A NEW VERSION OF NECESSARY CONDITIONS FOR OPTIMAL CONTROL PROBLEMS WITH DIFFERENTIAL ALGEBRAIC EQUATIONS*

IGOR KORNIENKO†, MATTHIAS GERDTS‡, AND MDR DE PINHO§

Abstract. Appealing to recent results for nonsmooth mixed constrained problems we derive new variants of necessary optimality conditions for optimal control problems involving differential algebraic equations. The analysis is quite suitable for index one problems with no need for the introduction of implicit functions. It is also suitable to some higher index problems.

Key words. optimal control, differential and algenraic equations, nonsmooth analysis, necessary conditions

AMS subject classifications. 49K21, 49J15 , 49J21

1. Introduction.

Although the literature on optimal control problems is quite vast, derivation of optimality conditions for problems involving differential and algebraic equations (DAE) have only recently deserved some attention (see, for example, [5, 6, 7, 9, 16, 18]) and remains a challenging area. Nowadays adopting models of this kind is routine in in areas like robotics, economics and process systems engineering.

Throughout this paper we focus on the following problem involving DAE's:

$$(P)\quad \begin{cases} \text{Minimize} & l(x(a),x(b)) \\ \text{subject to} & \\ & \dot{x}(t)=f(t,x(t),y(t),u(t)) \text{ a.e.} \\ & 0=g(t,x(t),y(t),u(t)) \text{ a.e.} \\ & u(t)\in U \text{ a.e.} \\ & (z(a),z(b))\in E \end{cases}$$

where $l:\mathbf{R}^n\times\mathbf{R}^n\to\mathbf{R}$, $f:[a,b]\times\mathbf{R}^n\times\mathbf{R}^m\times\mathbf{R}^k\to\mathbf{R}^n$, $g:[a,b]\times\mathbf{R}^n\times\mathbf{R}^m\times\mathbf{R}^k\to\mathbf{R}^m$, $U\subset\mathbf{R}^k$ is compact and $E\subset\mathbf{R}^n\times\mathbf{R}^n$ is a closed set. Typically the pair (x,y) is the state divided into "slow" variable x and "fast" variables y which can respond instantaneously to changes in control u; x has a derivative defined by the differential equation while y does not.

Usually necessary optimality conditions for (P) are derived following one of the two approaches: y is treated as a control or y is seen as a component of the state constraint. In the first case, the control is $v=(y,u)$, with y unconstrained while u takes values in a certain control set U. This amounts to consider (P) as a mixed constrained optimal control problem. In the second case, the state is considered to be $z=(x,y)$. For (P) this leads to assemble both differential and algebraic equation as a single equation of the form $h(t,z,u,\dot{z})=0$ where

$$h(t,z,u,\dot{z})=E\dot{z}(t)-\begin{bmatrix} f(t,z(t),u(t)) \\ -b(t,z(t),u(t)) \end{bmatrix}=0 \tag{1.1}$$

*This work has been supported by the European Union Seventh Framework Programme FP7-PEOPLE-2010-ITN under grant agreement 264735-SADCO and FCT Project PTDC/EEA-CRO/116014/2009.

†Universidade do Porto, Faculadade de Engenharia, DEEC, Portugal(`igor@fe.up.pt`).

‡Institut fr Mathematik und Rechneranwendung Universitt der Bundeswehr Mnchen, Germany(`matthias.gerdts@unibw.de`.

§Universidade do Porto, Faculadade de Engenharia, DEEC, Portugal

where $E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$, with I begin the identity matrix. It turns out that (1.1) is a special case of of a problem treated covered by Theorem 6.1 in [3].

In this paper we illustrate how nonsmooth developments reported in [3] for mixed constrained problems can be applicable to (P). We treat y as a control component. For problems with the so called index one DAE's, we get necessary conditions for problems with DAE's that, in some sense, improve and and extend existing results as, for example, those in [9],[5] and [16]. Furthermore, we show that [3] is also of help as far as maximum principles for some class of higher index problems.

Theorem 6.1 in [3] concerns DAE problems when y is viewed as a component of the state. Although it would be of interest to compare the information provided by this theorem to some literature (see, for example, [7], [18] and [14]), we do not dwell on such subject in this work for simplicity of exposition.

Optimal control problems with DAE's are frequently coupled with inequalities mixed constraints and pure state constraints both in the form of equalities and inequalities see for example [6] or [7] (see also [8])). We do not treat such general problems here. A simple inspection of our central tool of analysis, [3], will convince the reader that we can handle mixed inequalities constraints but not pure state constraint. It is our believe that future extensions of Theorem 2.1 in [3] to cover problems with pure state constraints may later come to help in this respect.

The interest of our findings does not dwell simply on the nonsmoothness of the data. Noteworthy, when when some smoothness is imposed on b, our results are directly comparable to many in the literature and, in some situations, hold under weaker assumptions. Our approach can be extended without much ado to cover situations where inequality mixed constraints of the form $g(t, x, y, u) \leq 0$ are added to DAE but we do not develop here.

2. Preliminaries. If g is a vector, $g \in \mathbf{R}^m$, the inequality $g \leq 0$ is interpreted componentwise.

We will denote by $\mathbf{B}$ the closed unit ball centered at the origin regardless of the dimension of the underlying space. Also $|\cdot|$ is the Euclidean norm or the induced matrix norm on $\mathbf{R}^{p\times q}$.

Take any $A \subset \mathbf{R}^n$. Then the *Euclidean distance function* with respect to A is defined as

$$d_A : \mathbf{R}^k \to \mathbf{R}, \qquad y \to d_A(y) = \inf \{|y - x| : x \in A\}.$$

Consider now a function $h : [a, b] \to \mathbf{R}^p$. We say that $h \in W^{1,1}([a, b]; \mathbf{R}^p)$ if and only if it is absolutely continuous; in $h \in L^1([a, b]; \mathbf{R}^p)$ iff h is integrable; and in $h \in L^\infty([a, b]; \mathbf{R}^p)$ iff it is essentially bounded. The norm of $L^1([a, b]; \mathbf{R}^p)$ is denoted by $\|\cdot\|_1$ and the norm of $L^\infty([a, b]; \mathbf{R}^p)$ is $\|\cdot\|_\infty$.

As it is clear from the introduction we shall make use of standard concepts from nonsmooth analysis. The basic concepts of nonsmooth analysis are well known so we refrain from stating them here. Instead we refer the reader to, for example, [1], [2], [17], [15] and [12]. Next we introduce the notation of some nonsmooth concepts we will use throughout this paper.

Let $A \subset \mathbf{R}^n$ to be a closed set with and consider $x^* \in A$. The *limiting normal cone to A at x^** (also known as *Mordukhovich* normal cone) is denoted by $N^L_A(x^*)$ while the *Clarke* normal cone is $N^C_A(x^*)$.

Take a lower semicontinuous function $f : \mathbf{R}^k \to \mathbf{R} \cup \{+\infty\}$ and a point $x^* \in \mathbf{R}^k$ where $f(x^*) < +\infty$. Then the *limiting subdifferential* of f at x^* is denoted by $\partial^L f(^*)$.

Recall that when the function f is Lipschitz continuous near x, the convex hull of the limiting subdifferential, co $\partial^L f(x)$, coincides with the *(Clarke) subdifferential*, denoted here by $\partial^C f(x)$.

3. Index One: Nonsmooth Case. We now concentrate on (P) and we explore the consequences of some results in [3] when y is treated as a control.

For simplicity of the exposition and analysis we impose assumptions on the data that are rather stronger than those in [3] but are still satisfied for a large class of real problems.

First we enforce the following *basic hypotheses*: the function l is locally Lipschitz, E is a closed set and $(t,(x,y,u)) \to f(t,(x,y,u)), g(t,(x,y,u))$ are $\mathcal{L} \times \mathcal{B}$ (relative to the σ-field generated by the product of Lebesgue measurable subsets in $\mathbf{R}$ and Borel measurable subsets in $R^n \times \mathbf{R}^m \times \mathbf{R}^k$).

For (P) we consider (x,y,u) to be a triple where both y and u are considered to be *essentially bounded* functions[1] and x, the state, an absolutely continuous function. As usual, a process is called admissible when it satisfies all the constraints of the problem. We say that an admissible process for (P) is (x^*,y^*,u^*) is a $W^{1,1}$-*local minimum*[2] when it minimizes the cost

$$J(x,y,u) := l(x(a),x(b))$$

over all the admissible processes (x,y,u) such that

$$|x(t) - x^*(t)| \leq \varepsilon, \quad \int_0^1 |\dot{x}(t) - \dot{x}^*(t)|\, dt \leq \varepsilon$$

for some $\varepsilon > 0$.

We now introdue some sets. Let

$$S(t) := \{(x,y,u) : g(t,x,y,u) = 0,\ u \in U\},$$

and

$$S^*_\varepsilon(t) := \{(x,y,u) \in S(t) :\ |x - x^*(t)| \leq \varepsilon\}.$$

Take a reference triple (x^*,y^*,u^*) and a parameter ϵ. Consider a function $\psi : [a,b] \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^k \to \mathbf{R}^N$. We say that ψ satisfies $[\mathbf{L}^*]$ if the following condition holds:

[$\mathbf{L}^*$] There exists a constant k_ψ such that, for almost every $t \in [a,b]$, for every (x_i,y_i,u_i) with $|x_i - x^*(t)| \leq \varepsilon$, we have

$$|\psi(t,x_1,y_1,u_1) - \psi(t,x_2,y_2,u_2)| \leq k_\psi[|x_1 - x_2| + |y_1 - y_2| + |u_1 - u_2|].$$

We shall impose $\mathbf{L}^*$ on both f and g. As for the algebraic equation and set control constraints we assume the following.

[**A1**] There exists constant M such that, for almost every $t \in [a,b]$, all $(x,y,u) \in S^*_\varepsilon(t)$, all $\lambda \in \mathbf{R}^m$, all $\xi \in N^L_U(u)$ we have

$$(3.1) \qquad (\alpha,\beta_1,\beta_2 - \xi) \in \partial^L_{x,y,u}\langle \lambda, g(t,x,y,u)\rangle \implies |\lambda| \leq M|(\beta_1,\beta_2)|.$$

[1]It is possible to work with measurable functions but the analysis is simpler if L^∞ functions are considered

[2]In [3] *local minimum of radius* R are considered. Taking $R = +\infty$, this reduces to our case. The $W^{1,1}$-*local minimum* is sometimes called the weak local minimum.

A simple application of Theorem 7.1 in [3] yields the following result:

THEOREM 3.1. (**Nonsmooth Case**) Let (x^*, y^*, u^*) be a $W^{1,1}$-local minimum for problem (P). If the basic assumptions are satisfied, f and g satisfy [$\mathbf{L}^*$] and [**A1**] holds, then there exist $p \in W^{1,1}([a,b];\mathbf{R}^n)$ and a scalar $\lambda_0 \geq 0$ satisfying the *nontriviality condition*:

$$||p||_\infty + \lambda_0 > 0, \tag{3.2}$$

the *Euler adjoint inclusion*: for almost every $t \in [0,1]$

$$(-\dot{p}(t), 0, 0) \in \partial^C_{x,y,u} \langle p(t), f(t, x^*(t), y^*(t), u^*(t))\rangle - N^C_{S(t)}(x^*(t), y^*(t), u^*(t)) \tag{3.3}$$

the global *Weierstrass condition*: for almost every $t \in [0,1]$ and all $(x^*(t), y, u) \in S(t)$

$$\langle p(t), f(t, x^*(t), y, u)\rangle \leq \langle p(t), f(t, x^*(t), y^*(t), u^*(t))\rangle, \tag{3.4}$$

and the *transversality condition*:

$$(p(a), -p(b)) \in N^L_E(x^*(a), x^*(b)) + \lambda_0 \partial^L l(x^*(a), x^*(b)). \tag{3.5}$$

Theorem 3.1 is a nonsmooth version of the maximum principle for (P) when y is seen as a control. Although of interest in itself, its applicability may be problematic because of the nature of (3.3); the set $N^C_{S(t)}(x^*(t), y^*(t), u^*(t))$ may be hard to handle. Fortunately if some differentiability is assumed on function g, then this normal cone can de dealt with in a simple way as we explain next.

4. Index one: Differential Case. Let us first assume that b satisfies [$\mathbf{L}^*$] and g is also strict differentiable at $(x^*(t), y^*(t), u^*(t))$ for almost every t. Then from nonsmooth calculus (see for example Proposition 4.1 in [3]), we know that, for any measurable function $\chi : [a,b] \to \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^k$ such that $\chi(t) \in N^C_{S(t)}(x^*(t), y^*(t), u^*(t))$ almost everywhere, there exist measurable functions $\lambda : [a,b] \to \mathbf{R}^m$, $\xi : [a,b] \to \mathbf{R}^k$ such that $\xi(t) \in N^C_U(u^*(t))$ and

$$\chi(t) = g_{x,y,u}(t, x^*(t), y^*(t), u^*(t))^T \lambda(t) + (0, 0, \xi(t)). \tag{4.1}$$

Taking into account the sum rule of nonsmooth calculus we deduce from the above that (3.3) in Theorem 4.1 is, for almost every t, of the form

$$\begin{aligned}(-\dot{p}(t), 0, \xi(t)) \in\ & \partial^C_{x,y,u}\langle p(t), f(t, x^*(t), y^*(t), u^*(t))\rangle \\ & - g_{x,y,u}(t, x^*(t), y^*(t), u^*(t))^T \lambda(t)\end{aligned} \tag{4.2}$$

where $\xi(t) \in N^C_U(u^*(t))$ almost everywhere.

Let us go a step further and **replace [$\mathbf{L}^*$] by the following assumption**:

[**CD**] The function $(x, y, u) \to g(t, x, y, u)$ is continuous differentiable for almost every $t \in [a,b]$.

Additionally let us consider the following assumption:

[**I1**] There exist constants c and m_g such that for almost every $t \in [a,b]$, all $(x, y, u) \in S^*_\varepsilon(t)$ we have

$$\det g_y(t, x, y, u) g_y(t, x, y, u)^T > c, \quad \text{and } |[g_y(t, x, y, u) b_y(t, x, y, u)^T]^{-1}| \leq m_g.$$

We claim that [[**I1**]] implies [[**A1**]]. To show such implication take any $\lambda \in \mathbf{R}^m$ and observe that

$$\partial^L_{x,y,u}\langle \lambda, b(t,x,y,u)\rangle = g_{x,y,u}(t,x,y,u)^T\lambda.$$

If $(\alpha, \beta_1, \beta_2 - \xi) \in \partial^L_{x,y,u}\langle \lambda, g(t,x,y,u)\rangle$ and $\xi \in N^C_U(u)$, then we have

$$\alpha = g_x(t,x,y,u)^T\lambda, \tag{4.3}$$

$$\beta_1 = g_y(t,x,y,u)^T\lambda, \tag{4.4}$$

$$\beta_2 = g_u(t,x,y,u)^T\lambda + \xi. \tag{4.5}$$

From [**I1**] we duduce that

$$\lambda = [g_y(t,x,y,u)g_y(t,x,y,u)^T]^{-1}\beta_1$$

and

$$|\lambda| \le m_g|\beta_1| \le m_g|(\beta_1,\beta_2)|.$$

This means that under our differentiable assumption [**CD**], [**I1**] implies [**A1**].

Recall that [**I1**] characterizes (P) as a problem involving **index 1** differential algebraic equations.

It is easy to deduce the following:

Corollary 4.1. (**Index 1 case**) Let (x^*, y^*, u^*) be a $W^{1,1}$-local minimum for problem (P). If the basic assumptions are satisfied, f satisfies [$\mathbf{L}^*$], g satisfies [**CD**] and [**I1**], then there exist $p \in W^{1,1}([a,b];\mathbf{R}^n)$ and a scalar $\lambda_0 \ge 0$ satisfying the *nontriviality condition*:

$$||p||_\infty + \lambda_0 > 0,$$

the *Euler adjoint inclusion*:

$$(-\dot p(t), 0, \xi(t)) \in \tag{4.6}$$

$$\partial^C_{x,y,u}\langle p(t), f(t,x^*(t),y^*(t),u^*(t))\rangle - g_{x,y,u}(t,x^*(t),y^*(t),u^*(t))^T\lambda(t) \quad \text{a.e.},$$

the global *Weierstrass condition*: for almost every $t \in [0,1]$ and all $(x^*(t), y, u)$ such that $u \in U$ and $g(t, x^*(t), y, u) = 0$

$$\langle p(t), f(t,x^*(t),y,u)\rangle \le \langle p(t), f(t,x^*(t),y^*(t),u^*(t))\rangle \quad \text{a.e.},$$

and the *transversality condition*:

$$(p(a), -p(b)) \in N^L_E(x^*(a), x^*(b)) + \lambda_0\partial^L l(x^*(a), x^*(b)).$$

where ξ is a measurable function satisfying $\xi(t) \in N^C_U(u^*(t))$ almost everywhere.

Important Note: It is also a simple matter to see that under the assumptions of the above theorem there exists a constant k_b such that

$$|\lambda(t)| \le m_g k_g L_f |p(t)| \quad \text{a.e.}$$

Also, the above result covers situations where $U = \mathbf{R}^k$. In such cases, the necessary consitions hold with $N_U^C(u^*(t)) = \{0\}$.

Remark: Corollary 4.1 provides a maximum principle for optimal control problems involving index 1 problems easily comparable with most of the literature. A special feature of our findings is that there is no explicit reference to implicit functions (see for example [18]). Moreover, although smoothness of g is assumed, a certain degree of nonsmoothness is allowed.

Results in the same vein were obtained in [5] and but now no convexity is assumed whereas in [5] is. Comparing Corollary 4.1 with [9] and [16] it is also clear that our result covers more general problems.

5. Higher Index Problems. Before proceeding let us recall that we are unable to deal with situations where the algebraic equation is reduced to $g(t, x(t)) = 0$, i.e., to equalities pure state constraints; such case is amply treated as a higher index case problem in the literature (see for example [7]). However our approach can still cover some classes of problems with index higher than 1. Indeed, for problems of index higher than one satisfying [**A1**], Theorem 4.1 can still provide necessary conditions for the smooth and nonsmooth case.

In what follows we analyse a particular case when $U = \mathbf{R}^k$. When the control u is unsconstrained there is no distinct treatment between y and u.

Let us keep assumption [**CD**] and focus on problems where

$$g(t, x, y, u) = \begin{pmatrix} g^1(t, x, y, u) \\ g^2(t, x, u) \end{pmatrix}.$$

Here $g^1(t, x, y, u) \in \mathbf{R}^{m_1}$, $g^2(t, x, y, u) \in \mathbf{R}^{m_2}$, $m_1 + m_2 = m$ and $m_1, m_2 \geq 1$. It is easy to see that [**I1**] does not hold since g_y is not of maximal rank. The optimal control problem involving such function g is now of higher index.

It is a simple matter to see that [**A1**] may hold. For example [**A1**] is satisfied when, for example, the lost of rank of g_y is compensated by information on g_u. In such situations it is easy to obtain necessay conditions in the spirit of Corollary 4.1 (we refrain from do so here).

6. A Companion Result. It is clear from our findings that we treat (y, u) as the control partioned into components where at most one of the components, u, is subject to set control constraints. Clearly this approach is well known for those familiar with optimal mixed state-control constrained problems (see for example, [10]). However, in some literature such approach implies different hypotheses in the y and the u components; this is the case with [5]). Theorem 3.2 in [3] is of help in this regard.

To set our scenary we start by stating alternative assumptions to [$\mathbf{L}^*$] and [**A2**]. Set

$$S(t, u) := \{(x, y) : g(t, x, y, u) = 0\}, \quad S_\varepsilon^*(t, u) := \{(x, y) \in S(t, u) : \ |x - x^*(t)| \leq \varepsilon\}.$$

[$\mathbf{L_2^*}$] There exists a constant k_ψ such that, for almost every $t \in [a, b]$, for every $u \in U$, every (x_i, y_i) with $|x_i - x^*(t)| \leq \varepsilon$, we have

$$|\psi(t, x_1, y_1, u) - \psi(t, x_2, y_2, u)| \leq k_\psi[|x_1 - x_2| + |y_1 - y_2|].$$

[**A2**] There exists a constant M such that, for almost every $t \in [a, b]$, every $u \in U$, $\lambda \in \mathbf{R}^N$, $\xi \in N_{U(t)}^L(u)$ and $(x, y) \in S_\varepsilon^*(t, u)$ we have

$$(\alpha, \beta) \in \partial_{x,y}^L \langle \lambda, g(t, x, v, u) \rangle \implies |\lambda| \leq M|\beta|.$$

A noteworthy aspect of imposing $[\mathbf{L_2^*}]$ on f and g is that we impose Lipchitz continuity with respect to y but not even continuity is imposed with respect to u.

Assume the basic hypotheses, that $[\mathbf{L_2^*}]$ is satisfied by both f and g, that [**A2**] holds and that $S(t,u)$ is closed. Then following the approach of the previous subsections we deduce a nonsmooth Theorem in the vein of Theorem 3.1.

For the differential case consider now

[**CD2**] The function $(x,y) \to g(t,x,y)$ is continuous differentiable for almost every $t \in [a,b]$.

Take $(\alpha,\beta) \in N^C_{S(t,u)}(x,y)$. If [**CD2**] holds, then there exists $\lambda : [a,b] \to \mathbf{R}^m$ such that

$$(\alpha,\beta) = g_{x,y}(t,x,y,u)^T\lambda(t).$$

It is now a simple matter to see that the following holds.

THEOREM 6.1. *Let (x^*,y^*,u^*) be a $W^{1,1}$-local minimum for problem (P). If the basic assumptions are satisfied, f and g satisfy $[\mathbf{L_2^*}]$, the set $S(t,u)$ is closed and [**A2**] and [**CD2**] hold, then there exist $p \in W^{1,1}([a,b];\mathbf{R}^n)$ and a scalar $\lambda_0 \geq 0$ satisfying (3.2) and (3.5) as well as the* Euler adjoint inclusion*:*

$$(-\dot{p}(t),0) \in \partial^C_{x,y}\langle p(t), f(t,x^*(t),y^*(t),u^*(t))\rangle - g_{x,y}(t,x^*(t),y^*(t),u^*(t))^T\lambda(t) \quad a.e.,$$

and the global Weierstrass condition*: for all $u \in U$ and $(x^*(t),y) \in S(t,u)$*

$$\langle p(t), f(t,x^*(t),y(t),u)\rangle \leq \langle p(t), f(t,x^*(t),y^*(t),u^*(t))\rangle \quad a.e.,$$

Application of the previous result to index one case is now easy to obtain.

7. Conclusions. We exploit results in [3] so as to obtain necessary optimality conditions for optimal control problems involving differential algebraic equations with a special semi-explicit structure and under some smoothness assumptions. We mainly focus on problems where the "fast" variable y is treated as a control. Although are results are not as general as we would like, we hope that future work extending results in [3] to cover problems with pure state constraints may be of help in the future.

REFERENCES

[1] F. CLARKE, *Optimization and Nonsmooth Analysis*, (1983), John Wiley, New York.

[2] F. CLARKE, Y. LEDYAEV, R. J. STERN, P. R. WOLENSKI, *Nonsmooth Analysis and Control Theory*, (1998), Springer-Verlag, New York.

[3] F. CLARKE, M. D. R. DE PINHO, *Optimal control problems with mixed constraints*, SIAM J. Control Optim. 48 (2010), pp. 4500–4524.

[4] F. CLARKE, Y. LEDYAEV, M. D. R. DE PINHO, *An extension of the schwarzkopf multiplier rule in optimal control*, 2011, SIAM J. Control Optim. 49 (2011), pp. 599-610.

[5] MDR DE PINHO AND R. VINTER, *Necessary conditions for optimal control problems involving nonlinear differential algebraic equations*, J. Math. Anal. Appl., 21 (1997), pp. 493–516.

[6] MATTHIAS GERDTS, *Local minimum principle for optimal control problems subject to differential-algebraic equations of index two*, J. Optim. Theory Appl., 130 (2006), pp. 441–460.

[7] MATTHIAS GERDTS, *Optimal Control of Odes and Daes*, (2012) Walter de Gruyter.

[8] MATTHIAS GERDTS, *Optimal Control of Ordinary Differential Equations and Differential-Algebraic Equations*, Habilitation Thesis (2006), Department of Mathematics University of Bayreuth, Germany.

[9] E. N . DEVDARYANI, Y. S. LEDYAEV Y. S, *Maximum principle for implicit control systems*, Appl. Math. Optim., 40 (1996), pp. 79–103.

[10] Dmitruk A. V. 1993. *Maximum principle for the general optimal control problem with phase and regular mixed constraints*. Comput. Math. Model., 4, 364–377.

[11] Philip Loewen, *Optimal Control via Nonsmooth Analysis* (1993), CRM, American Mathematical Society.

[12] B. Mordukhovich, *Variational analysis and generalized differentiation. Basic Theory*, Fundamental Principles of Mathematical Sciences 330 (2006), Springer-Verlag, Berlin.

[13] R. Pytlak, *Optimal Control of DifferentialAlgebraic Equations of Higher Index, Part 1: First-Order Approximations*, J Optim Theory Appl (2007) 134, pp. 6175.

[14] R. Pytlak, *Optimal Control of DifferentialAlgebraic Equations of Higher Index, Part 2: Necessary Optimality Conditions*, J. Optim. Theory Appl. (2007) 134, pp. 77-90

[15] R. T. Rockafellar B.and Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317 (1998), Springer-Verlag, Berlin.

[16] T. Roubicek and M. Valasek, *Optimal control of causal differential-algebraic systems*, (2002) J. Math. Anal. Appl., 269, pp. 616–641.

[17] R. Vinter, *Optimal Control*, 2000, Birkhäuser, Boston.

[18] Peter Kunkel and Mehrmann Volker, *Optimal control for unstructured nonlinear differential-algebraic equations of arbitrary index*, Math. Control Signals Syst., 22 (2008) pp. 227–269.