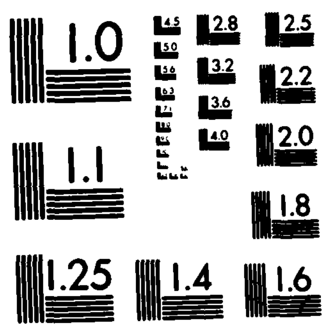
AD-A136 375 A HIGH-ORDER TEST FOR OPTIMALITY OF BANG-BANG CONTROLS 1/1
(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER
A BRESSAN NOV 83 MRC-TSR-2596 DAAG29-80-C-0041

UNCLASSIFIED F/G 12/1 NL

END
DATE
FILMED
2-84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MRC Technical Summary Report #2596

A HIGH-ORDER TEST FOR OPTIMALITY OF BANG-BANG CONTROLS

Alberto Bressan

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

November 1983

(Received September 20, 1983)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

A HIGH-ORDER TEST FOR OPTIMALITY OF BANG-BANG CONTROLS

Alberto Bressan*

Technical Summary Report #2596

November 1983

ABSTRACT

For control systems of the form $\dot{x} = X(x) + \sum_{i=1}^{k} Y_i(x)u_i$, a strengthened version of the classical Pontryagin Maximum Principle is proved. The necessary condition for optimality given here is obtained using functional analytic techniques and quite general high-order perturbations of the reference control. As shown by an example, our test is particularly effective when applied to bang-bang controls, a case where other high-order tests do not provide additional information.

AMS (MOS) Subject Classifications: 49B10, 49B27

Key Words: Admissible variational family, high-order tangent vector.

Work Unit Number 5 - Optimization and Large Scale Systems

---

* Istituto di Matematica Applicata, Università di Padova, ITALY.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

The determination of optimal controls is the primary problem in control theory. The most effective tool in this respect is Pontryagin's Maximum Principle, which provides necessary conditions for optimality, thus restricting the search to a small set of candidates among which the optimal control is more easily found.

In cases where Pontryagin's test does not yield sufficient information, more refined necessary conditions are known, for example Krener's High-order Maximum Principle. The high-order test discussed in this paper states further necessary conditions for optimality and can be particularly useful when applied to bang-bang controls, a case often encountered in the applications, where other high-order tests are generally ineffective.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

A HIGH-ORDER TEST FOR OPTIMALITY OF BANG-BANG CONTROLS

Alberto Bressan*

1. Introduction

Let $\mathcal{U}$ be a closed convex subset of the Banach space $L^1([0,T];\mathbb{R}^m)$ and consider a continuously Fréchet differentiable mapping $\phi : \mathcal{U} \to \mathbb{R}^n$. Given $\bar{u} \in \mathcal{U}$, in this paper we give a high-order sufficient condition for $\phi(\bar{u})$ to belong to the interior of the image $\phi(\mathcal{U})$. Problems of this kind arise frequently in control theory. Indeed, consider a control system of the form

$$\dot{x}(t) = f(x(t)) + G(x(t))u(t) \quad ,$$
$$x(0) = 0, \qquad u(t) \in \Omega \quad \text{for a.e.} \quad t \in [0,T] \quad , \tag{S}$$

where Ω is a compact convex subset of $\mathbb{R}^m$ and f, G are C^1 mappings from $\mathbb{R}^n$ into $\mathbb{R}^n$ and $\mathbb{R}^n \times \mathbb{R}^m$ respectively. If T is small enough, then (S) yields a C^1 map $\psi : u \to x(T,u)$ from the set $\mathcal{U}$ of admissible controls into $\mathbb{R}^n$. Here $x(T,u)$ is the point reached at time T by the solution of (S) corresponding to the control u. A classical problem is the following: given an admissible control $\bar{u}$, decide whether $\bar{u}$ is time optimal. This is often equivalent to showing that $x(T,\bar{u})$ lies on the boundary of the reachable set $\mathcal{R}(T)$.

A well known necessary condition for optimality is given by the Pontryagin Maximum Principle (PMP) [2,8]. Krener's high order maximum principle (HMP) [6] provides further conditions, obtained from the study of more general one parameter perturbations u_ξ of the control $\bar{u}$. If the first order variation at the terminal point of the trajectory

$$\lim_{\xi \to 0} [x(T,u_\xi) - x(T,\bar{u})]/\xi \tag{1.1}$$

---

* Istituto di Matematica Applicata, Università di Padova, ITALY.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

vanishes, a high order tangent vector can be generated, and additional necessary conditions for extremality are found. This method yielded several new results [3,4,5,6], especially concerning the problem of local stability. In this case, the reference control is $\bar{u}(t) \equiv 0$ and lies in the interior of $\Omega = [-1,+1]$. Hence there are several ways to locally perturb $\bar{u}$ and achieve a cancellation in the first order variation (1.1). The HMP can be here particularly effective. On the other hand, if $\bar{u}$ is bang-bang, $\bar{u}(t)$ already lies on the boundary of Ω, and only one-sided perturbations of $\bar{u}$ are admissible. As a result, in general there is no way of generating high order tangent vectors, as long as only the "instantaneous" control variations considered in [5,6] are used. In order to develop a genuine high order test for bang-bang controls, it is necessary to achieve the cancellation of the first order variation (1.1) by perturbing $\bar{u}$ simultaneously in the neighborhoods of two or more distinct times. This leads us to consider more general control variations.

In the following, the variable t always denotes time, while ξ, c are used as variational parameters: $u(\xi,\cdot)$ or $u(c,\xi,\cdot)$ will denote controls in $L^1([0,T];\mathbb{R}^m)$ depending continuously on the parameters ξ, c. In the abstract setting considered in sections 2 and 3, the control u is regarded merely as a point in a Banach space E, and we use the shorter notation $u(\xi)$ or $u(c,\xi)$ to indicate its dependence on one or two parameters.

Definition 1. A one-parameter admissible variational family of control functions (AVF) for a control $\bar{u}$ on $[0,T]$, generating a tangent vector $v \in \mathbb{R}^n$, is a continuous map $\gamma : \xi \to u(\xi,\cdot)$ from a nondegenerate interval $[0,\hat{\xi}]$ into $L^1([0,T];\mathbb{R}^m)$ such that

$$u(0,\cdot) = \bar{u}(\cdot) \quad , \quad u(\xi,\cdot) \in U \quad \forall\, \xi \in [0,\hat{\xi}] \quad , \tag{1.2}$$

$$\lim_{\xi \to 0} [x(T,u(\xi,\cdot)) - x(T,\bar{u}(\cdot))]/\xi = v \quad . \tag{1.3}$$

We say that the AVF γ has order h if there exist constants C_1, C_2 for which

$$0 < C_1 < \xi^{-1} \|u(\xi,\cdot) - \overline{u}(\cdot)\|^h_{L^1} < C_2, \quad \forall\ \xi \in (0,\hat{\xi}] \quad . \tag{1.4}$$

Notice that one can recover every high order tangent vector by means of the first order derivative (1.3), via a suitable change of the parameter ξ. As shown in [5], this method differs from Krener's only in computational ease. The above class of AVF is at the same time simpler and more general than those studied in [5,6], hence the corresponding family of tangent vectors can be much larger. One would like to use all of these vectors to derive a stronger HMP.

Assume that, given suitable variational families γ_i for $\overline{u}(i=0,\ldots,k)$, the positive span of the corresponding tangent vectors v_i is all of $\mathbb{R}^n$. To conclude that $x(T,\overline{u})$ lies in the interior of the reachable set $R(T)$, one has to construct approximate convex combinations of the v_i continuously depending on the parameters. More precisely, the γ_i should be <u>summable</u> in the sense of

<u>Definition 2</u>. Let $F = \{\gamma_0,\ldots,\gamma_k\}$ be a finite collection of AVF for the control $\overline{u}$, generating the tangent vectors $v_0,\ldots,v_k$. Set

$$\Delta^k = \{c = (c_0,\ldots,c_k);\ c_i \geq 0,\ \sum_{i=0}^{k} c_i = 1\} \quad . \tag{1.5}$$

F is summable if there exist $\overline{\xi} > 0$ and a continuous map $(c,\xi) \to u(c,\xi,\cdot)$ from $\Delta^k \times [0,\overline{\xi}]$ into $L^1([0,T];\mathbb{R}^m)$ such that, for all $c \in \Delta^k$

$$u(c,0,\cdot) = \overline{u}(\cdot),\ u(c,\xi,\cdot) \in U \quad \forall\ \xi \in [0,\overline{\xi}] \quad , \tag{1.6}$$

$$\lim_{\xi \to 0} [x(T,u(c,\xi,\cdot)) - x(T,\overline{u}(\cdot))]/\xi = \sum_{i=0}^{k} c_i v_i \tag{1.7}$$

uniformly on Δ^k .

This crucial property holds for control variations of the special kind considered in [5,6], but is not satisfied by an arbitrary collection of AVF (see §5 for a counterexample). Our key result is that if all but one of the $\gamma_i \in \mathcal{F}$ have order 1, then $\mathcal{F}$ is summable. This is first proven in an abstract setting, then stated for the control system (S). We thus obtain a strengthened version of the PMP which is particularly effective when applied to bang-bang controls. Indeed, our single high order variation is allowed to be quite arbitrary. An application of this technique is given in §5.

2. Notations, statement of the main results.

Consider a mapping ϕ from a neighborhood of a closed subset $\mathcal{U}$ of a Banach space E into $\mathbb{R}^n$ and denote by $D\phi(u)$ its differential at u. We say that ϕ is C^1 if the map $u \to D\phi(u)$ from E into the space of continuous linear operators $L(E,\mathbb{R}^n)$ is continuous. For the definition of the operator norm on $L(E,\mathbb{R}^n)$ and for the basic properties of differentials our general reference is Diêudonne [1].

If $\bar{u} \in \mathcal{U}$, by an <u>admissible variational family</u> (AVF) for $\bar{u}$, generating a tangent vector $v \in \mathbb{R}^n$, we mean a continuous map $\gamma : \xi \to u(\xi)$ from $[0,1]$ into $\mathcal{U}$ such that

$$u(0) = \bar{u},\ u(\xi) \in \mathcal{U} \quad \forall\, \xi \in [0,1] \ , \tag{2.1}$$

$$\lim_{\xi \to 0} [\phi(u(\xi)) - \phi(\bar{u})]/\xi = v \ . \tag{2.2}$$

If, for some $0 < C_1 < C_2 < \infty$ and all $\xi \in (0,1]$,

$$C_1 < \xi^{-1} \| u(\xi) - \bar{u} \|^h < C_2 \tag{2.3}$$

we say that γ has order h. We write $B(x,r)$ for the closed ball centered at x with radius r. The Euclidean norm on $\mathbb{R}^n$ and the operator norm on the set of $n \times m$ matrices are both written as $|\cdot|$, while double bars are used for the norms $\|\cdot\|$ in Banach spaces such as E or $L(E,\mathbb{R}^n)$. Int A, ∂A, $\overline{co}\, A$ denote the interior, the boundary and the convex closure of a set A. With these conventions we have

Theorem 1 - Let $\mathcal{U}$ be a closed convex subset of a Banach space E, and let ϕ be a C^1 mapping from a neighborhood of $\mathcal{U}$ into $\mathbb{R}^n$. Assume $\bar{u} \in \mathcal{U}$ and let $\gamma_i : \xi \to u_i(\xi)$ be AVF of $\bar{u}$ generating the tangent vectors v_i $(i=0,\dots,k)$. If $0 \in \text{int}\, \overline{co}\{v_0,\dots,v_k\} \subseteq \mathbb{R}^n$ and $\gamma_1,\dots,\gamma_k$ have order 1, then $\phi(\bar{u}) \in \text{int}\, \phi(\mathcal{U})$.

From this result, a sharper form of Pontryagin's maximum principle for the system (S) can be derived. To fix the ideas, assume that f and G are C^1 on $B(0,r) \subseteq \mathbb{R}^n$; f_x, G_x will denote partial derivatives. Let

$$\sup\{|f(x)|,\ |G(x)|;\ x \in B(0,r)\} < M_1$$

$$\sup\{|w|;\ w \in \Omega\} < M_2,\ 0 < T < (M_1 + M_1 M_2)^{-1} r \ .$$

This guarantees that, for every control $u \in L^1([0,T];\mathbb{R}^m)$ taking values in Ω, $\|u\| < M_2 T$ and there exists a unique solution $t \to x(t,u)$ of (S) defined on $[0,T]$, taking values inside $B(0,r)$. Notice that the open ball $\mathcal{B} = \{u \in L^1; \|u\| < M_2 T\}$ is a neighborhood of the set of admissible controls

$$\mathcal{U} = \{u \in L^1([0,T];\mathbb{R}^m);\ u(t) \in \Omega \ \text{ a.e. in } [0,T]\} \ .$$

We assume that Ω is closed, bounded and convex, thus the same holds for $\mathcal{U}$. The map $\psi : \mathcal{B} \to C^0([0,T];\mathbb{R}^n)$ that associates to each control u the corresponding solution $x(\cdot,u)$ of (S) is continuously Fréchet differentiable. Indeed ψ is implicitly defined by the equation $\psi(u) = \Psi(u,\psi(u))$, with

$$\Psi(u,x)(t) = \int_0^t f(x(s))ds + \int_0^t G(x(s))u(s)ds \ . \qquad (2.4)$$

The map Ψ can be thought of as the composition $\Psi_2 \circ \Psi_1$, defined by

$$\Psi_1(u,x)(t) = (u(t), f(x(t)), G(x(t))) \ ,$$

$$\Psi_2(u,y_1,y_2)(t) = \int_0^t y_1(s)ds + \int_0^t y_2(s)u(s)ds \ .$$

Clearly Ψ_1 is C^1 and Ψ_2 is bilinear. Hence Ψ is C^1 and the same holds for ψ, because of the implicit function theorem ([1], pg. 275). An application of Theorem 1 yields

Theorem 2 - Let $\bar{u}$ be an admissible control for the system (S) and assume that $x(t,\bar{u}) \in \partial R(T)$. Then, for every tangent vector v_0 generated by a (possible high-order) admissible variational family γ_0 for $\bar{u}$, there exists and absolutely continuous nonzero n-vector valued function $t \to \lambda(t)$ on $[0,T]$ which satisfies

$$\lambda(T) \cdot v_0 < 0 \tag{2.5}$$

$$\dot{\lambda}(t) = -\lambda(t)\left[f_x(x(t,\bar{u})) + G_x(x(t,\bar{u}))\bar{u}(t)\right]$$

$$\lambda(t)G(x(t,\bar{u}))\bar{u}(t) = \max\{\lambda(t)G(x(t,\bar{u}))u;\ u \in \Omega\} \tag{2.7}$$

for almost every t in $[0,T]$.

3. Proof of Theorem 1

Assume that $0 \in \operatorname{int} \overline{\operatorname{co}}\{v_0,\ldots,v_k\}$. From the family $\{v_0,\ldots,v_k\}$ choose $n+1$ nontrivial tangent vectors $v_0',\ldots,v_n'$ such that $0 \in \operatorname{int} \overline{\operatorname{co}}\{v_0',\ldots,v_n'\}$. Notice that, if v_0 is not one of the chosen vectors, any other vector can play its distinguished role. For notational convenience, assume that $0 \in \operatorname{int} \overline{\operatorname{co}}\{v_0,\ldots,v_n\}$. Relying on the fact that $\gamma_1,\ldots,\gamma_n$ have order 1 we now prove

Lemma 1. The family of admissible variations $F = \{\gamma_0,\ldots,\gamma_n\}$ is summable.

Proof. Define the scalar function α by

$$\alpha(\xi) = \sup\{\|u_0(\zeta) - \bar{u}\|^{1/2};\ 0 < \zeta < \xi\} \quad . \tag{3.1}$$

Clearly α is a continuous, nondecreasing function with $\alpha(0) = 0$. By (2.2), for $\xi > 0$ small enough we have

$$\|u_0(\xi) - \bar{u}\|/\xi > |v_0|/2\|D\phi(\bar{u})\| \quad . \tag{3.2}$$

Therefore there exists a $\bar{\xi} > 0$ such that

$$\left(\frac{|v_0|\xi}{2\|D\phi(\bar{u})\|}\right)^{1/2} < \alpha(\xi) < 1, \quad \xi/\alpha(\xi) < 1 \tag{3.3}$$

for all $\xi \in (0,\bar{\xi}]$. Define $u(c,\xi)$ on $\Delta^n \times [0,\bar{\xi}]$ by

$$u(c,\xi) = u_0(c_0\xi) + \sum_{i=1}^{n} c_i(\xi/\alpha(\xi))[u_i(\alpha(\xi)) - u_0(c_0\xi)]$$
$$\text{if} \quad 0 < \xi < \bar{\xi} \quad , \quad u(c,0) = \bar{u} \quad . \tag{3.4}$$

By (3.3), $u(c,\xi)$ is well defined and takes values inside U, being a convex combination of members of U. As $\xi \to 0$, $u_0(c,\xi)$ tends to $\bar{u}$ and each term inside the summation in (3.4) tends to zero uniformly w.r.t. c. Therefore u depends continuously on the parameters c,ξ. To show (1.7) we write

$$\frac{\phi(u(c,\xi)) - \phi(\bar{u})}{\xi} = \frac{\phi(u_0(c,\xi)) - \phi(\bar{u})}{\xi} + \frac{\phi(u(c,\xi)) - \phi(u_0(c_0\xi))}{\xi} \quad . \tag{3.5}$$

As $\xi \to 0$, the first term on the right hand side of (3.5) converges to $c_0 v_0$. The second term can be written as

$$\int_0^1 D\phi(\theta\cdot u(c,\xi) + (1-\theta)\, u_0(c_0\xi)) \cdot (u(c,\xi) - u_0(c_0\xi))d\theta$$
$$= \int_0^1 \left[D\phi(\bar{u}) + \chi(c,\xi,\theta)\right] \cdot \left\{\sum_{i=1}^{n} c_i(1/\alpha(\xi))\left[u_i(\alpha(\xi)) - u_0(c_0\xi)\right]\right\}d\theta \quad . \tag{3.6}$$

The continuous Fréchet differentiability of ϕ implies that $\chi(c,\xi,\theta) = D\phi(\theta\cdot u(c,\xi) + (1-\theta)\, u_0(c_0\xi)) - D\phi(\bar{u})$ is a continuous linear operator whose norm tends to zero uniformly in c, t as $\xi \to 0$. Observe that

$$(1/\alpha(\xi))\|u_i(\alpha(\xi)) - u_0(c_0\xi)\| < (1/\alpha(\xi))\, \|u_i(\alpha(\xi)) - \bar{u}\| +$$
$$(1/\alpha(\xi))\|u_0(c_0\xi) - \bar{u}\| < k_i + \|u_0(c_0\xi) - \bar{u}\|^{1/2} \tag{3.7}$$

for some finite constants k_i $(i=1,\ldots,n)$, because the u_i have order 1 and by (3.1). The limit as $\xi \to 0$ of the last term in (3.5) is therefore give.

$$\lim_{\xi \to 0} D\phi(\overline{u}) \cdot \sum_{i=1}^{n} (c_i/\alpha(\xi))\,(u_i(\alpha(\xi)) - \overline{u}) \quad . \tag{3.8}$$

By the definition (2.2) of tangent vector, one has

$$\begin{aligned} v_i &= \lim_{\xi \to 0} \frac{\phi(u_i(\xi)) - \phi(\overline{u})}{\xi} = \lim_{\xi \to 0} \frac{D\phi(\overline{u})\cdot(u_i(\xi) - \overline{u}) + o(\xi)}{\xi} \\ &= \lim_{\xi \to 0} D\phi(\overline{u})\cdot(u_i(\xi) - \overline{u})\cdot\xi^{-1} \quad . \end{aligned} \tag{3.9}$$

Indeed u_i is a first order AVF, hence the term $o(\xi)$, which is infinitesimal of higher order w.r.t. $\|u_i(\xi) - \overline{u}\|$ as $\xi \to 0$, is also of higher order w.r.t. ξ. Comparing (3.9) with (3.8) one concludes that

$$\lim_{\xi \to 0} [\phi(u(c,\xi) - \phi(\overline{u})]\,\xi^{-1} =$$

$$c_0 v_0 + \sum_{i=1}^{n} c_i \lim_{\xi \to 0} D\phi(\overline{u}) \cdot [u_i(\alpha(\xi)) - \overline{u}]/\alpha(\xi) = \sum_{i=0}^{n} c_i v_i \quad ,$$

uniformly in c.

Using the above lemma, the proof of Theorem 1 can now be completed by an application of Brouwer's fixed point theorem. Let $\delta = \text{dist}(0, \partial\overline{co}\{v_0,\dots,v_n\})$ and choose $\xi_0 > 0$ so small that

$$|\phi(u(c,\xi_0)) - \phi(\overline{u}) - \xi_0 \sum_{i=0}^{n} c_i v_i| < \xi_0\cdot\delta/2 \tag{3.10}$$

for all $c \in \Delta^n$. Consider the injective map $\sigma : \Delta^n \to \mathbb{R}^n$ defined by

$$\sigma(c) = \phi(\overline{u}) + \xi_0 \sum_{i=1}^{n} c_i v_i \quad .$$

For $x \in B(\phi(\overline{u}), \xi_0\delta)$ define $F(x) = \phi(u(\sigma^{-1}(x),\xi_0))$. By (3.10), $|F(x) - x| < \xi_0\delta/2$. For each $x_0 \in B(\phi(\overline{u}), \xi_0\delta/2)$, an application of Brouwer's theorem ([8] pg. 251) now implies the existence of some $x \in B(\phi(\overline{u}), \xi_0\delta)$ for which $F(x) = x_0$. Hence $B(\phi(\overline{u}), \xi_0\delta/2) \subseteq \phi(\mathcal{U})$. Q.E.D.

4. Proof of Theorem 2.

Suppose that the conclusion is false. Then there exists an admissible variational family γ_0 for $\bar{u}$ possibly high order, that generates a tangent vector v_0 such that, for every absolutely continuous $\lambda(\cdot)$ satisfying (2.5) and (2.6), one has

$$\lambda(t)\cdot G(x(t,\bar{u}))\cdot\bar{u}(t) < \max\{\lambda(t)\cdot G(x(t,\bar{u}))\cdot u;\ u \in \Omega\} \tag{4.1}$$

for t in a subset $J \subseteq [0,T]$ having positive measure. For each vector $\eta \neq 0$ with $\eta\cdot v_0 < 0$, let $\lambda_\eta(\cdot)$ be the unique solution of (2.6) for which $\lambda_\eta(T) = \eta$, and choose a control $u_\eta \in \mathcal{U}$ such that

$$\lambda_\eta(t)\cdot G(x(t,\bar{u}))\cdot u_\eta(t) = \max\{\lambda_\eta(t)\cdot G(x(t,\bar{u}))u;\ u \in \Omega\} \tag{4.2}$$

for a.e. $t \in [0,T]$. The continuity of λ_η, G and $x(\cdot,\bar{u})$ and a selection theorem [7] imply that such a measurable u_η exists. Define an AVF γ for $\bar{u}$ by setting

$$u(\xi,\cdot) = \xi u_\eta(\cdot) + (1-\xi)\bar{u}(\cdot), \qquad \forall\ \xi \in [0,1] \ . \tag{4.3}$$

Then, for every ξ, $u(\xi) \in \mathcal{U}$ because $\mathcal{U}$ is convex, and $\|u(\xi) - \bar{u}\|/\xi = \|u_\eta - \bar{u}\| \neq 0$, showing that γ has order one. Let $\Pi_T : C^0[0,T] \to \mathbf{R}^n$ be the linear projection $x \to x(T)$. From the remarks made in §2 it follows that the map $\xi \to x(t,u(\xi))$ is the composition of C^1 mappings, hence the tangent vector generated by the AVF (4.3) exists and is given by

$$v = \lim_{\xi\to 0}\ [x(T,u(\xi)) - x(T,\bar{u})]/\xi = \Pi_T\cdot D\psi(u)\cdot(u_\eta - \bar{u}) =$$
$$\int_0^T M(T,s)\cdot G(x(s,\bar{u}))\cdot(u_\eta(s) - \bar{u}(s))ds \tag{4.4}$$

where $s \to M(T,s)$ is the matrix fundamental solution of

$$\dot{z}(t) = [f_x(x(t,\bar{u})) + G_x(x(t,\bar{u}))\cdot\bar{u}(t)]z(t)$$

with $M(T,T) = I$, and where ψ is the input-output map defined above (2.4). By (2.6) the inner product of η and v is

$$\eta\cdot v = \int_0^T \lambda_\eta(T)M(T,s)\cdot G(x(s,u))\cdot(u_\eta(s) - u(s))ds =$$
$$\int_0^T \lambda_\eta(s)\cdot G(x(s,\bar{u}))\cdot(u_\eta(s) - \bar{u}(s))ds > 0 \tag{4.5}$$

because of (4.1). Hence, for every nontrivial vector η with $\eta\cdot v_0 < 0$, there exists a first order tangent vector v for which $\eta\cdot v > 0$. The positive span of the set of first order tangent vectors together with v_0 is thus the whole space $\mathbb{R}^n$. Theorem 1 applied to the C^1 map $\phi = \Pi_T\cdot\psi$: $u \to x(T,u)$ yields $x(T,\bar{u}) \in \text{int } R(T)$, a contradiction.

5. Examples.

The assumption on the order of the control variations in Theorem 1 is essential. Indeed, two arbitrary second order AVF need not be summable, as shown by

Example 1. Define a time dependent system on $\mathbb{R}^3$ by

$$\begin{aligned} &(\dot{x}_1(t), \dot{x}_2(t), \dot{x}_3(t)) = (\varphi_2(t)x_3(t)u_1(t), \varphi_2(t)x_3(t)u_2(t) \quad , \\ &\varphi_1(t)u_3(t)) \quad , \\ &(x_1(0), x_2(0), x_3(0)) = (0,0,0) \quad , \end{aligned} \tag{5.1}$$

where $t \in [0,3]$, the smooth functions φ_1, φ_2 satisfy

$$\begin{aligned} &\varphi_1(t) = 0,\ \varphi_2(t) > 0 \quad \text{for} \quad t \in [1,2] \quad , \\ &\varphi_2(t) = 0 \quad \text{for} \quad t \in [0,1] \cup [2,3] \quad , \\ &\int_0^1 \varphi_1(t)dt = \int_1^2 \varphi_2(t)dt = -\int_2^3 \varphi_1(t)dt = 1 \end{aligned} \tag{5.2}$$

and the controls satisfy the constraints

$$0 \le u_i(t) \le 1 \quad (i=1,2), \quad -\infty < u_3(t) < \infty \quad . \tag{5.3}$$

The reachable set at time $t = 3$ is then

$$R(3) = \{(x_1,x_2,x_3);\ x_1x_2 > 0\} \quad . \tag{5.4}$$

Let $\bar{u}$ be the null control. Consider the two AVF for $\bar{u}$:

$$u^{(1)}(\xi)(t) = (\xi^{1/2},0,\xi^{1/2}),\ u^{(2)}(\xi)(t) = (0,\xi^{1/2},-\xi^{1/2}) \quad ,$$

constant on the time interval $[0,3]$. Notice that for $i = 1,2$

$$\|u^{(i)}(\xi) - \bar{u}\|^2/\xi = (\int_0^3 |u^{(i)}(\xi)(t)|dt)^2/\xi = 18 \quad . \tag{5.5}$$

By setting $h = 2$, $C_1 = C_2 = 18$ in (1.4) one checks that $u^{(1)}$ and $u^{(2)}$ are both of second order. The endpoints of the corresponding trajectories are

$$x(3,u^{(1)}(\xi)) = (\xi,0,0),\ x(3,u^{(2)}(\xi)) = (0,-\xi,0) \quad . \tag{5.6}$$

Hence $u^{(1)}$ and $u^{(2)}$ generate the tangent vectors

$$v_1 = (1,0,0), \quad v_2 = (0,-1,0) \quad . \tag{5.7}$$

Comparing (5.7) with (5.4), it is clear that these two AVF cannot be summable. In this example, the set of high order tangent vectors of the special type considered in [6] is the cone $\Gamma = \{(0,0,x_3);\ x_3 \in \mathbb{R}\}$. This is of course convex and coincides here with the first order tangent cone. Notice that the time dependency can be easily removed by adjoining a new variable $x_0 = t$.

We now illustrate a non trivial application of Theorem 2 to the study of optimality of bang-bang controls.

Example 2. Consider the three dimensional autonomus system with scalar control $u(t) \in [-1,1]$:

$$\begin{aligned} (\dot{x}_1,\dot{x}_2,\dot{x}_3) &= (u,x_1,x_2+kx_1^2/2) \quad , \\ (x_1(0),x_2(0),x_3(0)) &= (0,0,0) \quad . \end{aligned} \tag{5.8}$$

The adjoint equations for this system are

$$(\dot{\lambda}_1,\dot{\lambda}_2,\dot{\lambda}_3) = (-\lambda_2-kx_1\lambda_3,\ -\lambda_3,0) \quad . \tag{5.9}$$

If $|k| < 1$, then a theorem of Sussmann [9] yields the existence of a $T > 0$ such that every time optimal control $u(\cdot)$ on $[0,T]$ is bang-bang with at most two switchings. If $|k| > 1$, the above result does not apply. Indeed, for every $T > 0$, there exist bang-bang controls u that satisfy Pontryagin's necessary conditions for optimality and have an arbitrarily large number of switchings on $[0,T]$. In order to construct a regular feedback synthesis for (5.8) it is important to rule out the optimality of these controls. In this direction we prove

Proposition 1. Assume $k > 1$. Then every bang-bang control $\bar{u}$ assuming the value $+1$ on a positive neighborhood of the origin is not optimal after its third switching time.

Proof. Let $\bar{u}$ be a bang-bang control which is initially $+1$ and has at least 3 switchings, and let $0 < t_1 < t_2 < t_3$ be its first three switching times. Fix any $T > t_3$, smaller then the fourth switching time if there is any. We will prove that $x(T,\bar{u}) \in \text{int } R(T)$. If the classical Pontryagin's necessary conditions do not hold for $\bar{u}$ on $[0,T]$, we are done. Otherwise, let $\lambda(t) = (\lambda_1(t),\lambda_2(t),\lambda_3(t))$ be a nontrivial adjoint variable satisfying (2.6) and (2.7), given in this case by (5.9) and

$$\bar{u}(t) = \text{sgn } \lambda_1(t) \quad \text{a.e. on} \quad [0,T] \tag{5.10}$$

respectively. Our first task is to compute $\lambda(T)$. Set $t_0 = 0$, $t_4 = T$. From (5.9) it follows that the map $t \to \lambda(t)$ is C^1 on $[0,T]$ and piecewise analytic on $[t_{i-1},t_i]$ $(i=1,\dots,4)$. In particular, we have

$$\lambda_3(t) = \lambda_3(0), \qquad \lambda_2(t) = \lambda_2(0) - t\lambda_3 \tag{5.11}$$

$$\ddot{\lambda}_1(t) = \lambda_3(1 - k \text{ sgn } \lambda_1(t)) \quad \text{a.e. on} \quad [0,T] \tag{5.12}$$

$$\lambda_1(t_i) = 0 \qquad (i = 1,2,3) \quad . \tag{5.13}$$

Hence $\lambda_1(t)$ is a polynomial of degree 2 in t on each subinterval $[t_{i-1},t_i]$. If $\lambda_1(t) = 0$ for some $t \in (t_1,t_2)$, then we would have $\lambda(t) \equiv 0$, against the assumptions. Thus $\lambda_1(t) \neq 0$ for $t_1 < t < t_2$. By

(5.13), $\ddot{\lambda}_1$ is not identically zero. Together with (5.12), this implies $\lambda_3 > 0$. Multiplying λ_3 by a positive scalar, we can therefore assume $\lambda_3(t) \equiv 1$. This, together with (5.12) and (5.13), determines $\lambda_1(t)$ uniquely:

$$\lambda_1(t) = \frac{1+k}{2}(t-t_1)(t-t_2), \quad \text{for} \quad t \in [t_1,t_2] \tag{5.14}$$

$$\lambda_1(t) = \frac{1-k}{2}(t-t_2)(t-t_3), \quad \text{for} \quad t \in [t_2,t_3] \quad . \tag{5.15}$$

The computation of $\dot{\lambda}_1(t_2)$ using alternatively (5.9), (5.14) and (5.15) yields

$$\begin{aligned}\dot{\lambda}_1(t_2) = -\lambda_2(t_2) - kx_1(t_2) = -\lambda_2(t_2) - k(2t_1-t_2) = \\ \frac{k+1}{2}(t_2-t_1) = \frac{k-1}{2}(t_3-t_2) \quad .\end{aligned} \tag{5.16}$$

Notice that the above expressions coincide because λ is C^1. From (5.9), (5.16) we deduce

$$\lambda_2(T) = t_1(1-3k)/2 + t_2(1+k)/2 - T \quad . \tag{5.17}$$

For notational convenience, set $a = t_1$, $b = t_2-t_1$, $c = t_3-t_2$, $d = T-t_3$. So far, we have proven that, up to a positive scalar factor, there exists a unique adjoint variable $\lambda(t)$ that satisfies (2.6) and (2.7) on $[0,T]$. In particular, (5.17) and the last equality in (5.16) yield

$$\lambda(T) = (\lambda_1(T), -ka + \frac{k-1}{2}b - c - d, 1) \quad , \tag{5.18}$$

$$(k+1)b = (k-1)c \quad . \tag{5.19}$$

The second part of the proof consists in the construction of a second order AFV for $\overline{u}$ generating at $t = T$ a tangent vector v having a positive inner product with $\lambda(T)$. A lengthy but elementary computation (see Appendix) shows that the control $\overline{u}$ steers the system from the origin to a point $x(T,\overline{u})$ whose coordinates are

$$
\begin{aligned}
x_1(T,\overline{u}) &= a - b + c - d \quad , \\
x_2(T,\overline{u}) &= T^2/2 - (b+c+d)^2 + (c+d)^2 - d^2 \quad , \\
x_3(T,\overline{u}) &= [T^3/2 - (b+c+d)^3 + (c+d)^3 - d^3]/3 + \\
&\quad [a^3 + (b-a)^3 + (c-b+a)^3 + (d-c+b-a)^3/2]k/3 \quad .
\end{aligned}
\tag{5.20}
$$

For $t \in [0,T]$ and $\xi > 0$ suitably small define

$$
\begin{aligned}
u(\xi)(t) &= 1 \quad \text{if} \quad t \in [0,a+\xi^{1/2}c) \cup [a+b+\xi^{1/2}(b+c),\, T-d+\xi^{1/2}b) \quad , \\
u(\xi)(t) &= -1 \quad \text{if} \quad t \in [a+\xi^{1/2}c, a+b+\xi^{1/2}(b+c)) \cup [T-d+\xi^{1/2}d, T] \quad .
\end{aligned}
$$

The coordinates of $x(T,u(\xi))$ are thus obtained from (5.20), replacing a,b,c,d by $a+\xi^{1/2}c$, $b+\xi^{1/2}b$, $c-\xi^{1/2}c$, $d-\xi^{1/2}b$ respectively. Using (5.19), one checks that in the expression of $x(T,u(\xi))$ all terms in $\xi^{1/2}$ cancel, hence the map $\xi \to u(\xi)$ is an AVF of $\overline{u}$ of order 2. The computation of the corresponding tangent vector v defined by (1.3) yields (see Appendix)

$$
v = (0,\ 2,\ b + c + 2d + k(2a - b + c))bc \quad . \tag{5.21}
$$

The inner product of (5.18) and (5.21) is

$$
\lambda(T)\cdot v = (k - 1)bc^2 > 0 \quad .
$$

This shows that the necessary conditions for extremality given in Theorem 2 do not hold for $\overline{u}$, hence $x(T,\overline{u}) \in \text{int } R(T)$. For any $T > t_3$, $\overline{u}$ is not time optimal after T, therefore $\overline{u}$ is not optimal after its third switching time.

APPENDIX

For the control $\overline{u}$ considered in Example 2, the coordinates of the point $x(T,\overline{u})$ are

$$x_1(T,\overline{u}) = \int_0^T u(s)ds = a-b+c-d$$

$$x_2(T,\overline{u}) = \int_0^T (T-s)u(s)ds = T^2/2 - (b+c+d)^2 + (c+d)^2 - d^2$$

$$x_3(T,\overline{u}) = \int_0^T \frac{(T-s)^2}{2} u(s)ds + \frac{k}{2}\int_0^T (x_1(s,\overline{u}))^2 ds$$

$$= \left[\int_0^a - \int_a^{a+b} + \int_{a+b}^{a+b+c} - \int_{a+b+c}^T\right] \frac{(T-s)^2}{2} ds$$

$$+ \frac{k}{2}\left[\int_0^a s^2 ds + \int_a^{a+b}(2a-s)^2 ds + \int_{a+b}^{a+b+c}(s-2b+2a)^2 ds + \int_{a+b+c}^T (2a-2b+2c-s)^2 ds\right]$$

$$= \left[\frac{T^3}{6} - \frac{(b+c+d)^3}{3} + \frac{(c+d)^3}{3} - \frac{d^3}{3}\right]$$

$$+ \frac{k^2}{2}\left[\frac{2a^3}{3} + \frac{2(a-b)^3}{3} + \frac{2(c-b+a)^3}{3} + \frac{(d-c+b-a)^3}{3}\right] .$$

The coordinates of $x(T,u(\xi))$ are:

$$x_1(T,u(\xi)) = x_1(T,\overline{u})$$

$$x_2(T,u(\xi)) = T^2/2 - (b+c+d) - c\xi^{1/2})^2 + (c+d-(b+c)\xi^{1/2})^2$$

$$- (d-b\xi^{1/2})^2 = x_2(T,\overline{u}) + 2bc\xi$$

$$x_3(T,u(\xi)) = \frac{1}{3}\left[T^3/2 - (b+c+d - c\xi^{1/2})^3 + (c+d - (b+c)\xi^{1/2})^3\right.$$

$$\left.- (d-b\xi^{1/2})^3\right] + \frac{k}{3}\left[(a+c\xi^{1/2})^3 + (b-a+(b-c)\xi^{1/2})^3\right.$$

$$\left.+ (c-b+a-b\xi^{1/2})^3 + (d-c+b-a)^3/2\right]$$

$$= x_3(T,\overline{u}) + [b^2c+bc^2]\xi^{1/2} + k[b^2c-bc^2]\xi^{1/2}$$

$$+ [bc^2+b^2c+2bcd]\xi + k[2abc-b^2c+bc^2] + O(\xi^{3/2})$$

$$= x_3(T,\overline{u}) + bc[b+c+2d+k(2a-b+c)]\xi$$

because, by (5.19), $b+c+k(b-c) = 0$. This yields (5.21).

REFERENCES

1. J. Diéudonne, Eléments d'analyse, Gauthier-Villars, Paris 1968.
2. H. Hermes and J. P. LaSalle, Functional analysis and time optimal control, Academic Press, New York 1969.
3. H. Hermes, Local controllability and sufficient conditions in singular problems, J. Diff. Eq. 20 (1976), pp. 213-232.
4. H. Hermes, Controlled stability, Ann. Mat. Pura ed Appl. CXIV (1977), pp. 103-119.
5. H. Hermes, Lie algebras of vector fields and local approximation of attainable sets, SIAM J. Control 16 (1978), pp. 715-727.
6. A. J. Krener, The high order maximal principle and its applications to singular extremals, SIAM J. Control, 15 (1977), pp. 256-293.
7. K. Kuratowski and C. Ryll-Nardzewski, A general theorem on selectors, Bull. Acad. Pol. Sc. Math. Astr. Phys., vol XIII, 6 (1965), pp. 397-403.
8. E. B. Lee and L. Markus, Foundations of optimal control theory, Wiley 1967.
9. H. J. Sussmann, A bang-bang theorem with bounds on the number of switchings, SIAM J. Control 17 (1979), pp. 629-651.

AB/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #2596	2. GOVT ACCESSION NO. AD-A136375	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A High-Order Test for Optimality of Bang-Bang Controls		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Alberto Bressan		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 5 - Optimization and Large Scale Systems
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE November 1983
		13. NUMBER OF PAGES 17
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Admissible variational family, high-order tangent vector

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

For control systems of the form $\dot{x} = X(x) + \sum_{i=1}^{k} Y_i(x)u_i$, a strengthened version of the classical Pontryagin Maximum Principle is proved. The necessary condition for optimality given here is obtained using functional analytic techniques and quite general high-order perturbations of the reference control. As shown by an example, our test is particularly effective when applied to bang-bang controls, a case where other high-order tests do not provide additional information.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)