On some recent developments on necessary conditions in Optimal Control

Piernicola Bettiol

University of Brest - UBO

Sample Days

Brest

8-10 September 2021

Outline of the talk

- Optimal control problems modeled as a differential inclusion: Clarke's necessary conditions
- Ioffe's refinement: a new Weierstrass condition
- An application: mixed constraint problems
- Examples
- Final remarks

Consider the optimal control problem:

$$(P) \begin{cases} \text{Minimize } g(x(0), x(T)) \\ \text{over } x(.) \in W^{1,1} \text{ satisfying} \\ \dot{x}(t) \in F(t, x(t)) \text{ a.e. } t \in [0, T] \\ (x(0), x(T)) \in C \ , \end{cases}$$

Data:
$g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ function - **cost function**
$C \subset \mathbb{R}^n \times \mathbb{R}^n$ closed set - **end-point constraint**
$F : [0, T] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ multifunction - **dynamic constraint**

(An alternative formulation of the classical optimal control problem associated with the Maximum Principle: a differential inclusion replaces the underlying controlled differential equation.)

$$(P)\begin{cases}\text{Minimize } g(x(0),x(T))\\ \text{over } x(.)\in W^{1,1} \text{ satisfying}\\ \dot{x}(t) \in F(t,x(t)) \text{ a.e.} t\in[0,T]\\ (x(0),x(T))\in C\,,\end{cases}$$

• $x(.)$ **Feasible** state trajectory :
a $W^{1,1}$ function s.t. $\dot{x}(t)\in F(t,x(t))$ a.e. and $(x(0),x(T))\in C$.

• $\bar{x}(.)$ is a $W^{1,1}$-**local minimizer**:
for some $\beta>0$,

$$g(x(0),x(T)) \ \geq\ g(\bar{x}(0),\bar{x}(T))$$

for all feasible state trajectories x s.t.

$$||x-\bar{x}||_{W^{1,1}}\leq\beta\,.$$

First order necessary conditions for problem (P)

A large literature:

- **Clarke**, *Optimization and Nonsmooth Analysis*, 1983.
- **Loewen and Rockafellar**, *Optimal control of unbounded differential inclusions*, SICON 1994.
- **Mordukhovich**, *Discrete approximations and refined Euler-Lagrange conditions for non-convex differential inclusions*, SICON 1995.
- **Ioffe**, *Euler-Lagrange and Hamiltonian formalisms in dynamic optimization*, SICON 1997.
- **Vinter and Zheng**, *The extended Euler-Lagrange condition for nonconvex variational problems*, SICON 1997.

. . . .

- **Clarke**, *Necessary Conditions in Dynamic Optimization.* AMS Memoirs 2005.
- **Ioffe**, *On Generalized Bolza Problem and Its Application to Dynamic Optimization*, JOTA 2019.

Some nonsmooth analysis constructs

$C \subset \mathbb{R}^n$ **closed set**, $\bar{x} \in C$,

The **proximal normal cone of C at $\bar{x}$**:

$N^P_C(\bar{x}) := \{\eta \in \mathbb{R}^n : \exists M > 0 \;\; s.t. \; \eta \cdot (x - \bar{x}) \leq M|x - \bar{x}|^2 \; for \; all \; x \in C\}.$

The (limiting) **normal cone of C at $\bar{x}$**:

$N_C(\bar{x}) := \{\lim_{i\to\infty} \eta_i \; : \; \eta_i \in N^P_C(x_i) \;\; and \; x_i \in C \; for \; all \; i, and \; x_i \to \bar{x}\}.$

Some nonsmooth analysis constructs...

$f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous function, take a point x s. t. $f(x) < +\infty$

The **proximal subdifferential of $f(.)$ at x**:

$$\partial_P f(x) := \{\zeta \in \mathbb{R}^n : \exists\ \sigma > 0, \epsilon > 0 \ \text{ s. t.}$$
$$f(y) - f(x) \geq \zeta \cdot (y - x) - \sigma|y - x|^2, \quad \forall y \in x + \epsilon\mathbb{B}\} .$$

The (limiting) **subdifferential of f at x**:

$$\partial f(x) := \{\lim_{i\to\infty} \zeta_i : \zeta_i \in \partial_P f(x_i), x_i \to x, f(x_i) \to f(x)\} .$$

The **partial subdifferential** $\partial_x f(\bar{x}, \bar{y})$: the subdifferential of $x \to f(x, \bar{y})$ at $\bar{x}$.

Mulifunctions - Lipschitz regularity

Take a **multifunction** $F : \mathbb{R}^k \rightsquigarrow \mathbb{R}^n$ and a point $\bar{x} \in \mathbb{R}^k$.
F is (locally) **Lipschitz continuous** w.r.t. the Hausdorff metric: there exist $\epsilon > 0$ and $k > 0$ s. t.

$$d_H(F(x), F(y)) \leq k|x - y| \ \text{ for all } x, y \in \bar{x} + \epsilon\mathbb{B}$$

Here $d_H(A, B)$ is the Hausdorff distance function:

$$d_H(A, B) := \max \left\{ \sup_{a \in A} d_B(a), \sup_{b \in B} d_A(b). \right\}$$

An equivalent statement of the condition is: there exist $\epsilon > 0$ and $k > 0$ s. t.

$$F(y) \subset F(x) + k|y - x|\mathbb{B} \ \text{ for all } x, y \in \bar{x} + \epsilon\mathbb{B}.$$

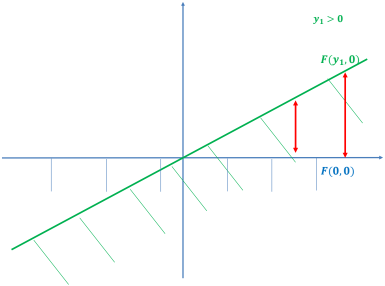
Rmk: this condition becomes overly restrictive for unbounded differential inclusions...

Example. Consider the multifunction $F : \mathbb{R}^2 \rightsquigarrow \mathbb{R}^2$ defined by

$$F(x_1, x_2) := \{(v_1, v_2) \in \mathbb{R}^2 \ : \ v_2 \ \leq \ x_1 v_1\}.$$

F is **NOT Lipschitz** continuous

$$d_H(F(x_1, x_2), F(y_1, y_2)) = +\infty \ \text{ for } \ y_1 \neq x_1,$$



Rmk: For two **unbounded** sets which are 'close' or 'regular' in an intuitive sense, the Hausdorff distance between them can be **very large**...

Pseudo-Lipschitz Continuity - Aubin Regularity

Take a multifunction $F : \mathbb{R}^k \rightsquigarrow \mathbb{R}^n$ and a point $(\bar{x}, \bar{v}) \in Gr\, F$. Take also numbers $\epsilon > 0$, $R > 0$ and $k \geq 0$. We say that F is **pseudo-Lipschitz** continuous near $(\bar{x}, \bar{v})$ (with parameters ϵ, R and k) if

$$F(y) \cap (\bar{v} + R\,\mathbb{B}) \subset F(x) + k|x - y|\mathbb{B}$$
$$\text{for all } x, y \in \bar{x} + \epsilon\,\mathbb{B}\,.$$

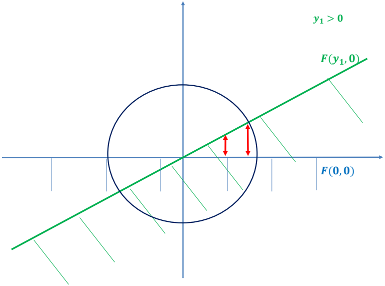
($Gr\, F$ is the graph of F)

Rmk: it allows to deal with unbounded velocity sets F.

Example... Consider the multifunction $F : \mathbb{R}^2 \rightsquigarrow \mathbb{R}^2$ defined by

$$F(x_1, x_2) := \left\{ (v_1, v_2) \in \mathbb{R}^2 \; : \; v_2 \; \leq \; x_1 v_1 \right\}.$$

F is **pseudo-Lipschitz** continuous near $(\bar{x} = 0, \bar{v} = 0)$ (with parameters $\epsilon = 1$, $R = 1$ and $k = 1$)



Standing assumptions

$\bar{x}$ is a **reference** state trajectory.

(G1) g is Lipschitz continuous on a neighborhood of $(\bar{x}(0), \bar{x}(T))$ and C is a closed set;

(G2) $F(t,x)$ is nonempty for each $(t,x) \in [0,T] \times \mathbb{R}^n$, Gr $F(t,.)$ is a closed set for each $t \in [0,T]$ and F is $\mathcal{L} \times \mathcal{B}^n$ measurable;

(G3) There exist $\epsilon > 0$ and a measurable function $R : [0,T] \to (0,\infty) \cup \{+\infty\}$ (a 'radius function') such that the following conditions are satisfied:

(a) (**Pseudo-Lipschitz Continuity**) There exists $k \in L^1$ s.t.

$$F(t,x') \cap (\dot{\bar{x}}(t) + R(t)\,\mathbb{B}) \subset F(t,x) + k(t)|x' - x|\mathbb{B},$$
$$\text{for all } x, x' \in \bar{x}(t) + \epsilon\mathbb{B}, \text{ a.e. } t \in [0,T];$$

(b) (**Tempered Growth**) There exist $r \in L^1(0,T)$, $r_0 > 0$ and $\gamma \in (0,1)$ s.t. $r_0 \le r(t)$, $\gamma^{-1} r(t) \le R(t)$ a.e. and

$$F(t,x) \cap (\dot{\bar{x}}(t) + r(t)\mathbb{B}) \neq \emptyset \text{ for all } x \in \bar{x}(t) + \epsilon\mathbb{B}, \text{ a.e. } t \in [0,T].$$

A simple consequence of **Clarke's 2005 Memoirs**:

Thm. 1 (The Euler Lagrange Inclusion)
Let $\bar{x}$ be a $W^{1,1}$ local **minimizer** for (P).
Then there exist an arc $p \in W^{1,1}([0,T];\mathbb{R}^n)$ and $\lambda \geq 0$, satisfying the following conditions:

(i) $(\lambda, p) \neq (0,0)$,

(ii) $\dot{p}(t) \in \text{co}\{\eta : (\eta, p(t)) \in N_{\text{Gr}\,F(t,.)}(\bar{x}(t), \dot{\bar{x}}(t))\}$ a.e. $t \in [0,T]$,

(iii) $(p(0), -p(T)) \in \lambda \partial g(\bar{x}(0), \bar{x}(T)) + N_C(\bar{x}(0), \bar{x}(T))$,

(iv) $p(t) \cdot \dot{\bar{x}}(t) \geq p(t) \cdot v$

for all $v \in F(t, \bar{x}(t)) \cap \left(\dot{\bar{x}}(t) + R(t)\mathring{\mathbb{B}}\right)$.

Weierstrass condition

Rmk: condition (iv) tells us that $v \to p(t) \cdot v$ is maximized at the optimal velocity $\dot{\bar{x}}(t)$, over the set

$$F(t, \bar{x}(t)) \cap \left(\dot{\bar{x}}(t) + R(t)\mathring{\mathbb{B}}\right),$$

in which R is the radius function of hypothesis (G3)

[Ioffe, JOTA 2019] provides a **refinement of the Weierstrass condition (iv) above**:

Thm. 2
Under the hypotheses of Thm. 1, the assertions of the theorem remain valid when (iv) is replaced by the **stronger** condition

$$\textbf{(iv)}' \quad p(t) \cdot \dot{\bar{x}}(t) \geq p(t) \cdot v, \ \text{for all } v \in \Omega_0(t), \ \text{a.e. } t \in [0, T]$$

The set of **regular admissible velocities at $\bar{x}(t)$**:

$$\Omega_0(t) := \{e \in F(t, \bar{x}(t)) \ : \ F(t, .) \text{ is pseudo-Lipschitz near } (\bar{x}(t), e)\} .$$

Rmk: Under the pseudo-Lipschitz hypothesis (G3)(a) of Thm. 1, we have

$$F(t, \bar{x}(t)) \cap \left(\dot{\bar{x}}(t) + R(t)\mathring{\mathbb{B}} \right) \subset \Omega_0(t) .$$

So the Ioffe refinement asserts that the $v \to p(t) \cdot v$ is **maximized over a larger set**.

Example 1.

$$(E1)\begin{cases} \text{Minimize } e \cdot x(1) \\ \text{over } x \in W^{1,1}([0,1];\mathbb{R}^n) \text{ such that} \\ \dot{x}(t) \in F(x(t)), \\ x(0) = 0, \end{cases}$$

in which $e := (1, 0, \ldots, 0)$ and

$$F(x) := \{0\} \cup \{v \ : \ |v| \geq 1 + |x|^{1/2}\}\,.$$

Take as nominal feasible F-trajectory $\bar{x} \equiv 0$.

Prop. 1 Concerning problem (E1),

(a): $\bar{x}$ is not a $W^{1,1}$ local minimizer.

(b): For any collection of radius function $R(t)$, $\epsilon > 0$ and integrable Lipschitz bound $k(t)$ such that the hypotheses of Thm. 1 are satisfied,

1. Conditions (i)-(iii) and (iv) of **Thm. 1** are satisfied.
2. Condition (iv)′ of **Thm. 2** (Weierstrass condition with refinement) is **NOT** satisfied.

An application: Mixed Constraint Problems

$$(M)\begin{cases} \text{Minimize } \; g(x(0),x(T)) + \int_0^T L(t,x(t),u(t))dt \\ \text{over } x \in W^{1,1}([0,T];\mathbb{R}^n) \text{ and meas.} \\ u:[0,T]\to\mathbb{R}^m \text{ such that} \\ \dot{x}(t) = f(t,x(t),u(t)), \; \text{a.e. } t\in[0,T], \\ h^{(1)}(t,x(t),u(t)) \le 0, h^{(2)}(t,x(t),u(t)) = 0, \\ \qquad\qquad \text{and } u(t)\in U, \; \text{a.e. } t\in[0,T], \\ (x(0),x(T)) \in C\,, \end{cases}$$

Data:

$g:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ **end-point cost**
$L:[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ **Lagrangian**
$C\subset\mathbb{R}^n\times\mathbb{R}^n$ closed set - **end-point constraint**
$f:[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ - **dynamic**
$h^{(1)}:[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^{\kappa_1}$, $h^{(2)}:[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^{\kappa_2}$ **mixed constraints**
$U\subset\mathbb{R}^m$ **control set**

First order necessary conditions for problem (M)

A large literature (and applications...):

- **Hestenes**, *Calculus of Variations and Optimal Control Theory*, 1966.
- **Dubovitskii, Milyutin**, *Theory of the principle of the maximum,* Methods of the Theory of Extremal Problems in Economics, 1981.
- **Dmitruk**, *Maximum principle for a general optimal control problem with state and regular mixed constraints*, Comput. Math. Model., 1993.

. . . .

- **Bonard, Faubourg, Launay, Trélat**, Optimal control with state constraints and the space shuttle reâentry problem, Journal of Dynamical and Control Systems, 2003

. . . .

- **Clarke, de Pinho**, *Optimal Control Problems with Mixed Constraints*, SICON, 2010.

[Clarke and de Pinho, SICON 2010] derived new general optimality conditions for mixed constraint optimal control problems.

Basic steps of their approach:

1. reduce the mixed constraint problem to a differential inclusion problem

2. apply the generalized Euler-Lagrange conditions of Clarke's AMS Memoirs 2005 (**Thm. 1**)

Admissible state-control pairs

Define

$$S(t) := \{(x,u) \,:\, h^{(1)}(t,x,u) \le 0, \quad h^{(2)}(t,x,u) = 0, u \in U\}$$

Ror given process $(\bar{x}, \bar{u})$ and parameters $\epsilon > 0$ and $R > 0$:

$$S^{\epsilon,R}(t) := \{(x,u) \in S(t) \,:\, |x - \bar{x}(t)| \le \epsilon, |u - \bar{u}(t)| \le R\}.$$

Define also the set of **admissible controls at state $\bar{x}(t)$**:

$$\Omega(t) := \{u \in \mathbb{R}^m \,:\, (\bar{x}(t), u) \in S(t)\}\,.$$

The ‘**Mangasarian-Fromovitz**’ condition:

given (t, x, u) such that $t \in [0, T]$ and $(x, u) \in S(t)$

$$(MF)_{t,x,u} : \begin{cases} \lambda_1 \in (\mathbb{R}^+)^{\kappa_1}, \lambda_2 \in \mathbb{R}^{\kappa_2}, \\ \lambda_1 \cdot h^{(1)}(t, x, u) = 0, \eta \in N_U(u) \\ \nabla_u(\lambda_1 \cdot h^{(1)} + \lambda_2 \cdot h^{(2)})(t, x, u) + \eta = 0 \end{cases}$$
$$\implies |(\lambda_1, \lambda_2)| = 0.$$

Standing assumptions

$(\bar{x}, \bar{u})$ is a **reference process**.
Assume that, for some $\epsilon > 0$ and positive measurable function $R \in L^\infty$, strictly bounded away from 0,

(H1) g is Lipschitz cont. on a neighb. of $(\bar{x}(0), \bar{x}(T))$;

(H2) for each $x \in \mathbb{R}^n$, $f(., x, .)$, $L(., x, .)$, $h^{(1)}(., x, .)$ and $h^{(2)}(., x, .)$ are $\mathcal{L} \times \mathcal{B}^m$ measurable; there exist integrable functions $k_x^{f,L}$ and $k_u^{f,L}$ such that, for a.e. $t \in [0, T]$,

$$|(f, L)(t, x_1, u_1) - (f, L)(t, x_2, u_2)| \leq k_x^{f,L}(t)|x_1 - x_2| + k_u^{f,L}(t)|u_1 - u_2|$$

for all (x_1, u_1) and (x_2, u_2) in a neighborhood of $S^{\epsilon, R(t)}(t)$;

(H3) For a.e. $t \in [0, T]$, $h^{(1)}(t, ., .)$ and $h^{(2)}(t, ., .)$ are continuously differentiable; there exists $k^h > 0$ s. t., for a.e. $t \in [0, T]$, $\forall (x_1, u_1)$, (x_2, u_2) in a neighb. of $S^{\epsilon, R(t)}(t)$;

$$|(h^{(1)}, h^{(2)})(t, x_1, u_1) - (h^{(1)}, h^{(2)})(t, x_2, u_2)| \leq k^h(|x_1 - x_2| + |u_1 - u_2|)$$

Standing assumptions...

(MF) $(MF)_{t,x,u}$ is satisfied, for every point (t, x, u) in closure $\{(t, x, u) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \, : \, (x, u) \in S^{\epsilon, R(t)}(t)\}$;

(H4) $\bar{u}$ is essentially bounded.

($\bar{u}(t)$ is interpreted as some version of the equivalence class of bounded, a.e. equal functions);

Thm. 3 (Clarke-de Pinho, 2010)
Let $(\bar{x}, \bar{u})$ be a $W^{1,1}$ local **minimizer** for (M).
Then there exist $p \in W^{1,1}([0,T], \mathbb{R}^n)$, $\lambda^0 \geq 0$ and integrable functions $\lambda^1 : [0,T] \to (\mathbb{R}^+)^{\kappa_1}$ and $\lambda^2 : [0,T] \to \mathbb{R}^{\kappa_2}$ such that

$$\lambda^1(t) \cdot h^{(1)}(t, \bar{x}(t), \bar{u}(t)) = 0, \text{ a.e. and}$$

(i) $(p, \lambda^0) \neq 0$,

(ii) $$(-\dot{p}(t), 0) \in \text{co}\, \partial_{x,u}\{p(t) \cdot f(t, \bar{x}(t), \bar{u}(t)) - \lambda^0 L(t, \bar{x}(t), \bar{u}(t))\} - \{0\} \times \text{co}\, N_U(\bar{u}(t)) - \lambda^1(t) \cdot \nabla_{x,u} h^{(1)}(t, \bar{x}(t), \bar{u}(t)) - \lambda^2(t) \cdot \nabla_{x,u} h^{(2)}(t, \bar{x}(t), \bar{u}(t)) \text{ a.e.,}$$

(iii) $(p(0), -p(T)) \in \lambda^0 \partial g(\bar{x}(0), \bar{x}(T)) + N_C(\bar{x}(0), \bar{x}(T))$,

(iv) $$p(t) \cdot f(t, \bar{x}(t), \bar{u}(t)) - \lambda^0 L(t, \bar{x}(t), \bar{u}(t)) \geq p(t) \cdot f(t, \bar{x}(t), u) - \lambda^0 L(t, \bar{x}(t), u)$$

for all $u \in \Omega(t) \cap \left(\bar{u}(t) + R(t)\mathring{\mathbb{B}}\right)$, a.e. $t \in [0,T]$.

Refined necessary conditions

What happens if we **apply Ioffe's refined condition** of Thm. 2?

The notion of '*regular admissible velocities*' of Thm. 2 gives rise in the mixed constraint setting to

the set of **regular admissible controls at state $\bar{x}(t)$**: for each $t \in [0, T]$

$\Omega_0(t) := \{u \in \mathbb{R}^m \,:\, (\bar{x}(t), u) \in S(t)$ and there exists $\rho > 0$ such that $(MF)_{t,x',u'}$ is satisfied, for all (x', u') in a neighborhood of $S(t) \cap \Big((\bar{x}(t), u) + \rho\mathbb{B} \times \rho\mathbb{B} \Big)$ relative to $S(t)\,\}$.

Refined necessary conditions...

Thm. 4
Under the hypotheses of Thm. 3, the assertions of Thm. 3 remain valid when the Weierstrass condition (iv) is replaced by the refined Weierstrass condition:

(iv)′ $p(t) \cdot f(t, \bar{x}(t), \bar{u}(t)) - \lambda^0 L(t, \bar{x}(t), \bar{u}(t)) \geq p(t) \cdot f(t, \bar{x}(t), u) - \lambda^0 L(t, \bar{x}(t), u)$

$$\text{for all } u \in \Omega_0(t), \quad \text{a.e. } t \in [0, T].$$

Rmk: Under the hypothesis (MF) of Thm. 4 we have

$$\boxed{\Omega(t) \cap (\bar{u}(t) + R(t))\mathring{\mathbb{B}} \subset \Omega_0(t)}.$$

[P.B.-R. Vinter, IEEE CDC 2021]

Example 2

$$(E2) \begin{cases} \text{Minimize } -x(T) - \int_0^T (0 \vee (u(t)-\pi))^2 dt \\ \text{over } (x,y) \in W^{1,1}([0,T];\mathbb{R}^2) \\ \quad \text{and meas } (u,v) : [0,T] \to \mathbb{R}^2 \\ \text{such that,} \\ (\dot{x}(t), \dot{y}(t)) = (\sin(u(t)), v(t)), \quad \text{a.e. } t \in [0,T], \\ \sin(u(t)) - y(t) \le 0 \text{ and } |v(t)| \le 1 \quad \text{a.e. } t \in [0,T], \\ x(0) = y(0) = 0. \end{cases}$$

Take as **nominal state-control** functions:

$$(\bar{x}(t), \bar{y}(t)) \equiv (\frac{1}{2} \times t^2, t), \text{ for } t \in [0, \pi/8]$$

$$(\bar{u}(t), \bar{v}(t)) = (\arcsin(t), 1), \text{ for } t \in [0, \pi/8].$$

The problem has one, time invariant, mixed inequality constraint

$$h^{(1)}(y,u) := \sin(u) - y \le 0\,.$$

Prop. 2
Concerning problem (E2), For any radius function $R(.)$ and parameter $\epsilon > 0$ such that the hypotheses of Thm. 3 are satisfied (with $((\bar{x}, \bar{y}), (\bar{u}, \bar{v}))$),

1. Conditions (i)-(iii) and (iv) of Thm. 3 are satisfied.
2. Condition (iv)$'$ of Thm. 4 (Weierstrass condition with refinement) are **NOT** satisfied.

$\Rightarrow ((\bar{x}, \bar{y}), (\bar{u}, \bar{v}))$ **is NOT a minimizer!**

Final remarks

1) Ioffe's 2019 paper provides a refinement improving the Weierstrass
2) We have provided an example of a optimal control problem for differential inclusions, where the new information in the refined Weierstrass condition is used to establish that a certain extremal (a feasible trajectory satisfying the necessary conditions of Thm.1) is not optimal.
3) We have shown that using the necessary conditions of Thm. 2 (that include the refined Weierstrass condition) we are able to derive improved necessary conditions for the mixed constraint problem.
4) We have provided an example of a mixed constraint optimal control problem for controlled differential equations, where the refined Weierstrass condition in Thm. 4 is used to establish that a certain extremal is not optimal.