Necessary conditions for a nonclassical control problem with state constraints ⋆

Chems Eddine Arroud * **Giovanni Colombo** **

* *Department of Mathematics, Jijel University, Jijel, Algeria, and Mila University Center, Mila, Algeria (e-mail: arroud.math@gmail.com)*
** *Dipartimento di Matematica "Tullio Levi-Civita", University of Padova, Padova, Italy (e-mail: colombo@math.unipd.it).*

Abstract: We consider the problem of minimizing the cost $h(x(T))$ at the endpoint of a trajectory x subject to the finite dimensional dynamics

$$\dot{x} \in -N_C(x) + f(x,u), \quad x(0) = x_0,$$

where N_C denotes the normal cone to the convex set C. Such differential inclusion is termed, after Moreau, *sweeping process*. We label it as a "nonclassical" control problem with state constraints, because the right hand side is discontinuous with respect to the state, and the constraint $x(t) \in C$ for all t is implicitly contained in the dynamics.
We prove necessary optimality conditions in the form of Pontryagin Maximum Principle by requiring, essentially, that C is independent of time. If the reference trajectory is in the interior of C, necessary conditions coincide with the usual ones. In the general case, the adjoint vector is a BV function and a signed vector measure appears in the adjoint equation.

Keywords: Moreau's sweeping process, Optimal control, Pontryagin Maximum Principle.

1. INTRODUCTION

The *sweeping process* was introduced by Moreau in the Seventies as a model for dry friction and plasticity (see Moreau (1974)) and later studied by several authors. In its perturbed version, it features the differential inclusion

$$\dot{x}(t) \in -N_{C(t)}(x(t)) + f(x(t)), \quad t \in [0,T] \tag{1}$$

coupled with the initial condition

$$x(0) = x_0 \in C(0). \tag{2}$$

Here $C(t)$ is a closed moving set, with normal cone $N_{C(t)}(x)$ at $x \in C(t)$. The space variable, in this paper, belongs to $\mathbb{R}^n$. If $C(t)$ is convex, or mildly non-convex (in a sense that will not be made precise here), and is Lipschitz as a set-valued map depending on t, and the perturbation f is Lipschitz as well, then it is well known that the Cauchy problem (1), (2) admits one and only one Lipschitz solution (see, e.g., Thibault (2003)). Observe that the state constraint $x(t) \in C(t)$ for all $t \in [0,T]$ is *built in the dynamics*, being $N_{C(t)}(x)$ empty if $x \notin C(t)$: should a solution $x(\cdot)$ exist, then automatically $x(t) \in C(t)$ for all t. If a control parameter u appears within f, then one is lead to study problems of the type

$$\dot{x}(t) \in -N_{C(t)}(x(t)) + f(x(t), u(t)), \quad u(t) \in U \tag{3}$$

subject to (2), aiming, for example, at

$$\text{minimizing } h(x(T)), \tag{4}$$

the final cost h being smooth. There is a clear difference with classical control problems with state constraints (see, e.g., Vinter (2000)), where the constraint does not appear explicitly in the dynamics: in this case the right hand side of the dynamics is not Lipschitz with respect to the state variable, but indeed has only closed graph. This fact is a source of major difficulties in deriving necessary optimality conditions for (3), (4).

In recent years (see, e.g., Bagagiolo (2002), Gudovich et al. (2011), Brokate et al. (2013), Colombo et al. (2016), Colombo et al. (2016), Arroud et al. (2016), and Cao et al. (2017), and references therein) some papers dealing with control problems involving the sweeping process were published, the control appearing in the perturbation f and/or in the moving set C. Several necessary conditions were established, under different kinds of assumptions, or a Hamilton-Jacobi characterization of value function was proved. The present paper is devoted to prove a result inspired by Arroud et al. (2016) and Brokate et al. (2013). More precisely, we prove necessary conditions of Pontryagin maximum principle type for (4) subject to (3) and (2), the control appearing only within f, in the case where $C(\cdot)$ is constant, smooth and convex (see Theorems 2 and 3). The case where C satisfies milder convexity assumptions and is not necessarily constant was treated in Arroud et al. (2016) with an extra assumption, while Brokate et al. (2013) contains results for a particular control problem involving a fixed smooth and *uniformly convex* set C. More preccisely, differently from Colombo et al. (2016) and Cao et al. (2017), where discrete approximations are used, in both Brokate et al. (2013) and Arroud et al. (2016) the authors use a penalization technique. The classical Moreau-Yosida regularization allows in Arroud et al.

⋆ The second author is partially supported by Padova University project PRAT2015 "Control of dynamics with active constraints" and by Istituto Nazionale di Alta Matematica.

(2016) to relax the uniform convexity assumption, at the price of requiring a strong *outward pointing condition* on f in order to treat the discontinuity of second derivatives of the squared distance function at the boundary of $C(t)$. In Brokate et al. (2013), the authors adopt a suitable smoothing of the distance, which on one hand needs $C(t)$ constant and uniformly convex and $0 \in C$, while on the other avoids imposing further compatibility assumptions between f and C. In this paper we adapt to our situation the method developed in Brokate et al. (2013) and remove the assumption of strict convexity on C. The main technical part is Section 4.

2. PRELIMINARIES AND ASSUMPTIONS

Notation. We define the distance from a set $C \subset \mathbb{R}^n$ as $d(x) = \inf\{\|y-x\| : y \in C\}$ and *signed* distance from C as $d_S(x) = d(x)$ if $x \notin C$ and $d_S(x) = -\inf\{\|y-x\| : y \notin C\}$ if $c \in C$. The normal cone to a convex set C is defined as $N_C(x) = \emptyset$ if $x \notin C$ and $N_C(x) = \{v \in \mathbb{R}^n : \langle v, y - x\rangle \le 0 \ \forall y \in C\}$ if $x \in C$.

Assumptions on the set C. Let

$$C = \{x \in \mathbb{R}^n : g(x) \le 0\},$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is of class $\mathcal{C}^2$ with gradient $\nabla g \neq 0$ on the boundary ∂C of C, and with the Hessian matrix $\nabla^2 g(x)$ positive semidefinite for all $x \in \mathbb{R}^n$. Assume furthermore that $g(\cdot)$ is coercive, so that C is compact (and convex) and that $g(0) < 0$, so that $0 \in C$ and C has nonempty interior. Observe that under our assumption the signed distance $d_S(x)$ from C is of class $\mathcal{C}^2$ in a neighborhood of ∂C.

Assumptions on the dynamics and the cost. The control set $U \subset \mathbb{R}^n$ is compact and f is continuous and bounded, say by a constant β, and is of class $\mathcal{C}^1$ with respect to x, with $\|\nabla_x f(x,u)\| \le L$ for all x, u. The cost h is smooth.

Let now $\psi(x)$ be a $\mathcal{C}^2$ smoothing of d_S in the interior of C (which is < 0 in $\operatorname{int} C$ and is such that $\nabla\psi(x)$ is the unit external normal to C at x for every $x \in \partial C$). Set also

$$\Psi(x) = \frac{1}{3}\psi^3(x)\, 1_{(0,+\infty)}(\psi(x)).$$

Observe that $\Psi(\cdot)$ is of class $\mathcal{C}^2$ and convex in the whole of $\mathbb{R}^n$ and that both $\nabla\Psi(\cdot)$ and $\nabla^2\Psi(\cdot)$ vanish on C. Moreover one has, for each $x \in [R]^n$,

$$\nabla\Psi(x) = d^2(x)\nabla d(x), \quad (5)$$

$$\nabla^2\Psi(x) = 2d(x)\nabla d(x) \otimes \nabla d(x) + d^2(x)\nabla^2 d(x), \quad (6)$$

because in C, and in particular at the points where $\nabla d(x)$ does not exist (namely, in ∂C), both sides of the above expressions vanish, and outside C they coincide.

3. THE REGULARIZED PROBLEM

Consider the regularized dynamics

$$\dot{x}(t) = \frac{-1}{\varepsilon}\nabla\Psi(x(t)) + f(x(t), u(t)),\ x(0) = x_0, \quad (7)$$

where $\varepsilon > 0$ and $u(t) \in U$ for all t. For each given u, this Cauchy problem admits a unique solution x_ε for each $\varepsilon > 0$ on a maximal interval of existence. It is not difficult to prove that this interval is $[0,T]$ (see the proof of Proposition 1).

For every $\varepsilon > 0$ and every *global* minimizer x_*, u_* of (4) subject to (3) and (2), we consider the *approximate problem* $P_\varepsilon(u_*)$

$$\text{minimize } h(x(T)) + \frac{1}{2}\int_0^T \|u(t) - u_*(t)\|^2\, dt, \quad (8)$$

over controls u, where x is a solution of (7). By standard results, $P_\varepsilon(u_*)$ admits a global minimizer u_ε, with the corresponding solution x_ε. Necessary conditions of the original problem will be obtained by passing to the limit along conditions for $P_\varepsilon(u_*)$.

3.1 A priori estimates for the regularized problem

Proposition 1. Let $\varepsilon_n \to 0$ and let (u_n, x_n) be a solution of the problem P_{ε_n}. Then, up to a subsequence, u_n converges strongly in $L^2(0,T)$ to u_* and x_n converges weakly in $W^{1,2}(0,T)$ to x_*.

Proof. Since $0 \in C$ and so $\nabla\Psi(0) = 0$, by the convexity of Ψ we obtain that $\langle\nabla\Psi(x), x\rangle \ge 0$ for all $x \in \mathbb{R}^n$. Thus

$$\|x_n(t)\| - \|x_0\| = \\ = \int_0^t \Big\langle \frac{x_n(s)}{\|x_n(s)\|}, \frac{-1}{\varepsilon_n}\nabla\Psi(x_n(s)) + f(x_n(s), u_n(s))\Big\rangle\, ds \le \beta,$$

which, in particular, implies that x_n is defined in the whole of $[0,T]$. Moreover,

$$\begin{aligned}\|\dot{x}_n\|^2_{L^2} &= \int_0^T \Big\langle \dot{x}_n(t), \frac{-1}{\varepsilon_n}\nabla\Psi(x_n(t)) + f(x_n(t), u_n(t))\Big\rangle\, dt \\ &= \int_0^T \Big(\frac{-1}{\varepsilon_n}\frac{d}{dt}\Psi(x_n(t)) + \langle f(x_n(t), u_n(t)), \dot{x}_n(t)\rangle\Big)\, dt \\ &= \frac{-1}{\varepsilon_n}\Psi(x_n(T)) + \frac{1}{\varepsilon_n}\Psi(x_0) + \beta\int_0^T \|\dot{x}_n(t)\|\, dt \\ &\le \beta\sqrt{T}\|\dot{x}_n\|_{L^2},\end{aligned}$$

where we have used the fact that $x_0 \in C$ and that $\psi(x_n(T)) \ge 0$. The above estimate implies that the sequence $\dot{x}_n$ is uniformly bounded in $L^2(0,T)$. Thus, up to a subsequence, x_n converges weakly in $W^{1,2}(0,T)$ to $\bar{x}$. Observe now that the from the uniform boundedness of $\|\dot{x}_n\|_{L^2(0,T)}$ and of f, we can deduce from (7), thanks to (5), that

$$\|d(x_n(\cdot))^2\|_{L^2(0,T)} \le K\varepsilon_n \quad (9)$$

for a suitable constant K. Thus $x(t) \in C$ for all t. Again up to a subsequence, u_n converges weakly in $L^2(0,T)$ to some $\bar{u}$. By using the very same argument of Proposition 4.3 in Arroud et al. (2016), one can prove that $\bar{x}$ is the solution of (3), (2) corresponding to $\bar{u}$, that $\bar{u} = u_*$, and so that $\bar{x} = x_*$, and that the convergence is indeed strong.

Remark. Eq. (9) implies that, up to a subsequence,

$$\|d(x_n(\cdot))\|_{L^2(0,T)} \le \sqrt{TK}\sqrt{\varepsilon_n}. \tag{10}$$

From the uniform convergence of x_n (again up to a subsequence) we also get $\|d(x_n(\cdot))\|_{L^\infty} \to 0$ for $n \to \infty$. This is an important difference between this approach and the use of Moreau-Yosida approximation, which instead yields the stronger estimate $\|d(x_n(\cdot))\|_{L^\infty} \sim \varepsilon_n$ (see Sene et al. (2014) or (Arroud et al., 2016, Proposition 4.1)). Observe that the assumption that C is constant appears essential in order to obtain uniform *a priori* estimates for $\|\dot{x}_n\|_{L^2}$ within the present approach, which essentially uses the weaker penalization d^3 instead of the Moreau-Yosida one, namely d^2.

3.2 Necessary conditions for the regularized problem

The approximate problem $P_\varepsilon(u_*)$ satisfies the assumptions for necessary conditions of classical unconstrained optimal control problems. The same computations of Section 6 in Arroud et al. (2016) yield that for every ε and every minimizer $(u_\varepsilon, x_\varepsilon)$ there exists an absolutely continuous adjoint vector $p_n : [0,T] \to \mathbb{R}^n$ such that

$$-\dot{p}_\varepsilon(t) = \Big(\frac{-1}{\varepsilon}\nabla^2\Psi(x_\varepsilon(t)) + \nabla_x f(x_\varepsilon(t), u_\varepsilon(t))\Big)p_\varepsilon(t) \tag{11}$$

a.e. on $[0,T]$, together with the final condition

$$-p_\varepsilon(T) = \nabla h(x_\varepsilon(T)), \tag{12}$$

and the maximality condition

$$\langle p_\varepsilon(t), \nabla_u f(x_\varepsilon(t), u_\varepsilon(t))u_\varepsilon(t)\rangle - \langle u_\varepsilon(t) - u_*(t), u_\varepsilon(t)\rangle = \max_{u\in U}\{\langle p_\varepsilon(t), \nabla_u f(x_\varepsilon(t), u_\varepsilon(t))u\rangle - \langle u_\varepsilon(t) - u_*(t), u\rangle\}$$

for a.e. $t \in [0,T]$.

4. PASSING TO THE LIMIT

From now on we consider a sequence $\varepsilon_n \to 0$ such that the minimum (u_n, x_n) of the approximate problem converges as in the statement of Proposition 1. We set $p_n := p_{\varepsilon_n}$.

4.1 A priori estimates for the adjoint vectors of the approximate problem

We obtain from (11) and (6) that

$$\|p_n(t)\| - \|p_n(T)\| = \frac{-1}{\varepsilon_n}\int_t^T\Bigg(\frac{\big\langle\nabla^2\Psi(x_n(s))p_n(s), p_n(s)\big\rangle}{\|p_n(s)\|} + \Big\langle\nabla_x f(x_n(s), u_n(s))p_n(s), \frac{p_n(s)}{\|p_n(s)\|}\Big\rangle\Bigg)\,ds \le$$

(since $\nabla^2\Psi$ is positive semidefinite and $\nabla_x f$ is bounded)

$$\le L\int_t^T \|p_n(s)\|\,ds.$$

Recalling (12), we obtain from the above inequality and Gronwall's lemma that there exists a constant K_1 independent of n such that

$$\|p_n\|_\infty \le K_1. \tag{13}$$

Now we address ourselves to prove an *a priori* estimate on $\|\dot{p}_n\|_{L^1}$. To this aim, we define

$$\xi_n(t) = \langle p_n(t), \nabla d(x_n(t))\rangle,$$

(whenever it makes sense, i.e., if $x_n(t) \notin \partial C(t)$), so that,

$$\dot{\xi}_n(t) = \langle\dot{p}_n(t), \nabla d(x_n(t))\rangle + \langle p_n(t), \nabla^2 d(x_n(t))\dot{x}_n(t)\rangle.$$

Set now

$$\delta_n(t) = d(x_n(t)),\ \delta_n'(t) = \nabla d(x_n(t)),\ \delta_n''(t) = \nabla^2 d(x_n(t)).$$

With this notation, thanks to (5) and (6), eq. (11) can be rewritten as

$$-\dot{p}_n(t) = -\frac{\delta_n^2(t)}{\varepsilon_n}\delta_n''(t)p_n(t) - 2\frac{\delta_n(t)\xi_n(t)}{\varepsilon_n}\delta_n'(t) + +\nabla_x f(x_n(t), u_n(t))p_n(t). \tag{14}$$

Inserting $\dot{p}_n$ and $\dot{x}_n$ in the expression for $\dot{\xi}_n$ we obtain (omitting the t-dependence and using the fact that $\nabla^2 d(x)\,\nabla d(x) = 0$ for all x where it makes sense, and that $\delta_n\|\nabla\psi(x_n)\|^2 = \delta_n\|\nabla d(x_n)\| = \delta_n$)

$$-\dot{\xi}_n + 2\frac{\delta_n\xi_n}{\varepsilon_n} = -\frac{\delta_n^2}{\varepsilon_n}\langle\delta_n''p_n, \delta_n'\rangle + \langle\nabla_x f(x_n, u_n)p_n, \delta_n'\rangle - \langle p_n, \delta_n'' f(x_n, u_n)\rangle.$$

Observe now that the first summand in the right hand side of the above expression is bounded in $L^1(0,T)$ uniformly with respect to n, because, recalling (10), $\frac{\|\delta_n^2\|_{L^1}}{\varepsilon_n}$ is uniformly bounded. In turn, the second and the third summands are seen to be bounded in $L^\infty(0,T)$, uniformly with respect to n, by invoking (13). By multiplying both sides by $\mathrm{sign}(\xi_n)$ and integrating, we thus obtain the second estimate

$$\frac{1}{\varepsilon_n}\int_t^T \delta_n(s)|\xi_n(s)|\,ds \le K_2, \tag{15}$$

for a suitable constant K_2, independent of n. As a consequence, all three summands in the right hand side of the adjoint equation (14) are bounded in $L^1(0,T)$, uniformly with respect to n, and so we reach our final estimate

$$\|\dot{p}_n\|_{L^1(0,T)} \le K_3 \tag{16}$$

for a suitable constant K_3 independent of n.

4.2 Passing to the limit along the adjoint equation

By possibly extracting a further subsequence, we can assume that the sequence of measures $\dot{p}_n\,dt$ converges weakly$*$ in the sense of Radon measures to a signed vector measure μ, which is the distributional derivative of the BV function $p(t) =: \lim p_n(t)$, i.e., $\mu = dp$, where the limit of the p_n is pointwise in $[0,T]$. By arguing as in (Arroud et al., 2016, Proposition 7.3), we obtain first that the sequence of measures

$$\frac{\delta_n(t)\xi_n(t)}{\varepsilon_n}\delta_n'(t)\,dt$$

converges weakly* to a finite signed vector Radon measure, which can be written as

$$\xi(t)n_*(t)\,d\nu,$$

where $\xi \in L^1_\nu(0,T)$, $\xi \geq 0$ ν-a.e., $n_*(t)$ denotes the unit outward normal vector to C at $x_*(t)$ if $x_*(t) \in \partial C$ and 0 if $x_*(t) \in \operatorname{int} C$, and ν is a finite vector measure. Moreover

$$\frac{\delta_n^2(t)}{\varepsilon_n}\delta_n''(t)p_n(t) \rightharpoonup \eta(t)\nabla^2 d(x_*(t))p(t)$$

in $L^2(0,T)$, where $\eta \in L^\infty_\nu(0,T)$ and $\eta \geq 0$ a.e., with $\eta \equiv 0$ when x_* is in $\operatorname{int} C$.

4.3 Passing to the limit along the maximality condition

By taking into account Proposition 1, we obtain that the limit adjoint vector p is such that

$$\langle p(t), \nabla_u f(x_*(t), u_*(t))u_*(t)\rangle = = \max_{u\in U}\{\langle p(t), \nabla_u f(x_*(t), u_*(t))u\rangle\} \text{ for a.e. } t \in [0,T]. \quad (17)$$

5. THE MAIN RESULT

We deduce from Sections 3 and 4 the following necessary conditions:

Theorem 2. Under the assumptions stated in Section 2, let (x_*, u_*) be a global minimizer for (4) subject to (3) and to (2). Then there exist a BV adjoint vector $p : [0,T] \to \mathbb{R}^n$, together with a finite signed Radon measure ν on $[0,T]$, and measurable vectors $\xi, \eta : [0,T] \to \mathbb{R}$ (with $\xi \in L^1_\nu(0,T)$, $\xi(t) \geq 0$ for ν-a.e. t, $\eta \in L^\infty(0,T)$, and $\eta(t) \geq 0$ for a.e. t) such that $\xi(t) = \eta(t) = 0$ for all t with $x_*(t) \in \operatorname{int}$ C, satisfying the following properties:

- (adjoint equation)

for all continuous functions $\varphi : [0,T] \to \mathbb{R}^n$

$$-\int_{[0,T]} \langle \varphi(t), dp(t)\rangle = -\int_{[0,T]} \langle \varphi(t), n_*(t)\rangle\xi(t)\, d\nu(t) - \int_{[0,T]} \langle \varphi(t), \nabla_x^2 d(x_*(t))p(t)\rangle\eta(t)\, dt + \int_{[0,T]} \langle \varphi(t), \nabla_x f(x_*(t), u_*(t))p(t)\rangle\, dt,$$

- (transversality condition)

$$-p(T) = \nabla h(x_*(T)),$$

- (maximality condition)

$$\langle p(t), \nabla_u f(x_*(t), u_*(t))u_*(t)\rangle = = \max_{u\in U}\langle p(t), \nabla_u f(x_*(t), u_*(t))u\rangle \quad \text{for a.e. } t \in [0,T].$$

Some more precise statements on the measure ν require some assumptions on the reference trajectory x_*. In fact, consider the sets

$$E_0 := \{t \in [0,T] : x_*(t) \in \operatorname{int} C\}$$
$$E_\partial := \{t \in [0,T] : x_*(t) \in \partial C\}.$$

Of course, E_0 is open and E_∂ is closed, but one has to take into account the possibility that E_∂ be irregular (e.g., totally disconnected). Such phenomenon, in stratified state constrained control theory, is sometimes referred to as Zeno phenomenon, namely the switching from a stratum (the boundary of C, in this case) to other strata (the interior of C in this case) occurs at a complicated set (see, e.g., Barnard et al. (2013)).

The following is the second part of our necessary conditions. It is only a partial result, with respect to the rich set of conditions proved in Brokate et al. (2013).

Theorem 3. Under the assumptions stated in Section 2, let (x_*, u_*) be a global minimizer for (4) subject to (3) and to (2), and let p the adjoint vector given by Theorem 2. Define $p^N(t) = \langle p(t), n_*(t)\rangle$, $t \in [0,T]$. The following properties hold:

(1) $p^N(t) = 0$ for all $t \in E_0$, and p is absolutely continuous on E_0, where it satisfies the classical adjoint equation
$$-\dot p(t) = \nabla_x f(x_*(t), u_*(t))\, p(t). \quad (18)$$

(2) At every interior (or such that a left or right neighborhood is contained in E_∂) point t of E_∂, jumps of p may occur only in the normal direction $n_*(t)$, namely
$$p(t-) - p(t+) = \big(p^N(t-) - p^N(t+)\big)n_*(t).$$

(3) The adjoint vector p is absolutely continuous on every open interval contained in E_∂, and for a.e. t in such interval we have
$$-\dot p(t) = \langle \dot n_*(t), p(t)\rangle n_*(t) + \Gamma(t)p(t) - -\langle \Gamma(t)p(t), n_*(t)\rangle n_*(t), \quad (19)$$
where $\Gamma(t) = \nabla_x f(x_*(t), u_*(t)) - \eta(t)\nabla_x^2 d(x_*(t))$.

Proof. *(1).* The first assertion is obvious, since on E_0 we have $n_* \equiv 0$ and the absolute continuity together with (18) follow from the properties of the functions ξ and η proved in Theorem 2.

(2). Since $n_*(t)$ is continuous on E_∂, there exist $n-1$ continuous unit vectors $v_1(t),\ldots, v_{n-1}(t)$ such that $\mathbb{R}^n = \mathbb{R}n_*(t) \oplus \operatorname{span}\langle v_1(t), \ldots, v_{n-1}(t)\rangle$ for all $t \in E_\partial$. Let $t \in E_\partial$ and $\sigma > 0$ be such that $[t-\sigma, t+\sigma] \subset E_\partial$. Let $\varphi : [0,T] \to \mathbb{R}^n$ be continuous, with support contained in $[t-\sigma, t+\sigma]$. Set $\varphi^T(t) = \varphi(t) - \langle \varphi(t), n_*(t)\rangle n_*(t)$. By putting $\varphi^T(t)$ in place of φ in the adjoint equation we obtain

$$-\int_{t-\sigma}^{t+\sigma} \langle \varphi^T(s), dp(s)\rangle + \int_{t-\sigma}^{t+\sigma} \langle \varphi^T(s), n_*(s)\rangle\xi(s)\, d\nu = \int_{t-\sigma}^{t+\sigma} \langle \varphi^T(s), \nabla_x f(x_*(s), u_*(s))p(s)\rangle\, ds - \int_{t-\sigma}^{t+\sigma} \langle \varphi^T(s), \eta(s)\nabla^2 d(x_*(s))p(s)\rangle\, ds. \quad (20)$$

Observe now that $\langle \varphi^T, n_*(t)\rangle \equiv 0$, so that, by letting $\sigma \to 0$ in the above equation and using the continuity of φ^T, we obtain $\langle \varphi^T(t), p(t+) - p(t-)\rangle = 0$, namely $\langle p(t+)-p(t-), \varphi(t)\rangle = \big\langle p(t+)-p(t-), \langle \varphi(t), n_*(t)\rangle n_*(t)\big\rangle$. By taking subsequently φ such that $\varphi(t) = n_*(t)$, and $\varphi(t) = v_i(t)$, $i = 1, \ldots, n-1$, we obtain that $p(t-) -$

$p(t+) = \big(p^N(t-) - p^N(t+)\big)n_*(t)$, namely jumps of p may occur only in the direction $n_*(t)$, for all t in the interior of E_∂. If t is a left or a right endpoint of E_∂, one can extend n_* as a constant to the left or to the right of t and repeat the same argument.

(3). Fix now an interval $[s,t] \subseteq E_\partial$. The regularity condition on ∂C allows us to integrate by parts on (s,t), so that

$$\int_s^t \langle n_*(\tau), dp(\tau)\rangle + \int_s^t \langle \dot n_*(\tau), p(\tau)\rangle\, d\tau = \\ = \langle n_*(t+), p(t+)\rangle - \langle n_*(s-), p(s-)\rangle = 0. \quad (21)$$

Observe that both summands in the right hand side of (21) vanish, as a consequence of *(1)* and of the fact that p has bounded variation, since s and t belong to the interior of I_∂. In other words, the two measures $\langle n_*, dp\rangle$ and $\langle \dot n_*, p\rangle\, dt$ coincide in any open interval contained in E_∂. Therefore, for all continuous φ with support contained in (s,t), we obtain from (20) and (21) that

$$-\int_s^t \langle \varphi(\tau), dp(\tau)\rangle = \\ = -\int_s^t \langle \varphi(\tau), n_*(\tau)\rangle\, \langle n_*(\tau), dp(\tau)\rangle - \int_s^t \langle \varphi^T(\tau), dp(\tau)\rangle \\ = \int_s^t \langle \varphi(\tau), n_*(\tau)\rangle \langle p(\tau), \dot n_*(\tau)\rangle d\tau \\ + \int_s^t \Big\langle \varphi(\tau) - \langle \varphi(\tau), n_*(\tau)\rangle n_*(\tau), \nabla_x f(x_*(\tau), u_*(\tau)) p(\tau)\Big\rangle d\tau \\ - \int_s^t \Big\langle \varphi(\tau) - \langle \varphi(\tau), n_*(\tau)\rangle n_*(\tau), \eta(\tau)\nabla_x^2 d(x_*(\tau)) p(\tau)\Big\rangle d\tau$$

Since φ and the interval (s,t) are arbitrary, we obtain (19).

6. AN EXAMPLE

The state space is $\mathbb{R}^2 \ni (x,y)$, the constraint is $C := \{(x,y) : y \geq 0\}$, the upper half plane.

We wish to minimize $h(x(1), y(1)) = x(1) + y(1)$ subject to

$$\big(\dot x(t), \dot y(t)\big) \in -N_C\big(x(t), y(t)\big) + \big(u^x(t), u^y(t)\big) \\ \big(x(0), y(0)\big) = \big(0, y_0\big), \qquad y_0 \geq 0,$$

where the controls $\big(u^x(t), u^y(t)\big)$ belong to $[-1,1] \times [-1,1] =: U$.

This problem satisfies all our assumptions.

Observe first that if $y_0 \geq 1$, the constraint C does not play any role, and the optimality of the control $(-1,-1)$ is straightforward. Moreover, since the state constraint is not active, necessary optimality conditions are well known. If instead $0 \leq y_0 < 1$, then our analysis becomes relevant. An inspection to the level sets of the cost h shows that as long as the trajectory does not hit ∂C the only optimal control is $(-1,-1)$. Therefore, there exists at most one $\bar t$ such that the optimal solution hits ∂C and after $\bar t$ it remains on ∂C. On ∂C, namely for $x = 0$, we have $\nabla_x^2 d_C\big((0,y)\big) \equiv 0$. Thanks to Theorems 2 and 3 we obtain, for the optimal trajectory (x_*, y_*) corresponding to the optimal control (u_*^x, u_*^y) an adjoint vector (p^x, p^y), such that (p^x, p^y) is absolutely continuous on $(0,\bar t) \cup (\bar t, 1)$, and such that $\dot p^x = 0$, $\dot p^y = 0$ a.e. on $[0,T]$, $p^x(1) = p^y(1) = -1$, p^x is continuous at $t = 1$ and $p^y(1-) + 1 = 1$, namely $p^y(1-) = 0$. Thus the adjoint vector (p^x, p^y) satisfies:

$$p^x(t) = -1 \quad \text{for all } t \in [0,1] \\ p^y(t) \text{ is constant on } [0,\bar t) \cup (\bar t, 1) \\ p^y(1) = -1.$$

The maximum condition reads as

$$\langle(-1,-1), (u_*^x, u_*^y)\rangle = \max_{|u_1|\leq 1, |u_2|\leq 1} \langle(-1,-1), (u_1, u_2)\rangle \\ \text{for } t = 1 \\ \langle(-1, p^y(t)), (u_*^x, u_*^y)\rangle = \max_{|u_1|\leq 1, |u_2|\leq 1} \langle(-1, p^y(t)), (u_1, u_2)\rangle \\ \text{for } 0 \leq t < 1, t \neq \bar t,$$

which gives $u_*^x = -1$, while no information is available for $u_*^y(t)$. If we assume that u_*^y is constant, then an expected optimal control $u_*^y = -1$ is found. Of course all other optimal controls u_*^y, namely $u_*^y(t) = -1$ for $0 \leq t < \bar t$ and $u_*^y(t) \leq 0$ for $\bar t < t < 1$ satisfy our necessary conditions, and in this case $\bar t = y_0$.

REFERENCES

Ch. Arroud and G. Colombo. A Maximum Principle for the controlled Sweeping Process. *SVVA*, in print. DOI: 10.1007/s11228-017-0400-4.

F. Bagagiolo. Dynamic programming for some optimal control problems with hysteresis. *NoDEA*, 9:149–174, 2002.

R.C. Barnard and P.R. Wolenski. Flow invariance on stratified domains. *Set-Valued Var. Anal.*, 21:377–403, 2013.

M. Brokate and P. Krejčí. Optimal control of ODE systems involving a rate independent variational inequality. *Discrete and continous dynamical systems series B*, 18: 331–348, 2013.

Tan H. Cao and B.Sh. Mordukhovich. Optimality conditions for a controlled sweeping process with applications to the crowd motion model. *Discrete and continous dynamical systems series B*, 22:267–306, 2017.

G. Colombo, R. Henrion, N.D. Hoang, and B.Sh. Mordukhovich. Optimal control of the sweeping process over polyhedral controlled sets. *J. Differential Equations*, 260:3397–3447, 2016.

G. Colombo and M. Palladino. The minimum time function for the controlled Moreau's sweeping process. *SIAM J. Control*, 54:2036–2062, 2016.

A. Gudovich and M. Quincampoix. Optimal control with hysteresis nonlinearity and multidimensional play operator. *SIAM J. Control Optim.* 49:788–807, 2011.

J.-J. Moreau. On unilateral constraints, friction and plasticity. In *New variational techniques in mathematical physics* (Centro Internaz. Mat. Estivo (C.I.M.E.), II

Ciclo, Bressanone, 1973):171–322. Edizioni Cremonese, Rome, 1974.

M. Sene and L. Thibault. Regularization of dynamical systems associated with prox-regular moving sets. *J. Nonlin. and Convex Anal.*, 15:647–663, 2014.

L. Thibault. Sweeping process with regular and nonregular sets. *J. Differential Equations*, 193:1–26, 2003.

R.B. Vinter. *Optimal control*. Birkhäuser, Boston, 2000.