V}} \text{ and } (v, w) \in \text{var}_2(u_*, \mathcal{U}_v).$$

Furthermore, if $\nu = 3$, then

- For each $v \in \hat{\mathcal{V}}^3$, Condition (4.&) is equivalent to the following multiplier statement:

$$(4.\#) \text{ for every finite subset } S \text{ of } \text{var}_{3,1}(u_*, v, \mathcal{U}_v) \text{ there exists a covector } \bar{\mu}^{v,S} \in Y_{1,v,S}^* \text{ such that } \langle \bar{\mu}^{v,S}, \omega \rangle \leq 0, \text{ and the inequality } \langle \bar{\lambda}, \hat{P}(v, w, z) \rangle + \langle \bar{\mu}^{v,S}, \tilde{P}(v, w) \rangle \geq 0 \quad (14) \text{ holds for all } (w, z) \in S.$$

- Condition (4.&)—or its equivalent form (4.#)—implies the following statement:

$$(4.\%) \text{ for every } v \in \hat{\mathcal{V}}^3 \text{ there exist a covector } \bar{\mu}^v \in Y_{1,v}^* \text{ and a nonnegative number } \theta^v \text{ such that}$$

$$(4.\%.a) \text{ the inequalities}$$

$$\theta^v \langle \bar{\lambda}, P_1 \cdot z \rangle \geq 0 \quad (15)$$

$$\theta^v \langle \bar{\lambda}, \hat{P}(v, w, z) \rangle + \langle \bar{\mu}^v, \tilde{P}(v, w) \rangle \geq 0 \quad (16)$$

$$\text{hold for all } (w, z) \in \text{var}_{3,1}(u_*, v, \mathcal{U}_v),$$

$$(4.\%.b) \quad \langle \bar{\mu}^v, \omega \rangle \leq 0,$$

$$(4.\%.c) \text{ either } \theta^v > 0 \text{ or } \bar{\mu}^v \in Y_{1,v}^* \setminus \{0\}.$$

5. Necessary conditions for a minimum

For each $v \in \hat{\mathcal{V}}^3$ and each subset S of $\text{var}_{3,1}(u_*, v, \mathcal{U}_v)$, we let $X_{1,v,S}$ be the linear span of all the vectors $\tilde{P}(v, w)$, for all $(w, z) \in S$, and write $X_{1,v} \stackrel{\text{def}}{=} X_{1,v,\text{var}_{3,1}(u_*, v, \mathcal{U}_v)}$.

We write

$$\tilde{p}(v, w) \stackrel{\text{def}}{=} p_1 \cdot w + \frac{1}{2} p_2(v, v), \quad (17)$$

$$\hat{p}(v, w, z) \stackrel{\text{def}}{=} \frac{1}{6} p_3(v, v, v) + p_2(v, w) + p_1 \cdot z, \quad (18)$$

Theorem 5.1 Assume that $\nu \in \{1, 2, 3\}$, and the data $\mathcal{U}^a, X, \mathcal{U}, u_* \in \mathcal{U}, \mathcal{E}, \mathcal{I}, \mathcal{A}, \mathcal{F}, P_1, \dots, P_\nu, p_1, \dots, p_\nu$, satisfy Assumptions (M1) to (M4) of §3. If $\nu > 1$, assume that $\hat{\mathcal{V}}$ and $\{\mathcal{U}_v\}_{v \in \hat{\mathcal{V}}}$, are such that (A5) and (A6) hold, as well as the second-order cancellation and cross-term conditions (M7) and (M8). If $\nu = 3$, assume that the third-order cancellation and cross-term conditions (M3.9.a), (M3.9.b), (M3.9.c) are satisfied.

Let $x_* = \mathcal{E}(u_*)$, and assume that (MIN) holds. Then there exist a covector $\bar{\lambda} \in X^*$ and a $\lambda_0 \in \mathbb{R}$ such that

- (1) $\lambda_0 \geq 0$ and $\|\bar{\lambda}\| + \lambda_0 > 0$,
- (2) $\langle \bar{\lambda}, P_1 \cdot v \rangle + \lambda_0 p_1 \cdot v \geq 0$ for all $v \in \overline{\text{var}}_1(u_*, \mathcal{U})$,
- (3) if $\nu > 1$ then $\langle \bar{\lambda}, \tilde{P}(v, w) \rangle + \lambda_0 \tilde{p}(v, w) \geq 0$ whenever $v \in \hat{\mathcal{V}}$ and $w \in \text{var}_{2,1}(u_*, v, \mathcal{U}_v)$,
- (4) if $\nu = 3$ then for every $v \in \hat{\mathcal{V}}^3$,

$$(4.\&) \quad \langle \bar{\lambda}, \hat{P}(v, w, z) \rangle + \lambda_0 \hat{p}(v, w, z) \geq 0 \text{ whenever } (w, z) \in \text{var}_{3,1}(u_*, \mathcal{U}_v), \tilde{P}(v, w) = 0 \text{ and } \tilde{p}(v, w) \leq 0.$$

In addition, if $\nu > 1$, then (3) implies, in particular, that

$$(3.\#) \quad \langle \bar{\lambda}, P_1 \cdot w \rangle + \lambda_0 p_1 \cdot w \geq 0 \text{ whenever } v \in \hat{\mathcal{V}} \text{ and } (v, w) \in \text{var}_2(u_*, \mathcal{U}_v).$$

Furthermore, if $\nu = 3$, then

- For each $v \in \hat{\mathcal{V}}^3$, Condition (4.&) is equivalent to the following multiplier statement:

$$(4.\#) \text{ for every finite subset } S \text{ of } \text{var}_{3,1}(u_*, v, \mathcal{U}_v) \text{ there exists a covector } \bar{\mu}^{v,S} \in X_{1,v,S}^* \text{ and a nonnegative real number } \mu_0^{v,S} \text{ such that } \langle \bar{\lambda}, \hat{P}(v, w, z) \rangle + \langle \bar{\mu}^{v,S}, \tilde{P}(v, w) \rangle + \lambda_0 \left(\hat{p}(v, w, z) \right) + \mu_0^{v,S} \left(\tilde{p}(v, w) \right) \geq 0 \quad (19) \text{ for all } (w, z) \in S.$$

- (4.&)—or its equivalent form (4.#)—implies:

$$(4.\%) \text{ for every } v \in \hat{\mathcal{V}}^3 \text{ there exist a covector } \bar{\mu}^v \in X_{1,v}^* \text{ and nonnegative numbers } \mu_0^v, \theta^v \text{ such that}$$

$$(4.\%.a) \text{ the inequalities}$$

$$\theta^v \langle \bar{\lambda}, P_1 \cdot z \rangle \geq 0, \quad (20)$$

$$\theta^v \langle \bar{\lambda}, \hat{P}(v, w, z) \rangle + \langle \bar{\mu}^v, \tilde{P}(v, w) \rangle + \mu_0^v \left(\tilde{p}(v, w) \right) \geq 0 \quad (21)$$

$$\text{hold for all } (w, z) \in \text{var}_{3,1}(u_*, v, \mathcal{U}_v),$$

$$(4.\%.b) \text{ either } \theta^v \neq 0, \text{ or } \mu_0^v \neq 0, \text{ or } \bar{\mu}^v \in X_{1,v}^* \setminus \{0\},$$

$$(4.\%.c) \theta^v \lambda_0 = 0.$$

6. A high-order maximum principle

We now study the same optimal control problem as in §3, and state the necessary condition for a minimum that follows by applying Theorem 5.1.

Following [2], we write

$$R_2[f][h, v] = \frac{1}{2} f_{xx}^*(h, h) + f_{xu}^*(h, v) + \frac{1}{2} f_{uu}^*(v, v), \quad (22)$$

$$R_2[L][h, v] = \frac{1}{2} L_{xx}^*(h, h) + L_{xu}^*(h, v) + \frac{1}{2} L_{uu}^*(v, v), \quad (23)$$

$$R_3[f](h, v) = \frac{1}{6} f_{xxx}^*(h, h, h) + \frac{1}{2} f_{xxu}^*(h, h, v) + \frac{1}{2} f_{xuu}^*(h, v, v) + \frac{1}{6} f_{uuu}^*(v, v, v), \quad (24)$$

$$R_3[L](h, v) = \frac{1}{6} L_{xxx}^*(h, h, h) + \frac{1}{2} L_{xxu}^*(h, h, v) + \frac{1}{2} L_{xuu}^*(h, v, v) + \frac{1}{6} L_{uuu}^*(v, v, v), \quad (25)$$

and let

$$(\nabla^1 f(h, v))_x \cdot g = f_{xx}^*(g, h) + f_{xu}^*(g, v), \quad (26)$$

$$(\nabla^1 L(h, v))_x \cdot g = L_{xx}^*(g, h) + L_{xu}^*(g, v), \quad (27)$$

$$(\nabla^1 f(h, v))_u \cdot w = f_{xu}^*(h, w) + f_{uu}^*(v, w), \quad (28)$$

$$(\nabla^1 L(h, v))_u \cdot w = L_{xu}^*(h, w) + L_{uu}^*(v, w), \quad (29)$$

if $g \in L^\infty([a, b], \mathbb{R}^n)$, $w \in L^\infty([a, b], \mathbb{R}^m)$.

Theorem 6.1 Assume that $\nu \in \{1, 2, 3\}$, and conditions (C1) to (C7) of §3 hold. Then there exist an absolutely continuous field of covectors $[a, b] \ni t \mapsto \lambda(t) \in \mathbb{R}_n$ (where \mathbb{R}_n denotes the space of real n -dimensional row vectors) and a nonnegative $\lambda_0 \in \mathbb{R}$ such that

1. $\lambda_0 + \|\lambda(b)\| > 0$;
2. the adjoint equation $\dot{\lambda}(t) = -\lambda(t) \cdot f_x^*(t) - \lambda_0 L_x^*(t)$ is satisfied for almost all $t \in [a, b]$;
3. for almost all $t \in [a, b]$, the inequality

$$(\lambda_0 L_u^*(t) + \lambda(t) \cdot f_u^*(t)) \cdot (u - u_*(t)) \geq 0 \quad (30)$$

holds for all $u \in U$;

4. if $\nu > 1$, then for every $(h, v) \in \mathcal{H}$, the inequalities

$$\int_a^b (\lambda_0 L_u^* + \lambda \cdot f_u^*) \cdot w \, dt \geq 0, \quad (31)$$

$$\int_a^b \left((\lambda_0 L_u^* + \lambda \cdot f_u^*) \cdot w + \lambda_0 R_2[L](h, v) + \lambda(t) \cdot R_2[f](h, v) \right) dt \geq 0 \quad (32)$$

hold for every $w \in \text{var}_{2,1}(u_*, v, \mathcal{U})$ such that $\{t : w(t) \neq 0\} \subseteq I_{h,v}$;

5. if $\nu = 3$ then for every $(h, v) \in \mathcal{H}$ the inequality

$$\begin{aligned} & \int_a^b (\lambda_0 L_u^* + \lambda \cdot f_u^*) \cdot z \, dt \\ & + \int_a^b \lambda_0 \left((\nabla^1 L(h, v))_x \cdot \mathbf{y}^{v,w} + (\nabla^1 L(h, v))_u \cdot w \right) dt \\ & + \int_a^b \lambda \cdot \left((\nabla^1 f(h, v))_x \cdot \mathbf{y}^{v,w} + (\nabla^1 f(h, v))_u \cdot w \right) dt \\ & + \int_a^b \left(\lambda_0 R_3[L](h, v) + \lambda \cdot R_3[f](h, v) \right) dt \geq 0 \end{aligned} \quad (33)$$

holds for every $(w, z) \in \text{var}_{3,1}(u_*, v, \mathcal{U})$ such that $\{t : w(t) \neq 0\} \subseteq I_{h,v}$, $\{t : z(t) \neq 0\} \subseteq I_{h,v}$, and $\mathbf{y}^{v,w}(b) = 0$, where $\mathbf{y}^{v,w}$ is the solution of the initial value problem

$$\begin{cases} \dot{\mathbf{y}}^{v,w} &= f_x^* \cdot \mathbf{y}^{v,w} + f_u^* \cdot w + R_2[f](h, v), \\ \mathbf{y}^{v,w}(a) &= 0; \end{cases}$$

6. if $\nu = 3$ then for every member (h, v) of the set \mathcal{H} there exist absolutely continuous functions $[a, b] \ni t \mapsto \psi^{h,v}(t) \in \mathbb{R}^n$, $[a, b] \ni t \mapsto \rho^{h,v}(t) \in \mathbb{R}^n$, and nonnegative numbers $\psi_0^{h,v}$, $\theta^{h,v}$ such that

$$6.i. \quad \dot{\rho}^{h,v}(t) = -\rho^{h,v}(t) \cdot f_x^*(t) \text{ a.e. on } [a, b],$$

6.ii. the extended adjoint equation

$$\dot{\psi}^{h,v} = -\psi^{h,v} \cdot f_x^* - \psi_0^{h,v} L_x^* - \rho^{h,v} \cdot (\nabla^1 f(h, v))_x \quad (34)$$

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

$$6.iii. \quad \rho^{h,v}(t) = \theta^{h,v} \lambda(t) \text{ for all } t \in [a, b],$$

6.iv. the inequality

$$\begin{aligned} & \int_a^b \left(\psi^{h,v} \cdot f_u^* \cdot w + \psi_0^{h,v} L_u^* \cdot w + \rho^{h,v} \cdot f_{xu}^*(h, w) \right. \\ & \left. + \rho^{h,v} \cdot f_{uu}^*(v, w) + \rho^{h,v} \cdot f_u^* \cdot z \right) dt \\ & + \int_a^b \left(\psi^{h,v} \cdot R_2[f](h, v) + \psi_0^{h,v} R_2[L](h, v) \right. \\ & \left. + \rho^{h,v} \cdot R_3[f](h, v) \right) dt \geq 0 \end{aligned} \quad (35)$$

holds for all $(w, z) \in \text{var}_{3,1}(u_*, v, \mathcal{U})$ such that $w \equiv 0$ and $z \equiv 0$ outside $I_{h,v}$,

- 6.v. either $\theta^{h,v} \neq 0$, or $\mu_0^{h,v} \neq 0$, or the restriction of μ^v to $X_{1,v}$ does not vanish identically, where $X_{1,v}$ is the linear span of the vectors $\frac{1}{2}\mathcal{E}''(u_*)(v, v) + \mathcal{E}'(u_*)(w)$ for all pairs $(w, z) \in \text{var}_{3,1}(u_*, v, \mathcal{U})$ such that $w \equiv 0$ and $z \equiv 0$ outside $I_{h,v}$,

$$6.vi. \quad \theta^{h,v} \lambda_0 = 0,$$

7. for every member (h, v) of \mathcal{H} such that the restricted terminal point map \mathcal{E}_v^0 is 2-regular at $(0, v, \mathcal{U}_v^a, \mathcal{U}_v^0)$, the multipliers $\psi^{h,v}$, $\rho^{h,v}$, $\psi_0^{h,v}$, $\theta^{h,v}$, satisfy:

$$7.i. \quad \rho^{h,v}(t) \cdot f_u^*(t) \equiv 0 \text{ for almost all } t \in I_{h,v},$$

7.ii. the inequality

$$\begin{aligned} & \int_a^b \left(\psi^{h,v} \cdot f_u^* \cdot w + \psi_0^{h,v} L_u^* \cdot w + \rho^{h,v} \cdot f_{xu}^*(h, w) \right. \\ & \left. + \rho^{h,v} \cdot f_{uu}^*(v, w) \right) dt + \int_a^b \left(\psi^{h,v} \cdot R_2[f](h, v) \right. \\ & \left. + \psi_0^{h,v} R_2[L](h, v) + \rho^{h,v} \cdot R_3[f](h, v) \right) dt \geq 0 \end{aligned} \quad (36)$$

is true for all $w \in \text{var}_{2,1}(u_*, v, \mathcal{U})$ such that $w \equiv 0$ outside $I_{h,v}$,

- 7.iii. either $\psi_0^{h,v} \neq 0$, or the functional

$$\begin{aligned} & L^\infty(I_{h,v}, \mathbb{R}^m) \ni w \mapsto \int_{I_{h,v}} \left(\psi^{h,v} \cdot f_u^* \cdot w + \psi_0^{h,v} L_u^* \cdot w \right. \\ & \left. + \rho^{h,v} \cdot f_{xu}^*(h, w) + \rho^{h,v} \cdot f_{uu}^*(v, w) \right) dt \in \mathbb{R} \end{aligned}$$

does not vanish identically on $L^\infty(I_{h,v}, \mathbb{R}^m)$.

7. A simple example

We consider the control system in \mathbb{R}^2 given by

$$\begin{cases} \dot{x}_1 &= u_1 + \varphi_1(u_1, u_2, x_1, x_2), \\ \dot{x}_2 &= u_1 u_2 + \varphi_2(u_1, u_2, x_1, x_2), \end{cases}$$

where φ_1, φ_2 are real-valued functions on \mathbb{R}^4 of class C^3 and such that φ_1 and φ_2 vanish at $(0, 0, 0, 0)$ together with all their partial derivatives of order ≤ 3 .

We choose a fixed time interval, namely, $[0, 1]$. The controls are all the bounded measurable functions $[0, 1] \ni t \mapsto (u_1(t), u_2(t)) \in \mathbb{R}^2$. Therefore the space U of control values is \mathbb{R}^2 , and the space \mathcal{U} of controls is $L_\infty([0, 1], \mathbb{R}^2)$.

We let Ξ be the set of all trajectory-control pairs $(x(\cdot), u(\cdot))$, so the members of Ξ are the pairs $(x(\cdot), u(\cdot))$

such that $u(\cdot) \in \mathcal{U}$ and $x(\cdot)$ is an absolutely continuous map $[0, 1] \ni t \mapsto x(t) = (x_1(t), x_2(t))$ such that

$$\begin{aligned}\dot{x}_1(t) &= u_1(t) + \varphi_1(u_1(t), u_2(t), x_1(t), x_2(t)) \\ \dot{x}_2(t) &= u_1(t)u_2(t) + \varphi_2(u_1(t), u_2(t), x_1(t), x_2(t))\end{aligned}$$

for almost every $t \in [0, 1]$.

We define a cost functional $J : \Xi \mapsto \mathbb{R}$ by letting

$$J(x(\cdot), u(\cdot)) = \int_0^1 x_1(t)^3 dt \quad \text{for } (x(\cdot), u(\cdot)) \in \Xi.$$

Our optimal control problem is that of minimizing the cost among all trajectory-control pairs $(x(\cdot), u(\cdot)) \in \Xi_0$, where Ξ_0 is the set of all $(x(\cdot), u(\cdot)) \in \Xi$ such that $x_1(0) = x_2(0) = x_1(1) = x_2(1) = 0$.

We take our reference trajectory-control pair to be $\xi_* = (x_*, u_*)$, where $x_*(t) \equiv (0, 0)$ and $u_*(t) \equiv (0, 0)$. It is clear that (x_*, u_*) is not optimal, since, for example, it is easy to connect $(0, 0)$ to $(0, 0)$ by means of a trajectory $[0, 1] \ni t \mapsto (x_1(t), x_2(t))$ such that $x_1(t) < 0$ and $u_2(t) = 0$ whenever $0 < t < 1$, and such a trajectory will have a strictly negative cost. (That the trajectory exists is completely obvious if, say, $\varphi_1 \equiv \varphi_2 \equiv 0$, and is quite easy to prove in the general case).

Our goal is to see how well our necessary conditions perform in ruling out the optimality of ξ_* . The condition of Theorem 6.1 says that for ξ_* to be optimal there have to exist absolutely continuous functions $\lambda_1 : [0, 1] \mapsto \mathbb{R}$ and $\lambda_2 : [0, 1] \mapsto \mathbb{R}$, and a nonnegative λ_0 such that a nontriviality condition and certain inequalities hold, and the adjoint equation is satisfied.

It is clear that $f_x^* \equiv 0$ and $L_x^* \equiv 0$. Therefore the adjoint equation says that $\lambda(t) \equiv \text{constant} \equiv (\lambda_1, \lambda_2)$, where λ_1 and λ_2 are real numbers. Clearly,

$$f_u^* \equiv \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad L_u^* \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (37)$$

Therefore (30) implies that $\lambda_1 u_1 \geq 0$ for all pairs $(u_1, u_2) \in \mathbb{R} \times [0, \infty[$. It follows that $\lambda_1 = 0$.

If $h = (h_1, h_2)$ and $v = (v_1, v_2)$ we have

$$R_2[f](h, v) = \begin{bmatrix} 0 \\ v_1 v_2 \end{bmatrix}, \quad R_2[L](h, v) = 0. \quad (38)$$

To satisfy Hypothesis (C6) we have to make sure that every pair $(h, v) = ((h_1, h_2), (v_1, v_2)) \in \mathcal{H}$ is such that

$$h_1(t) = \int_0^t v_1(s) ds, \quad h_2(t) \equiv 0, \quad h_1(1) = 0. \quad (39)$$

This means that we can choose v_1 and v_2 to be arbitrary, provided only that $\int_0^1 v_1(t) dt = 0$, and then h_1 and h_2 are determined by (39).

The second-order inequality (32) says that

$$\begin{aligned} \int_0^1 \left((\lambda_0 L_u^* + \lambda \cdot f_u^*) \cdot w + \lambda_0 R_2[L](h, v) \right. \\ \left. + \lambda(t) \cdot R_2[f](h, v) \right) dt \geq 0. \end{aligned} \quad (40)$$

Then (37) and (38) imply, since $\lambda_1 = 0$, that

$$\lambda_2 \int_0^1 v_1(t) v_2(t) dt \geq 0. \quad (41)$$

We can choose \mathcal{H} to consist of two pairs (h^1, v^1) , (h^2, v^2) , where $v^1 = (v_1^1, v_2^1)$, $v^2 = (v_1^2, v_2^2)$, v_1^1, v_1^2 are functions supported in the intervals $I_{h^1, v^1} = [0, \frac{1}{2}]$, $I_{h^2, v^2} = [\frac{1}{2}, 1]$, not identically zero, and such that

$$\int_0^1 v_1^1(t) dt = \int_0^1 v_1^2(t) dt = 0, \quad v_2^1 \equiv v_1^1, \quad \text{and } v_2^2 \equiv -v_1^2.$$

Then (41) says that

$$\lambda_2 \int_0^1 (v_1^1(t))^2 dt \geq 0 \quad \text{and} \quad \lambda_2 \int_0^1 (v_1^2(t))^2 dt \leq 0. \quad (42)$$

Therefore $\lambda_2 = 0$. Then the third-order inequality (33) reduces to the assertion that $\lambda_0 \int_0^1 R_3[L](h^i, v^i) dt \geq 0$ for all $i \in \{1, 2\}$ such that there exists a $w = (w_1, w_2) \in L_\infty([0, 1], \mathbb{R}^2)$ having the property that $\int_0^1 w_1(t) dt = 0$, $\int_0^1 w_2(t) dt = -\int_0^1 v_1^i(t) v_2^i(t) dt$, and $\{t : w(t) \neq 0\} \subseteq I_{h^i, v^i}$.

It is clear that w does exist for $i = 1, 2$, since we can take $w_1(t) \equiv 0$, $w_2(t) = v_1^i(t) v_2^i(t)$. Moreover, $R_3[L](h^i, v^i) = (h_1^i)^3$. So the inequality

$$\lambda_0 \int_0^1 h_1^i(t)^3 dt \geq 0 \quad (43)$$

must be true for $i = 1, 2$. Clearly, we can choose (h^1, v^1) such that $h_1^1(t) \leq 0$ for all t . (Take, e.g., $v_1^1(t) = -\sin 4\pi t$ for $0 \leq t \leq \frac{1}{2}$.) Then $\int_0^1 h_1^1(t)^3 dt < 0$, so (43) implies that $\lambda_0 \leq 0$. Since $\lambda_0 \geq 0$, we conclude that $\lambda_0 = 0$, contradicting the nontriviality condition, and proving that ξ_* is not optimal. \diamond

Remark 7.1 For comparison, we point out that Theorem 6.1 of Ledzewicz-Schättler [2], as stated, is *not* applicable to our example, since the theorem requires that the first derivative of the cost functional not be identically zero (Hypothesis (iii) in Page 833 of [2]). \diamond

References

- [1] Ledzewicz, U., and H. Schättler, “An extended maximum principle.” *Nonlinear Analysis, Theory, Methods and Applications* **29**, 1997, pp. 159-183.
- [2] Ledzewicz, U., and H. Schättler, “A high-order generalized local maximum principle.” *SIAM J. Control Optim.* **38**, 2000, pp. 823-854.
- [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] Rampazzo, F., and H.J. Sussmann, “Set-valued differentials and a nonsmooth version of Chow’s Theorem.” In *Proc. 40th IEEE Conf. Decision and Control, Orlando, Fla., December 2000*, IEEE publications, New York, 2001.
- [5] Sussmann, H. J., “A maximum principle for hybrid optimal control problems.” In *Proc. 38th IEEE Conf. Decision and Control, Phoenix, AZ, Dec. 1999*. IEEE publications, New York, 1999, pp. 425-430.
- [6] Sussmann, H. J., “Résultats récents sur les courbes optimales.” In *15^e Journée Annuelle de la Société Mathématique de France (SMF)*, Publications de la SMF, Paris, 2000, pp. 1-52.
- [7] Sussmann, H. J., “New theories of set-valued differentials and new versions of the maximum principle of optimal control theory.” In *Nonlinear Control in the year 2000*, A. Isidori, F. Lamnabhi-Lagarigue and W. Respondek Eds., Springer-Verlag, London, 2000, pp. 487-526.
- [8] Sussmann, H. J., “Set-valued differentials and the hybrid maximum principle.” In *Proc. 39th IEEE Conf. Decision and Control, Sydney, Australia, December 12-15, 2000*, IEEE publications, New York, 2000.