Viscous Approximations of Non-Convex Sweeping Processes in the Space of Regulated Functions

Original
Viscous Approximations of Non-Convex Sweeping Processes in the Space of Regulated Functions / Krejci, P.; Monteiro, G. A.; Recupero, V.. - In: SET-VALUED AND VARIATIONAL ANALYSIS. - ISSN 1877-0541. - ELETTRONICO. - 31:4(2023), pp. 1-38. [10.1007/s11228-023-00695-y]

Availability:
This version is available at: 11583/2982545 since: 2023-09-28T10:38:17Z

Publisher:
Springer

Published
DOI:10.1007/s11228-023-00695-y

Viscous approximations of non-convex sweeping processes in the space of regulated functions

Pavel Krejčí[*], Giselle Antunes Monteiro[†], Vincenzo Recupero[‡§]

Abstract

Vanishing viscosity approximations are considered here for discontinuous sweeping processes with non-convex constraints. It is shown that they are well-posed for sufficiently small viscosity parameters, and that their solutions converge pointwise, as the viscosity parameter tends to zero, to the left-continuous solution to the sweeping process in the Kurzweil integral setting. The convergence is uniform if the input is continuous.

Keywords: Evolution variational inequalities, Sweeping processes, Vanishing viscosity, Prox-regular sets, Regulated functions
2010 AMS Subject Classification: 49J40, 47J20, 34A60, 34G25

Introduction

Sweeping processes have been introduced in [27] as an abstract setting of problems arising for example in elastoplasticity modeling, where the constitutive relation can be formulated as a constrained evolution system. Typically, the functional framework consists in assuming that

$$X \quad \text{is a real Hilbert space} \tag{0.1}$$

endowed with scalar product $\langle \cdot, \cdot \rangle$ and norm $|x| = \sqrt{\langle x, x \rangle}$ for $x \in X$, and one considers a family $C(t) \subset X$ of nonempty closed subsets of X parameterized by the time variable $t \in [0, T]$, where $T > 0$ is some given final time. The problem is to find a function $\xi : [0, T] \to X$ with a prescribed initial condition $\xi(0) = \xi_0 \in C(0)$, such that $\xi(t) \in C(t)$

[*]Faculty of Civil Engineering, Czech Technical University, Thákurova 7, 16629 Praha 6, Czech Republic (pavel.krejci@cvut.cz).
[†]Institute of Mathematics, Czech Academy of Sciences, Žitná 25, 11567 Praha 1, Czech Republic, (gam@math.cas.cz).
[‡]Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy (vincenzo.recupero@polito.it).
[§]Vincenzo Recupero is a member of GNAMPA-INdAM.

for all $t \in [0,T]$ and its derivative at time t points in the inward normal direction to $C(t)$ at the point $\xi(t)$. Formally, this can be stated as

$$-\dot{\xi}(t) \in N_{C(t)}(\xi(t)) \quad \text{for} \quad t \in (0,T), \quad \xi(0) = \xi_0, \tag{0.2}$$

where both the "time derivative" $\dot{\xi}(t)$ and the outward normal cone $N_{C(t)}(\xi(t))$ to $C(t)$ at the point $\xi(t)$ have to be given an appropriate meaning.

In the paper [27], this problem is uniquely solved provided that $C(t)$ is convex for every time t and that the mapping $t \mapsto C(t)$ is absolutely continuous in terms of the Hausdorff distance. In this case the solution ξ turns out to be absolutely continuous and (0.2) is satisfied almost everywhere. In [28] the analysis of sweeping processes was then extended to the case when the convex moving set $C(t)$ has bounded variation with respect to the Hausdorff metric. Under this weaker assumption, inclusion (0.2) has to be properly interpreted in the sense of the differential measures and it is shown to admit a unique solution of bounded variation.

The study of non-convex sweeping processes started with M. Valadier [33] and, since then, has called the attention of many other authors, e. g., [3, 10, 31]. An important concept which allows to get around the convexity of sets is the notion of uniform prox-regularity. These are closed sets having a neighborhood where the projection exists uniquely and in a continuous way. Sets with such a property appear in the literature under different terminologies; being introduced under the name of 'positively reached sets' by H. Federer [15] in finite dimensional setting. A series of properties as well as the connection between sets and functions was deeply investigated in [34] (therein called 'weak convex sets/functions'). The notion of prox-regularity was later extended to infinite dimensional spaces ([9, 30]), and appears to lead to an appropriate class of non-convex sets for which one can prove existence and uniqueness results for sweeping processes, see for instance [1, 4, 5, 8, 14]. Especially, for a given parameter $r > 0$, we will consider the notion of r-prox-regular set Z, which in particular enjoys of the property that a projection is uniquely defined in the r-neighborhood of Z. Here, we show in Lemma 1.3 that the parameter r can also be interpreted as the maximal radius of exterior balls tangent to Z.

A recent paper [29] presents a fairly general result for BV sweeping processes with prox-regular constraints. The case when the moving uniform prox-regular constraint has unbounded variation was instead dealt with in [12] where it is assumed that $C(t)$ is continuous in time: In this paper a so-called interior cone condition is also required, which essentially means that cusps are not admitted on the boundary.

In [24], we have separated the effects of translation and shape change of the set $C(t)$ by representing it in the form $C(t) = u(t) - Z(w(t))$ for given $u : [0,T] \to X$ and $w(t) : [0,T] \to A$, where A is a closed set of parameters in a Banach space W. The motivation comes from application in plasticity theory, where u and w have different meanings. Typically, u can be interpreted as stress tensor, and w can be, for example, a fatigue parameter, see [18, 32]. In this case, $Z(w)$ is the admissible stress domain and its boundary $\partial Z(w)$ is the yield surface.

If we allow both u and w to be regulated functions, that is, both right and left limits exist at each point of the domain of definition, then the Kurzweil theory of integration offers an efficient formalism for interpreting sweeping processes as Kurzweil integral variational inequalities. A detailed exposition of the method in the context of convex constraints can be found, e. g., in [22].

We denote by $G(0,T;X)$ and $G(0,T;W)$ the space of regulated functions with values in X (in W, respectively) endowed with the sup-norm, and by $BV(0,T;X)$ the dense subset of $G(0,T;X)$ of functions of bounded variation. Under the assumption that $u \in G(0,T;X)$ and $w \in G(0,T;W)$ are right-continuous, and $\{Z(w) : w \in W\}$ is a family of r-prox-regular sets depending continuously on w in terms of the Hausdorff distance, we have proved that $u(t)$ can be split into a sum $u(t) = x(t) + \xi(t)$ of right-continuous regulated functions in such a way that $x(t) \in Z(w(t))$ for all $t \in [0,T]$, $\xi \in BV(0,T;X)$, and the sweeping process (0.2) reformulated in terms of the Kurzweil integral with $V(\xi)(t) := \mathrm{Var}_{[0,t]}(\xi)$

$$\int_0^T \langle x(t) - z(t), \mathrm{d}\xi(t)\rangle + \frac{1}{2r}\int_0^T |x(t) - z(t)|^2 \,\mathrm{d}V(\xi)(t) \ge 0$$
$$\forall z \in G(0,T;X),\ z(t) \in Z(w(t)) \text{ for all } t \in [0,T], \tag{0.3}$$

admits a unique solution (ξ, x) for each given initial condition $x(0) = x_0 \in Z(w(0))$.

The aim of this paper is to study viscous approximations of non-convex Kurzweil sweeping processes of the form (0.3) under an additional hypothesis that the set Z does not depend on w, that is, the motion of $C(t)$ is driven only by translation. More specifically, we assume that $Z \subset X$ is an r-prox-regular set (see Definition 1.1 below), and consider the mapping which with a given $u \in G(0,T;X)$ and with a given initial condition $x_0 \in Z$ associates the solution $\xi \in G(0,T;X)$ of the Kurzweil integral variational inequality

$$\int_0^\tau \langle u(t+) - \xi(t+) - z(t), \mathrm{d}\xi(t)\rangle + \frac{1}{2r}\int_0^\tau |u(t+) - \xi(t+) - z(t)|^2 \,\mathrm{d}V(\xi(t)) \ge 0 \tag{0.4}$$

for all $z \in G(0,T;Z)$ and all $\tau \in [0,T]$ under the constraint $u(t) - \xi(t) \in Z$ for all $t \in [0,T]$, and with initial condition $\xi(0) = u(0) - x_0$.

By a viscous approximation we mean the solution ξ^ε of the differential equation

$$\varepsilon\dot{\xi}^\varepsilon(t) = \frac{f(|D(u(t) - \xi^\varepsilon(t))|)}{|D(u(t) - \xi^\varepsilon(t))|} D(u(t) - \xi^\varepsilon(t)), \quad \xi^\varepsilon(0) = \xi(0), \tag{0.5}$$

where f is a convex function such that $f(0) = 0$, and D is the distance mapping defined below in Section 1. We call (0.5) "viscous approximation" of (0.4) by analogy to applications in viscoelastoplasticity, where $f(p) = p$ for $p \ge 0$ and ε is the viscosity coefficient. A power law $f(p) = p^{m-1}$ for $m \geq 2$ is used as penalty function in [11] in the study of well-posedness of a sweeping process under integral forcing. More general classes of functions f turn out to be useful in applications to optimal control problems, see [7, 17, 21]. The reason is that in optimal control, the limit as $\varepsilon \to 0$ has to be taken simultaneously in the constitutive relation and in the associated Euler-Lagrange condition, which behave differently near the critical boundary. As the main result here, we prove that the solution ξ^ε of (0.5) converge pointwise as $\varepsilon \searrow 0$ to the left-continuous representative of the solution ξ of (1.1).

The fact that the viscous limit gives the left continuous representative of ξ independently of whether or not the input u is left continuous has already been observed in [22, 23] for the case of convex constraints and regulated inputs. Note that the case $f(p) = p$ in (0.5) corresponds also to the Moreau-Yosida approximation of the sweeping process. The first result on the topic goes back to Moreau's original work [27] on convex sweeping processes with absolutely continuous inputs. The convergence of Moreau-Yosida approximations was then further explored by other authors, e. g., [25]. Recently, a similar result in the non-convex case was proved in [20] under an additional compactness assumption. Let us also mention another approach to regularization which has been exploited in [13], where the approximation does not rely on the distance function, and the approximated solution lies always in the interior of the constraint.

The structure of the paper is as follows. In Section 1, we recall the properties of r-prox-regular sets and prove in Lemma 1.3 an alternative equivalent definition based on the geometric idea of tangent balls of radius r. The main results are stated in Section 2. The most substantial step is made in Section 3, where we prove that for sufficiently small ε, Problem (0.5) admits a global solution for every left-continuous regulated input with sufficiently small jumps, and that the total variations of the solutions are bounded independently of ε. This is indeed nontrivial, as the distance function is defined only in the r-neighborhood of the set Z, and one has to make sure that the solution stays globally within this range. In Section 4, we prove that the mapping which with u associates the solution ξ^ε to (0.5) is locally 1/2-Hölder continuous with constants independent of ε. Section 5 is devoted to an explicit formula for the solution of (0.5) in the case of piecewise constant input u. These results are then used in Section 6 for the proof of the pointwise convergence of ξ^ε toward the solution of the Kurzweil integral variational inequality (0.4). The convergence is uniform if u is continuous, and this fact is proved in Section 7.

1 Prox-regular sets in a Hilbert space

There are several equivalent approaches to prox-regularity – see, e. g., [30]. In particular, the definition below corresponds to items (a) and (g) of Theorem 4.1 in [30].

Definition 1.1. Let $Z \subset X$ be a closed connected set and let $d(y) := \operatorname{dist}(y, Z) = \inf\{|x - z| : z \in Z\}$ denote the distance of a point $y \in X$ from the set Z. Let $r > 0$ be given. We say that Z is r*-prox-regular* if the following condition hold.

$$\forall y \in X: \ d(y) \in (0,r) \ \ \exists x \in Z : \mathrm{dist}\left(x + \frac{r}{d(y)}(y-x), Z\right) = \frac{r}{d(y)}|y-x| = r. \tag{1.1}$$

The following variational characterization of r-prox-regularity is proved for example in [24, Lemma 1.3].

Lemma 1.2. *A set $Z \subset X$ is r-prox-regular if and only if for every $y \in X$ such that $d(y) = \mathrm{dist}(y, Z) < r$ there exists a unique $x \in Z$ such that $|y-x| = d(y)$ and*

$$\langle y-x, x-z\rangle + \frac{|y-x|}{2r}|x-z|^2 \geq 0 \quad \forall z \in Z. \tag{1.2}$$

Formula (1.2) can be used for introducing the concept of projection $Q: Z + B_r(0) \to Z$, where $B_r(0)$ denotes the open ball centered at 0 with radius r, which with a given $y \in Z + B_r(0)$ associates $x \in Z$ satisfying (1.2).

We further define the *distance mapping* $D: Z + B_r(0) \to B_r(0)$ by the formula

$$D(y) = y - Q(y) \tag{1.3}$$

for $y \in Z + B_r(0)$. For $y \in Z$ we have indeed $D(y) = 0$, and for $y \in Z + B_r(0) \setminus Z$ we have $|D(y)| = d(y) > 0$. We prove below in Lemma 1.4 that both Q and D are locally Lipschitz continuous in $Z + B_r(0)$.

There exists a simple geometric characterization of r-prox-regularity which implies in particular that a nonempty closed set $Z \subset X$ is r-prox-regular for every $r > 0$ if and only if it is convex. The exact statement reads as follows.

Lemma 1.3. *Let $Z \subset X$ be a nonempty closed set. Then the following two conditions are equivalent.*

(i) *Z is r-prox-regular;*

(ii) $x, y \in Z$, $|x-y| < 2r \implies \mathrm{dist}(\frac{1}{2}(x+y), Z) \leq r - \sqrt{r^2 - \frac{1}{4}|x-y|^2}$.

Proof. Assume first that Z is r-prox-regular, let $x, y \in Z$ be chosen such that $d := \frac{1}{2}|x-y| < r$, put $\bar{x} = \frac{1}{2}(x+y)$, $\rho = \mathrm{dist}(\bar{x}, Z)$, and assume that $\rho > 0$. We have indeed $\rho \leq |\bar{x} - x| = \frac{1}{2}|x-y| < r$. Let $z := Q(\bar{x})$. By definition of prox-regularity the point $\bar{y} := z + \frac{r}{\rho}(\bar{x} - z)$ has the property that $\mathrm{dist}(\bar{y}, Z) = r$. Using the identity $|u+v|^2 + |u-v|^2 = 2(|u|^2 + |v|^2)$ for $u = \bar{x} - \bar{y}$, $v = \frac{1}{2}(x-y)$ we have

$$4r^2 \leq 2\left(|\bar{y} - x|^2 + |\bar{y} - y|^2\right) = |x-y|^2 + 4|\bar{y} - \bar{x}|^2 = |x-y|^2 + 4(r-\rho)^2,$$

which yields $(r-\rho)^2 \geq r^2 - \frac{1}{4}|x-y|^2$, and (ii) follows.

Conversely, assume that (ii) holds, and let $y \in X$ be chosen such that $d = \operatorname{dist}(y, Z) \in (0, r)$. We find a sequence $\{x_n : n \in \mathbb{N}\} \subset Z$ such that $\varepsilon_n := |y - x_n| - d \searrow 0$. We define $\varepsilon^* > 0$ as the positive solution of the equation $r - \sqrt{r^2 - (d + \varepsilon^*)^2} = d - \varepsilon^*$, that is,

$$\varepsilon^* = \sqrt{\frac{r^2}{4} + rd - d^2} - \frac{r}{2},$$

and we may assume that

$$0 < \varepsilon_n < \varepsilon^* \quad \forall n \in \mathbb{N}. \tag{1.4}$$

Observe that $\varepsilon^* < r - d$, thus we have $|x_n - x_m| \le 2d + \varepsilon_n + \varepsilon_m < 2d + 2\varepsilon^* < 2r$, and from Hypothesis (ii) and inequality (1.4) it follows for all $m, n \in \mathbb{N}$ that

$$\rho_{nm} := \operatorname{dist}\left(\frac{1}{2}(x_n + x_m), Z\right) \le r - \sqrt{r^2 - \frac{1}{4}|x_n - x_m|^2} < r - \sqrt{r^2 - (d + \varepsilon^*)^2} = d - \varepsilon^*. \tag{1.5}$$

Let us recall the identity

$$2\left(|y - x_n|^2 + |y - x_m|^2\right) = |x_n - x_m|^2 + 4\left|y - \frac{1}{2}(x_n + x_m)\right|^2. \tag{1.6}$$

We find $z_{nm} \in Z$ such that $|\frac{1}{2}(x_n + x_m) - z_{nm}| < \rho_{nm} + \varepsilon_n + \varepsilon_m < d$ by virtue of (1.5). The triangle inequality yields

$$\left|y - \frac{1}{2}(x_n + x_m)\right| \ge |y - z_{nm}| - (\rho_{nm} + \varepsilon_n + \varepsilon_m),$$

and from (1.6) we obtain

$$\frac{1}{2}\left((d + \varepsilon_n)^2 + (d + \varepsilon_m)^2\right) \ge \frac{1}{4}|x_n - x_m|^2 + (d - (\rho_{nm} + \varepsilon_n + \varepsilon_m))^2, . \tag{1.7}$$

which implies in particular that

$$\frac{1}{4}|x_n - x_m|^2 \le 3d(\varepsilon_n + \varepsilon_m) + 2d\rho_{nm}. \tag{1.8}$$

On the other hand, noting that by (1.5) we have

$$(r - \rho_{nm})^2 \ge r^2 - \frac{1}{4}|x_n - x_m|^2, \tag{1.9}$$

by adding (1.9) to (1.7) we obtain

$$r^2 - 2r\rho_{nm} + \rho_{nm}^2 + d^2 + d(\varepsilon_n + \varepsilon_m) + \frac{1}{2}(\varepsilon_n^2 + \varepsilon_m^2) \ge r^2 + d^2 - 2d(\rho_{nm} + \varepsilon_n + \varepsilon_m) + (\rho_{nm} + \varepsilon_n + \varepsilon_m)^2,$$

which yields

$$2(r-d)\rho_{nm} \leq 3d(\varepsilon_n + \varepsilon_m). \tag{1.10}$$

Therefore, from (1.8) we get

$$\frac{1}{4}|x_n - x_m|^2 \leq \frac{3rd}{r-d}(\varepsilon_n + \varepsilon_m). \tag{1.11}$$

We conclude that $\{x_n\}$ is a Cauchy sequence in X, and its limit $\bar{x} = \lim_{n\to\infty} x_n$ is the projection of y onto Z.

It remains to prove that if $\operatorname{dist}(y, Z) = d \in (0, r)$ and $|x-y| = d$, then $\operatorname{dist}(x + (s/d)(y-x), Z) = s$ for every $s \in (d, r]$. Let $d \in (0, r)$ be given and let

$$\varepsilon_0 := \frac{d(r-d)}{r+d} < \varepsilon^* \tag{1.12}$$

with ε^* from (1.4). We first prove the following implication:

$$\operatorname{dist}(y, Z) = d,\ x = Qy \ \Rightarrow\ \operatorname{dist}\left(x + \frac{d+\varepsilon}{d}(y-x), Z\right) = d + \varepsilon \quad \forall \varepsilon \in (0, \varepsilon_0]. \tag{1.13}$$

Indeed, this implies that $\sup\{s \in [d, r] : \operatorname{dist}(x + (s/d)(y-x), Z) = s\} = r$ which is the desired statement. It suffices to define the sequences $d_0 = d$, $\varepsilon_j = d_j(r-d_j)/(r+d_j)$, $d_{j+1} = d_j + \varepsilon_j$ for $j \in \mathbb{N}$, apply repeatedly (1.13) replacing d with d_j and ε_0 with ε_j, and check that $d_j \to r$ as $j \to \infty$.

To prove (1.13), we proceed by contradiction. We define the unit vector $n = (y-x)/d$. Assume that there exists $\varepsilon < \varepsilon_0$ and $x_\varepsilon \in Z$ such that

$$|x + (d+\varepsilon)n - x_\varepsilon| < d + \varepsilon,$$

that is,

$$|x - x_\varepsilon|^2 + 2(d+\varepsilon)\langle n, x - x_\varepsilon\rangle < 0. \tag{1.14}$$

Let $z_\varepsilon = Q(\frac{1}{2}(x + x_\varepsilon))$. We have $|x + dn - z_\varepsilon| > d$ and

$$\left|\frac{x + x_\varepsilon}{2} - z_\varepsilon\right| \leq r - \sqrt{r^2 - \frac{1}{4}|x - x_\varepsilon|^2}. \tag{1.15}$$

Hence, by the triangle inequality

$$d < \left|\frac{x + x_\varepsilon}{2} - z_\varepsilon\right| + \left|\frac{x - x_\varepsilon}{2} + dn\right|. \tag{1.16}$$

We use (1.14) to estimate

$$\left|\frac{x - x_\varepsilon}{2} + dn\right|^2 = d^2 + \frac{1}{4}|x - x_\varepsilon|^2 + d\langle n, x - x_\varepsilon\rangle < d^2 - \frac{\lambda_\varepsilon}{4}|x - x_\varepsilon|^2 = d^2 - \lambda_\varepsilon \Delta_\varepsilon^2,$$

where we set $\lambda_\varepsilon = \frac{d-\varepsilon}{d+\varepsilon} > 0$ and $\Delta_\varepsilon = \frac{1}{2}|x - x_\varepsilon|$. From (1.15)–(1.16) we thus get

$$d < r - \sqrt{r^2 - \Delta_\varepsilon^2} + \sqrt{d^2 - \lambda_\varepsilon \Delta_\varepsilon^2},$$

or, equivalently,

$$\sqrt{r^2 - \Delta_\varepsilon^2} - \sqrt{d^2 - \lambda_\varepsilon \Delta_\varepsilon^2} < r - d. \tag{1.17}$$

We have by (1.14) that $\Delta_\varepsilon < d + \varepsilon$ and $\lambda_\varepsilon \Delta_\varepsilon^2 < d^2 - \varepsilon^2$. The fact that the left-hand side of (1.17) is positive follows from the sequence of inequalities

$$r^2 - d^2 - (1 - \lambda_\varepsilon)\Delta_\varepsilon^2 > r^2 - d^2 - \frac{1 - \lambda_\varepsilon}{\lambda_\varepsilon}(d^2 - \varepsilon^2) > r^2 - d^2 - 4\varepsilon d > 0$$

as a consequence of (1.12) and of the fact that $\varepsilon \le \varepsilon_0$. We square both sides of (1.17) and get

$$rd - \frac{1 + \lambda_\varepsilon}{2}\Delta_\varepsilon^2 < \sqrt{r^2 - \Delta_\varepsilon^2}\sqrt{d^2 - \lambda_\varepsilon \Delta_\varepsilon^2}. \tag{1.18}$$

The left-hand side of (1.18) can be estimated from below by $d(r - d - \varepsilon) > 0$. The square of (1.18) yields that

$$\lambda_\varepsilon r^2 + d^2 - (1 + \lambda_\varepsilon) rd < \left(\lambda_\varepsilon - \left(\frac{1 + \lambda_\varepsilon}{2}\right)^2\right)\Delta_\varepsilon^2 = -\frac{\varepsilon^2}{(d + \varepsilon)^2}\Delta_\varepsilon^2,$$

that is,

$$d(r - d) < \varepsilon(r + d) - \frac{\varepsilon^2}{(d + \varepsilon)(r - d)}\Delta_\varepsilon^2,$$

which contradicts the choice of ε. This completes the proof of Lemma 1.3. ■

The distance mapping D is locally Lipschitz continuous in the following sense.

Lemma 1.4. *Let $y_1, y_2 \in X$ be such that*

$$d(y_i) = |D(y_i)| \le \frac{r}{(1 + \kappa)^2}$$

for some $\kappa > 0$ and $i = 1, 2$. Then

$$|D(y_1) - D(y_2)| \le \left(1 + \frac{\sqrt{3}}{\kappa}\right)|y_1 - y_2|.$$

Proof. For $i = 1, 2$ put $\xi_i = D(y_i)$. We have by (1.2) that

$$\langle \xi_1, y_1 - \xi_1 - y_2 + \xi_2 \rangle + \frac{|\xi_1|}{2r}|y_1 - \xi_1 - y_2 + \xi_2|^2 \ge 0,$$
$$\langle \xi_2, y_2 - \xi_2 - y_1 + \xi_1 \rangle + \frac{|\xi_2|}{2r}|y_1 - \xi_1 - y_2 + \xi_2|^2 \ge 0.$$

Summing up the above inequalities we get

$$|\xi_1 - \xi_2|^2 \le \langle \xi_1 - \xi_2, y_1 - y_2\rangle + \frac{1}{(1+\kappa)^2}\big(|y_1 - y_2| + |\xi_1 - \xi_2|\big)^2. \tag{1.19}$$

The Young inequality yields

$$\frac{2}{(1+\kappa)^2}|y_1 - y_2|\,|\xi_1 - \xi_2| \le \frac{\kappa}{(1+\kappa)^2}|\xi_1 - \xi_2|^2 + \frac{1}{\kappa(1+\kappa)^2}|y_1 - y_2|^2,$$
$$\langle \xi_1 - \xi_2, y_1 - y_2\rangle \le \frac{\kappa}{2(1+\kappa)}|\xi_1 - \xi_2|^2 + \frac{1+\kappa}{2\kappa}|y_1 - y_2|^2.$$

From (1.19) we obtain

$$|\xi_1 - \xi_2|^2 \le \frac{(1+\kappa)^2 + 2}{\kappa^2}|y_1 - y_2|^2,$$

and the assertion follows from the inequality $(1+\kappa)^2 + 2 \le (\kappa + \sqrt{3})^2$. ■

Let us mention a superposition formula for the distance mapping which will be used later on in Section 5.

Lemma 1.5. *Let Z be r-prox-regular and let D be the mapping defined by* (1.3). *Then for every $y \in Z + B_r(0)$ and every $\delta \in (-1, d^*(y))$ we have*

$$D(y + \delta D(y)) = (1+\delta)D(y), \tag{1.20}$$

where

$$d^*(y) = \frac{r}{d(y)} - 1 \quad \text{for} \quad d(y) > 0, \qquad d^*(y) = \infty \quad \text{for} \quad d(y) = 0.$$

Proof. It suffices to prove that

$$Q(Q(y) + (1+\delta)D(y)) = Q(y) \tag{1.21}$$

for every $y \in Z + B_r(0)$. Indeed, if (1.21) holds, then

$$\begin{aligned} D(y + \delta D(y)) &= D(Q(y) + (1+\delta)D(y)) = Q(y) + (1+\delta)D(y) - Q(Q(y) + (1+\delta)D(y)) \\ &= (1+\delta)D(y), \end{aligned}$$

and (1.20) follows. To prove (1.21), we denote $v = Q(y)$, $w = Q(Q(y) + (1+\delta)D(y))$, and use (1.2) to obtain

$$\langle y - v, v - w\rangle + \frac{|y - v|}{2r}|v - w|^2 \ge 0,$$
$$\langle v + (1+\delta)(y - v) - w, w - v\rangle + \frac{|v + (1+\delta)(y-v) - w|}{2r}|v - w|^2 \ge 0.$$

We add to the second inequality above the $(1+\delta)$-multiple of the first inequality and obtain

$$|v-w|^2 \le \frac{(1+\delta)|D(y)| + |v+(1+\delta)D(y)-w|}{2r}|v-w|^2.$$

We have by hypothesis $(1+\delta)|D(y)| < r$, $|v+(1+\delta)D(y)-w| < r$, hence $v = w$, which we wanted to prove. ■

The following result recalls the known relation between the distance function d and the distance mapping D, see e.g. [30]. For the reader's convenience, the result is stated here in a way which suits our discussion and we give an elementary proof.

Lemma 1.6. *For $y \in Z + B_r(0)$ put $\psi(y) = \frac{1}{2}d^2(y)$. Then the directional derivative*

$$\psi'(y;v) := \lim_{\delta\to 0}\frac{1}{\delta}(\psi(y+\delta v) - \psi(y))$$

exists for every $y \in Z + B_r(0)$ and $v \in X$, and we have $\psi'(y;v) = \langle D(y), v\rangle$, that is,

$$D(y) = \nabla\psi(y) \quad \forall y \in Z + B_r(0). \tag{1.22}$$

Proof. For all $\delta \in (-\delta_0, \delta_0)$, where $\delta_0 > 0$ is such that $\delta_0|v| < r - d(y)$, we have

$$\begin{aligned}\frac{1}{\delta}(\psi(y+\delta v) - \psi(y)) &= \frac{1}{2\delta}\langle D(y+\delta v) - D(y), D(y+\delta v) + D(y)\rangle \\ &= \frac{1}{2\delta}|D(y+\delta v) - D(y)|^2 + \frac{1}{\delta}\langle D(y+\delta v) - D(y), D(y)\rangle. \end{aligned} \tag{1.23}$$

We have $|D(y+\delta v) - D(y)| \le C_0\delta|v|$ with a constant $C_0 > 0$ independent of δ, hence the first term on the right hand side of (1.23) tends to 0 as $\delta \to 0$. To handle the second term, we use the fact that by definition of D, we have

$$\langle D(y), y - D(y) - z\rangle + \frac{|D(y)|}{2r}|y - D(y) - z|^2 \ge 0 \quad \forall z \in Z. \tag{1.24}$$

We may choose in particular $z = Q(y+\delta v) = (y+\delta v) - D(y+\delta v)$ and obtain

$$\langle D(y), D(y+\delta v) - D(y)\rangle \ge \delta\langle D(y), v\rangle - C_1\delta^2|v|^2 \tag{1.25}$$

where $C_1 > 0$ is a constant independent of δ. We similarly have

$$\langle D(y+\delta v), y+\delta v - D(y+\delta v) - z\rangle + \frac{|D(y+\delta v)|}{2r}|y+\delta v - D(y+\delta v) - z|^2 \ge 0 \quad \forall z \in Z,$$

which for $z = y - D(y)$ yields

$$\langle D(y+\delta v), D(y) - D(y+\delta v)\rangle \ge -\delta\langle D(y+\delta v), v\rangle - C_2\delta^2|v|^2$$

with a constant $C_2 > 0$ independent of δ, and consequently

$$\langle D(y), D(y+\delta v) - D(y)\rangle \leq \delta \langle D(y), v\rangle + C_3\delta^2|v|^2 \tag{1.26}$$

with a constant $C_3 > 0$ independent of δ. Combining (1.25) with (1.26) we obtain the assertion. ∎

We make now an additional assumption about the geometry of the set Z.

Hypothesis 1.7. *The set Z is r-prox-regular with some $r > 0$, and there exist constants $R \geq 3$ and $\rho > 0$ such that*

$$\rho < \frac{r}{4(R+6)^2} \tag{1.27}$$

$$\forall x \in Z\ \exists x^* \in Z : |x - x^*| \leq R\rho, \quad B_{3\rho}(x^*) \subset Z. \tag{1.28}$$

Conditions (1.27)–(1.28) may look a bit complicated, however, that is just another formulation of the *interior cone condition* introduced in [12] in the form

$$\exists h > \sigma > 0\ \forall x \in Z\ \exists \bar{x} \in Z\ \forall \alpha \in [0,1] : (1-\alpha)x + \alpha\bar{x} + B_{\alpha\sigma}(0) \subset Z,\ |x - \bar{x}| \leq h. \tag{1.29}$$

The equivalence proof is given in [24] for different values of the constants. For the reader's convenience, we modify the proof for the present case.

Lemma 1.8. *Let Hypothesis 1.7 hold. Then the interior cone condition* (1.29) *is satisfied for $\sigma = \rho$ and $h = R\rho$. Conversely, if the set Z satisfies* (1.29), *then conditions* (1.27)–(1.28) *are fulfilled with R, ρ given in terms of h, σ.*

Proof. Assume for contradiction that conditions (1.27)–(1.28) hold and that (1.29) with $\sigma = \rho$ and $\bar{x} = x^*$ is violated. In other words, there exists $\alpha_0 \in [0,1]$ and $z \in B_1(0)$ with $z \neq 0$, such that $x_{\alpha_0} := x + \alpha_0(\bar{x} - x + \rho z) \notin Z$. Since $x_1 = \bar{x} + \rho z$ belongs to Z by hypothesis, we have that $\alpha_0 < 1$, and the segment connecting x_{α_0} and x_1 necessarily intersects the boundary ∂Z of Z. There exists therefore $\alpha \in (0,1)$ such that

$$x_\alpha := x + \alpha(\bar{x} - x) + \alpha\rho z \in \partial Z.$$

By Lemma 1.5 of [24], there exists $\xi \in X$, $|\xi| = 1$ such that $\operatorname{dist}(x_\alpha + r\xi, Z) = r$. By hypothesis, both x and $\bar{x} + 3\rho\xi$ belong to Z, hence

$$|x_\alpha + r\xi - x| \geq r,$$
$$|x_\alpha + (r - 3\rho)\xi - \bar{x}| \geq r.$$

In other words,

$$|\alpha(\bar{x} - x) + \alpha\rho z + r\xi| \geq r,$$
$$|(1-\alpha)(\bar{x} - x) - (r - 3\rho)\xi - \alpha\rho z| \geq r,$$

hence, using the triangle inequality and the fact that $|z| \leq 1$,

$$|\alpha(\bar{x} - x) + r\xi| \geq r - \alpha\rho,$$
$$|(1-\alpha)(\bar{x} - x) - (r - 3\rho)\xi| \geq r - \alpha\rho,$$

and, squaring both inequalities,

$$\alpha^2|\bar{x} - x|^2 + 2\alpha r \langle \xi, \bar{x} - x \rangle \geq -2\alpha r\rho + \alpha^2\rho^2,$$
$$(1-\alpha)^2|\bar{x} - x|^2 - 2(1-\alpha)(r - 3\rho) \langle \xi, \bar{x} - x \rangle \geq 6r\rho - 9\rho^2 - 2\alpha r\rho + \alpha^2\rho^2.$$

Taking into account that $r - 3\rho > 0$ by (1.27), we sum up suitable positive multiples of the two above inequalities and eliminate the term $\langle \xi, \bar{x} - x \rangle$ to obtain

$$(1-\alpha)\left(1 - \alpha + \alpha\frac{r - 3\rho}{r}\right)|\bar{x} - x|^2 \geq r\rho\left(6 - 2\alpha - 2(1-\alpha)\frac{r-3\rho}{r}\right) + \rho^2\left(-9 + \alpha^2 + \alpha(1-\alpha)\frac{r - 3\rho}{r}\right). \tag{1.30}$$

Therefore, since $(r - 3\rho)/r < 1$, from (1.28) and (1.30) we infer that

$$R^2\rho^2 \geq |\bar{x} - x|^2 \geq 4r\rho - 9\rho^2.$$

This implies that $(R^2 + 9)\rho \geq 4r$, which contradicts (1.27).

To prove the opposite implication, assume that (1.29) holds and that $x \in Z$ is arbitrarily chosen. It suffices to put $\rho = \alpha\sigma/3$, $x^* = (1-\alpha)x + \alpha\bar{x}$, $R = \alpha h/\rho = 3h/\sigma$, where $\alpha = \alpha(h, \sigma) > 0$ is chosen in such a way that

$$(R + 6)^2\rho = 3\alpha\left(\frac{h}{\sigma} + 2\right)^2\sigma < \frac{r}{4},$$

and the proof is complete. ■

2 Statement of the problem

We denote by $G(0, T; X)$ the space or regulated functions $u : [0, T] \to X$, that is, functions which admit both one-sided limits $u(t+)$, $u(t-)$ at each point $t \in [0, T]$ with the convention $u(0-) = u(0)$, $u(T+) = u(T)$. This is a Banach space with norm (see [2], Chapter II of [6], or [16, Proposition 4.1],)

$$|u|_{[0,T]} = \sup_{t\in[0,T]} |u(t)|. \tag{2.1}$$

The main object studied in this paper is the ODE

$$\varepsilon\dot{\xi}^\varepsilon(t) = \frac{f(|D(u(t) - \xi^\varepsilon(t))|)}{|D(u(t) - \xi^\varepsilon(t))|}D(u(t) - \xi^\varepsilon(t)), \qquad \xi^\varepsilon(0) = \xi^\varepsilon_0, \tag{2.2}$$

in a time interval $t \in (0,T)$ with a given initial condition $\xi_0^\varepsilon \in X$ such that $d(u(0+)-\xi_0^\varepsilon) < r$, and with unknown function ξ^ε, where $\varepsilon > 0$ is a a small parameter, $f : [0,r) \to \mathbb{R}_+$ is an increasing convex function with $f(0) = 0$, and $u \in G(0,T;X)$ is a given input function satisfying an additional hypothesis on the size of jumps. It is understood that $\varepsilon\dot{\xi}^\varepsilon(t) = 0$ whenever $D(u(t) - \xi^\varepsilon(t)) = 0$, thus (2.2) can be formally restated as the following problem.

Problem 2.1. *Assume that $\varepsilon > 0$, $\xi_0^\varepsilon \in X$, and that $f : [0,r) \to \mathbb{R}_+$ is an increasing convex function with $f(0) = 0$. Let $g : B_r(0) \to X$ be defined by the formula*

$$g(v) := \frac{f(|v|)}{|v|}v \quad \text{for } v \in B_r(0) \setminus \{0\}, \quad g(0) := 0. \tag{2.3}$$

Find an absolutely continuous function $\xi^\varepsilon : [0,T] \to X$ such that

$$\varepsilon\dot{\xi}^\varepsilon(t) = g(D(u(t) - \xi^\varepsilon(t))) \tag{2.4}$$

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

$$\xi^\varepsilon(0) = \xi_0^\varepsilon. \tag{2.5}$$

The mapping g is Lipschitz continuous on the ball $B_\delta := \{v \in X : |v| \le (1-\delta)r\}$ for each $\delta \in (0,1)$. Indeed, for $v, w \in B_\delta$ we have

$$|g(v) - g(w)|^2 \le (f(|v|) - f(|w|))^2 + \frac{f(|v|)f(|w|)}{|v||w|}|v - w|^2,$$

and the Lipschitz continuity follows from the convexity of f. Hence, by Lemma 1.4, also the mapping $g \circ D$ is Lipschitz continuous on the set $Z_\delta := \{y \in X : d(y) \le (1-\delta)r\}$. We put

$$K_\delta = \sup_{s \in (0,(1-\delta)r)} f(s), \quad L_\delta = \sup_{\substack{y_1,y_2 \in Z_\delta \\ y_1 \ne y_2}} \frac{|g(D(y_1)) - g(D(y_2))|}{|y_1 - y_2|}. \tag{2.6}$$

We now prove the existence and uniqueness of a local solution in an interval $[t_0, t_1]$ to (2.2) for each $\varepsilon > 0$ and each initial condition $\xi^\varepsilon(t_0)$ as long as $|D(u(t) - \xi^\varepsilon(t))| < r$.

Lemma 2.2. *Let $\alpha, \beta, \gamma, \delta, \varepsilon$ be positive numbers such that $\alpha + \beta + \gamma + \delta \le 1$, and assume that $u \in G(0,T;X)$, $\xi_*^\varepsilon \in X$, and $0 \le t_0 < t_0 + \sigma \le T$ are given such that $|u(t) - u(t_0+)| \le \alpha r$ for $t \in (t_0, t_0 + \sigma]$, and $d(u(t_0+) - \xi_*^\varepsilon) \le \beta r$. Put*

$$t_1 = t_0 + \min\left\{\sigma, \frac{\varepsilon\gamma r}{K_\delta}\right\},$$

where K_δ is defined in (2.6). Then there exists a unique absolutely continuous function ξ^ε in the set

$$\Sigma = \{\eta \in C([t_0,t_1];X) : \eta(t_0) = \xi_*^\varepsilon, \ |\eta(t) - \xi_*^\varepsilon| \le \gamma r \ \ \forall t \in [t_0,t_1]\}$$

such that (2.4) holds for a.e. $t \in [t_0, t_1]$. Moreover $d(u(t) - \xi^\varepsilon(t)) \le (1-\delta)r$ for every $t \in [t_0, t_1]$.

Proof. We prove that the mapping $\eta \in \Sigma \mapsto \xi^\varepsilon \in C([t_0, t_1]; X)$ defined by the formula

$$\varepsilon(\xi^\varepsilon(t) - \xi^\varepsilon_*) = \int_{t_0}^{t} g(D(u(\tau) - \eta(\tau)))\,\mathrm{d}\tau, \quad t \in [t_0, t_1] \tag{2.7}$$

is a contraction on Σ. The right-hand side of (2.7) is meaningful provided that $d(u(\tau) - \eta(\tau)) < r$ for $\tau \in [t_0, t_1]$. This is indeed the case. For $z \in Z$ we have that

$$\begin{aligned} |u(\tau) - \eta(\tau) - z| &\le |u(t_0+) - \eta(t_0) - z| + |u(\tau) - u(t_0+)| + |\eta(\tau) - \eta(t_0)| \\ &\le |u(t_0+) - \eta(t_0) - z| + (\alpha + \gamma)r, \end{aligned}$$

hence $d(u(\tau) - \eta(\tau)) \le (\alpha + \beta + \gamma)r \le (1 - \delta)r$. Moreover for $t \in [t_0, t_1]$ we have

$$|\xi^\varepsilon(t) - \xi^\varepsilon_*| \le \frac{(t - t_0)K_\delta}{\varepsilon} \le \gamma r, \tag{2.8}$$

hence $\xi^\varepsilon \in \Sigma$. To prove that the mapping $\eta \mapsto \xi^\varepsilon$ is a contraction, we consider $\eta_1, \eta_2 \in \Sigma$ and the corresponding $\xi^\varepsilon_1, \xi^\varepsilon_2$ given by (2.7). We have $u(t) - \eta_i(t) \in Z_\delta$ for $i = 1, 2$ and $t \in (t_0, t_1]$. Hence, by definition (2.6) of L_δ, we have for $t \in [t_0, t_1]$ the inequality

$$\varepsilon|\xi^\varepsilon_1(t) - \xi^\varepsilon_2(t)| \le L_\delta \int_{t_0}^{t} |\eta_1(\tau) - \eta_2(\tau)|\,\mathrm{d}\tau. \tag{2.9}$$

We define for $\nu > 0$ the following complete norm in $C([t_0, t_1]; X)$ by the formula

$$\|\eta\|_\nu = \max_{t \in [t_0, t_1]} \mathrm{e}^{-\nu t}|\eta(t)|. \tag{2.10}$$

It follows from (2.9) that

$$\mathrm{e}^{-\nu t}|\xi^\varepsilon_1(t) - \xi^\varepsilon_2(t)| \le \frac{L_\delta}{\varepsilon}\mathrm{e}^{-\nu t}\|\eta_1 - \eta_2\|_\nu \int_{t_0}^{t} \mathrm{e}^{\nu\tau}\,\mathrm{d}\tau = \frac{L_\delta}{\nu\varepsilon}(1 - \mathrm{e}^{-\nu(t - t_0)})\|\eta_1 - \eta_2\|_\nu.$$

It suffices to choose $\nu > L_\delta/\varepsilon$ to check that the mapping $\eta \mapsto \xi^\varepsilon$ is a contraction with respect to the norm $\|\cdot\|_\nu$, and the proof is complete. ■

Let us mention the following immediate consequence of Lemma 2.2.

Corollary 2.3. *Let $u \in G(0, T; X)$, $\varepsilon > 0$, and $\xi^\varepsilon_* \in X$ be given such that $d(u(0+) - \xi^\varepsilon_*) < r$. Then there exist $\tau^\varepsilon \in (0, T]$ and a unique absolutely continuous function $\xi^\varepsilon : [0, \tau^\varepsilon] \to X$ such that $\xi^\varepsilon(t_0) = \xi^\varepsilon_*$, (2.4) holds for a.e. $t \in [0, \tau^\varepsilon)$, and one of the following two alternatives holds:*

(i) $\tau^\varepsilon = T$;

(ii) $\tau^\varepsilon < T$, $d(u(\tau^\varepsilon+) - \xi^\varepsilon(\tau^\varepsilon)) \ge r$.

Proof of Corollary 2.3. The classical argument of [19, Theorem 3.1] guarantees the existence of a maximal solution to (2.5)–(2.4) characterized by the conditions (i), (ii), and its uniqueness follows from the contraction argument in Lemma 2.2. ■

In fact, we can restrict our considerations to left-continuous inputs u only. For the sake of completeness, we mention the following easy result.

Lemma 2.4. *Let $u \in G(0,T;X)$ be given, and let ξ^ε satisfy the identity*

$$\varepsilon(\xi^\varepsilon(t) - \xi^\varepsilon(t_0)) = \int_{t_0}^{t} g(D(u(\tau) - \xi^\varepsilon(\tau)))\,\mathrm{d}\tau \tag{2.11}$$

for $t \in [t_0, t_1]$. Put $\tilde{u}(t) = u(t-)$ for $t \in (t_0, t_1]$ and $\tilde{u}(t_0) = u(t_0)$ with the convention $u(0-) = u(0)$. Then we have

$$\varepsilon(\xi^\varepsilon(t) - \xi^\varepsilon(t_0)) = \int_{t_0}^{t} g(D(\tilde{u}(\tau) - \xi^\varepsilon(\tau)))\,\mathrm{d}\tau. \tag{2.12}$$

This is indeed an immediate consequence of the fact that u and $\tilde{u}$ coincide almost everywhere.

It is worth mentioning that equation (2.2) can be regarded as a gradient flow. Indeed, using Lemma 1.6, we derive for $y \in (Z + B_r(0)) \setminus Z$ the identity

$$\frac{f(|D(y)|)}{|D(y)|} D(y) = \frac{f(\sqrt{2\psi(y)})}{\sqrt{2\psi(y)}} \nabla\psi(y),$$

with $\psi(y) = 1/2d^2(y)$, and as a consequence (2.2) can be rewritten as

$$\varepsilon\dot{\xi}^\varepsilon(t) = \nabla\Psi(u(t) - \xi^\varepsilon(t)), \quad \Psi(y) = F(\psi(y)), \quad F(s) = \int_0^{\sqrt{2s}} f(\sigma)\,\mathrm{d}\sigma \ \text{ for } \ s \ge 0. \tag{2.13}$$

The gradient flow setting might be useful in some applications. Instead, in what follows, we shall systematically use the following equivalent variational formulation of Eq. (2.2).

Lemma 2.5. *An absolutely continuous function ξ^ε is a solution of (2.2) if and only if it satisfies almost everywhere the variational inequality*

$$\left\langle \dot{\xi}^\varepsilon(t), u(t) - \xi^\varepsilon(t) - \phi^\varepsilon(t) S(\dot{\xi}^\varepsilon(t)) - z \right\rangle + \frac{|\dot{\xi}^\varepsilon(t)|}{2r} |u(t) - \xi^\varepsilon(t) - \phi^\varepsilon(t) S(\dot{\xi}^\varepsilon(t)) - z|^2 \ge 0 \tag{2.14}$$

for every $z \in Z$ and $t \in [0,t]$, where $S(y) = 0$ for $y = 0$, $S(y) = y/|y|$ for $y \neq 0$, and $\phi^\varepsilon(t) = f^{-1}(\varepsilon|\dot{\xi}^\varepsilon(t)|)$ with f^{-1} denoting the inverse function to f.

Proof. Firstly observe that $\varepsilon|\xi^\varepsilon(t)| = f(|D(u(t) - \xi^\varepsilon(t))|)$ and consequently

$$\phi^\varepsilon(t) = f^{-1}(\varepsilon|\dot{\xi}^\varepsilon(t)|) = d(u(t) - \xi^\varepsilon(t)). \tag{2.15}$$

Therefore, an equivalent form of equation (2.2) reads

$$\phi^\varepsilon(t)S(\dot{\xi}^\varepsilon(t)) = D(u(t) - \xi^\varepsilon(t)) = u(t) - \xi^\varepsilon(t) - Q(u(t) - \xi^\varepsilon(t)) \quad \text{a. e.} \tag{2.16}$$

By using this together with Lemma 1.2, for $|\dot{\xi}^\varepsilon(t)| \neq 0$ we have

$$\left\langle \phi^\varepsilon(t)S(\dot{\xi}^\varepsilon(t)), u(t) - \xi^\varepsilon(t) - \phi^\varepsilon(t)S(\dot{\xi}^\varepsilon(t)) - z \right\rangle + \frac{\phi^\varepsilon(t)}{2r}|u(t) - \xi^\varepsilon(t) - \phi^\varepsilon(t)S(\dot{\xi}^\varepsilon(t)) - z|^2 \geq 0$$

for every $z \in Z$. Thus, the variational inequality (2.14) can be obtained by simply multiplying the inequality above by $|\dot{\xi}^\varepsilon(t)|/\phi^\varepsilon(t)$. ■

Following Lemma 2.4, we restrict the set of admissible inputs to a subset $\mathcal{U}$ of the set $G_L(0,T;X)$ of left-continuous regulated functions defined as follows:

$$u \in \mathcal{U} \iff |u(t+) - u(t)| < \frac{r}{6} \quad \forall t \in [0,T). \tag{2.17}$$

In the next sections, we prove the following results.

Theorem 2.6. *Let Hypothesis 1.7 hold, and let $u^* \in \mathcal{U}$ and $\xi_0^\varepsilon \in X$ be given such that $d(u^*(0+) - \xi_0^\varepsilon) < (r - \rho)/4$. We denote*

$$\mathcal{U}^* = \left\{ u \in \mathcal{U} : |u^* - u|_{[0,T]} \leq \frac{\rho}{4} \right\}.$$

Then there exists $\varepsilon_0 > 0$ such that for every $\varepsilon \in (0, \varepsilon_0)$ and for every $u \in \mathcal{U}^*$ Problem 2.1 has a unique global solution $\xi^\varepsilon \in W^{1,\infty}(0,T;X)$ such that $d(u(t) - \xi^\varepsilon(t)) \leq r/3$ for all $t \in [0,T]$. Moreover there exists a constant $V_0 > 0$ such that for every $u \in \mathcal{U}^*$ and every $\varepsilon \in (0, \varepsilon_0)$ we have

$$\operatorname*{Var}_{[0,T]} \xi^\varepsilon \leq V_0.$$

It is worth highlighting that the uniform bound for the output variation is obtained thanks to the interior cone condition similarly as in [12, 24]. If the solution is constructed via a time discretization process, then the uniform bound of the input variation automatically implies an uniform bound for the output variation, see [24]. The question whether this is valid also in the case of viscous approximations deserves to be studied in detail.

The mapping which with $u \in \mathcal{U}^*$ associates the solution ξ^ε to Problem 2.1 is locally Hölder continuous in the following sense.

Theorem 2.7. *Let the hypotheses of Theorem 2.6 hold. Then there exist* $\varepsilon_0 > 0$ *and a constant* $L > 0$ *independent of* ε *such that for every* $\varepsilon \in (0, \varepsilon_0)$ *and every* $u_1, u_2 \in \mathcal{U}^*$ *the solutions* $\xi_1^\varepsilon, \xi_2^\varepsilon$ *to Problem 2.1 associated with* u_1, u_2 *satisfy the inequality*

$$|\xi_1^\varepsilon - \xi_2^\varepsilon|_{[0,T]}^2 \le L\left(|\xi_1^\varepsilon(0) - \xi_2^\varepsilon(0)|^2 + |u_1 - u_2|_{[0,T]} + |u_1 - u_2|_{[0,T]}^2\right). \tag{2.18}$$

The uniqueness and existence of the solution to (0.4) stated in the previous Theorem 2.8 were proved in [24, Theorem 5.2] for right-continuous inputs. The conversion to the left-continuous case is easy and is shown in Section 6. The following result shows that the solution of (0.4) coincides with the viscous limit of ξ^ε as $\varepsilon \to 0$.

Theorem 2.8. *Given* $u \in \mathcal{U}$, *there exists a unique solution* $\xi \in G_L(0,T;X)$ *of the Kurzweil integral variational inequality* (0.4) *with initial condition* $\xi_0 = u(0) - x_0$. *Moreover, let* $\varepsilon_0 > 0$ *be given as in Theorem 2.6, and let* $\xi^\varepsilon \in W^{1,\infty}(0,T;X)$ *for* $\varepsilon \in (0, \varepsilon_0)$ *be the solution of Problem 2.1 with initial condition* $\xi_0^\varepsilon \in X$ *such that* $\xi_0^\varepsilon \to \xi_0$ *as* $\varepsilon \to 0$. *Then*

$$\lim_{\varepsilon \to 0} \xi^\varepsilon(t) = \xi(t) \quad \forall t \in [0,T].$$

The convergence in Theorem 2.8 cannot be expected to be uniform, since ξ^ε are continuous and ξ is not in general. The situation is different if u is continuous, and the convergence result reads as follows.

Theorem 2.9. *Given* $u \in C(0,T;X)$, *let* $\xi \in C(0,T;X)$ *be the solution of the Stieltjes integral variational inequality* (0.4) *with initial condition* $\xi_0 = u(0) - x_0$, *and let* $\xi^\varepsilon \in W^{1,\infty}(0,T;X)$ *be the solution of Problem 2.1 with initial condition* $\xi_0^\varepsilon \in X$ *such that* $\xi_0^\varepsilon \to \xi_0$ *as* $\varepsilon \to 0$. *Then*

$$\lim_{\varepsilon \to 0} |\xi^\varepsilon - \xi|_{[0,T]} = 0.$$

The proofs of Theorems 2.6, 2.7, 2.8, and 2.9 are given in the next sections.

3 Proof of Theorem 2.6

The proof of Theorem 2.6 is based on an iterative procedure related to the sequence of positive numbers defined recursively by the formula

$$\mu_i = (\mu_{i-1} + \theta)^2 \quad \text{for} \quad i \in \mathbb{N}, \qquad \mu_0 = \frac{1}{4} \in (\mu_*, \mu^*), \tag{3.1}$$

where $\theta \in (0, 1/4)$ is given and $0 < \mu_* < \mu^* < 1$ are the roots of the equation $\mu = (\mu + \theta)^2$. We have indeed

$$\mu^* = \frac{1}{2} - \theta + \sqrt{\frac{1}{4} - \theta} \in (1 - 3\theta, 1 - 2\theta), \ \mu_* = \frac{1}{2} - \theta - \sqrt{\frac{1}{4} - \theta} \in \left(\frac{\theta^2}{1 - 2\theta}, \theta\right). \tag{3.2}$$

By induction it is easily seen that the sequence $\{\mu_i\}$ is decreasing, thus $\lim_{i\to\infty}\mu_i = \mu_*$. For $\ell > 0$ we further denote

$$\varepsilon_\ell = \frac{\ell f(\mu_* r)}{r-\rho}, \tag{3.3}$$

and choose

$$\theta = \frac{R\rho}{r} < \frac{R}{4(R+6)^2} < \frac{1}{4R+48} \in \left(0, \frac{1}{60}\right) \tag{3.4}$$

by virtue of (1.27). In order to restrict the number of special cases to be distinguished, we assume that the domain of definition of functions $u \in \mathcal{U}$ is extended to $[0,\infty)$ by putting $u(t) = u(T)$ for $t > T$.

Lemma 3.1. *Let $0 \le a < T$ and $u \in \mathcal{U}$ be given, and let ξ^ε be the solution to Problem 2.1 in $[0,a]$ with the convention that only the initial condition $\xi^\varepsilon(0) = \xi_0^\varepsilon$ is prescribed if $a = 0$. Put $u_a = u(a+)$, $\xi_a^\varepsilon = \xi^\varepsilon(a)$, and let $\ell > 0$ be such that*

$$|u(t) - u_a| < \rho \quad \text{for} \quad t \in (a, a+\ell], \tag{3.5}$$

$$d(u_a - \xi_a^\varepsilon) = \lambda_0 r \tag{3.6}$$

with ρ from Hypothesis 1.7 and $\lambda_0 < \mu_i$ for some $i \in \mathbb{N}\cup\{0\}$, with μ_i defined in (3.1). Then for $\varepsilon < \varepsilon_\ell$ with ε_ℓ given by (3.3) the solution ξ^ε to (2.4)–(2.5) exists on $[0, a+\ell]$ and we have

$$d(u(t) - \xi^\varepsilon(t)) \le \lambda_0 r + (R+2)\rho \quad \text{for} \quad t \in [a, a+\ell]. \tag{3.7}$$

Furthermore, one of the following two situations occurs.

(i) $\lambda_0 \le \frac{1}{2}\sqrt{\frac{\rho}{r}}$. *Then*

$$\int_a^{a+\ell} |\dot{\xi}^\varepsilon(t)|\,\mathrm{d}t \le r - \rho.$$

(ii) $\lambda_0 > \frac{1}{2}\sqrt{\frac{\rho}{r}}$. *Then there exists a continuity point $a_1 \in (a, a+\ell)$ of u such that for μ_i from (3.1) with θ as in (3.4) we have*

$$d(u(a_1) - \xi^\varepsilon(a_1)) < \mu_{i+1} r,$$

$$\int_a^{a_1} |\dot{\xi}^\varepsilon(t)|\,\mathrm{d}t \le r - \rho.$$

Proof of Lemma 3.1. Note first that by (3.5) we have

$$|u(t) - u(s)| < 2\rho \quad \text{for} \quad a < s < t \le a+\ell. \tag{3.8}$$

By the hypothesis we have $d(u(a+) - \xi^\varepsilon(a)) = \lambda_0 r \le r/4$, thus thanks to Corollary 2.3, the solution ξ^ε to (2.4) can be uniquely extended from $[0,a]$ to a maximal interval $[0, t_\varepsilon]$.

We use Hypothesis 1.7 to find $x^* \in Z$ such that $|Q(u_a - \xi_a^\varepsilon) - x^*| \le R\rho$ and $B_{3\rho}(x^*) \subset Z$. Then, with the notation from (2.16) and using (3.5), we see that the point

$$z := u(t) - u_a + \rho S(\dot{\xi}^\varepsilon(t)) + x^*$$

is such that $|z - x^*| \le 2\rho$, hence it belongs to Z by Hypothesis 1.7. Consequently, z is an admissible choice in the variational inequality (2.14) and we obtain

$$\begin{aligned} &\left\langle \dot{\xi}^\varepsilon(t), u_a - \xi^\varepsilon(t) - (\rho + \phi^\varepsilon(t))S(\dot{\xi}^\varepsilon(t)) - x^* \right\rangle \\ &\quad + \frac{|\dot{\xi}^\varepsilon(t)|}{2r}|u_a - \xi^\varepsilon(t) - (\rho + \phi^\varepsilon(t))S(\dot{\xi}^\varepsilon(t)) - x^*|^2 \ge 0 \end{aligned} \tag{3.9}$$

for a. e. $t \in (a, t_\varepsilon)$. We rewrite (3.9) as

$$\begin{aligned} \left((\rho + \phi^\varepsilon(t)) - \frac{1}{2r}(\rho + \phi^\varepsilon(t))^2\right)|\dot{\xi}^\varepsilon(t)| &\le \left(1 - \frac{1}{r}(\rho + \phi^\varepsilon(t))\right)\left\langle \dot{\xi}^\varepsilon(t), u_a - \xi^\varepsilon(t) - x^* \right\rangle \\ &\quad + \frac{|\dot{\xi}^\varepsilon(t)|}{2r}|u_a - \xi^\varepsilon(t) - x^*|^2, \end{aligned} \tag{3.10}$$

or, putting $U(t) = |u_a - \xi^\varepsilon(t) - x^*|^2$,

$$|\dot{\xi}^\varepsilon(t)|\big((\rho + \phi^\varepsilon(t))(2r - (\rho + \phi^\varepsilon(t))) - U(t)\big) \le -(r - (\rho + \phi^\varepsilon(t)))\dot{U}(t). \tag{3.11}$$

By (2.15), (3.5) and (3.6) we have

$$\phi^\varepsilon(t) \le |u(t) - \xi^\varepsilon(t) - x^*| \le |u(t) - u_a| + |u_a - \xi^\varepsilon(t) - x^*| \le \rho + U^{1/2}(t), \tag{3.12}$$

$$U^{1/2}(a) = |u_a - \xi_a^\varepsilon - x^*| \le d(u_a - \xi_a^\varepsilon) + |Q(u_a - \xi_a^\varepsilon) - x^*| \le \lambda_0 r + R\rho. \tag{3.13}$$

Considering first Case (i), let us show that U is decreasing. The fact that $\lambda_0 \le \frac{1}{2}\sqrt{\frac{\rho}{r}}$ together with (3.13) and (1.27), yields

$$\begin{aligned} \rho(2r - \rho) - U(a) &\ge \rho(2r - \rho) - (\lambda_0 r + R\rho)^2 \ge \rho(2r - \rho) - \left(\frac{1}{2}\sqrt{\rho r} + R\rho\right)^2 \\ &= \rho\left(\frac{7}{4}r - (R^2 + 1)\rho - R\sqrt{\rho r}\right) \ge \rho r\left(\frac{7}{4} - \frac{R^2 + 1}{4(R + 6)^2} - \frac{R}{2(R + 6)}\right) \\ &\ge \rho r. \end{aligned} \tag{3.14}$$

The function $p \mapsto p(2r - p)$ is increasing in $[0, r]$, therefore from (3.11) we get

$$|\dot{\xi}^\varepsilon(t)|\big(\rho(2r - \rho) - U(t)\big) \le -(r - (\rho + \phi^\varepsilon(t)))\dot{U}(t) \qquad \text{for a.e. } t \in (a, t_\varepsilon). \tag{3.15}$$

By (3.14) there exists s_ε which is the largest number in $(a, t_\varepsilon]$ such that $\rho(2r-\rho)-U(t) > 0$ for $t \in (a, s_\varepsilon)$. Therefore, using (3.12) and (1.27) we get for $t \in (a, s_\varepsilon)$ that

$$r-(\rho+\phi^\varepsilon(t)) > r-2\rho-U^{1/2}(t) > r-2\rho-\sqrt{\rho(2r-\rho)} \geq r-2\rho - \frac{\sqrt{2}r}{2(R+6)} > 0, \qquad (3.16)$$

and (3.15) can be rewritten as

$$\frac{\dot{U}(t)}{\rho(2r-\rho)-U(t)} \leq -\frac{1}{r-(\rho+\phi^\varepsilon(t))}|\dot{\xi}^\varepsilon(t)| \leq -\frac{1}{r-\rho}|\dot{\xi}^\varepsilon(t)| \leq 0, \qquad (3.17)$$

for a.e. $t \in (a, s_\varepsilon)$. Hence, U is decreasing in (a, s_ε), and from (3.14) we infer that

$$\rho(2r-\rho) - U(t) \geq \rho(2r-\rho) - U(a) \geq \rho r, \qquad (3.18)$$

which shows that $s_\varepsilon = t_\varepsilon$. Moreover, we have by (3.12) and (3.13) that

$$d(u(t)-\xi^\varepsilon(t)) = \phi^\varepsilon(t) \leq \rho + U^{1/2}(a) \leq \lambda_0 r + (R+1)\rho \quad \text{for} \quad t \in (a, t_\varepsilon], \qquad (3.19)$$

which implies (3.7). In addition, according to Corollary 2.3, we must have $t_\varepsilon = a+\ell$. Indeed, if $t_\varepsilon < a+\ell$, the inequality above together with condition (i) and (3.8) would imply

$$d(u(t_\varepsilon+)-\xi^\varepsilon(t_\varepsilon)) \leq \phi^\varepsilon(t_\varepsilon) + |u(t_\varepsilon+)-u(t_\varepsilon)| < \frac{1}{2}\sqrt{\rho r} + (R+4)\rho < r,$$

which contradicts the maximality of t_ε. Besides, by (3.14), (3.17), and (3.18) we have

$$\begin{aligned}\int_a^{a+\ell} |\dot{\xi}^\varepsilon(t)|\,\mathrm{d}t &\leq \int_a^{a+\ell} \frac{(\rho-r)\dot{U}(t)}{\rho(2r-\rho)-U(t)}\,\mathrm{d}t = (r-\rho)\log\left(\frac{\rho(2r-\rho)-U(a+\ell)}{\rho(2r-\rho)-U(a)}\right)\\ &\leq (r-\rho)\log\left(\frac{2r-\rho}{r}\right) \leq (r-\rho)\log 2 \leq r-\rho. \end{aligned} \qquad (3.20)$$

which concludes the proof of Case (i).

Before passing to Case (ii), let us check that with the notation (3.2), for θ as in (3.4) we have

$$\lambda^* := \frac{1}{2}\sqrt{\frac{\rho}{r}} > \theta > \mu_*. \qquad (3.21)$$

Indeed, the fact that $\mu_* < \theta$ follows from (3.2). Furthermore, by (1.27) we have that $\theta/\lambda^* = 2R\sqrt{\rho/r} \leq R/(R+6) < 1$, and (3.21) follows.

Let us now consider Case (ii) and assume that $\mu_i > \lambda_0 > \lambda^*$ for some $i \in \mathbb{N}\cup\{0\}$. Put $\lambda_1 := (\lambda_0+\theta)^2 \in (\mu_*, \mu_{i+1})$. Since $d(u_a - \xi^\varepsilon_a) = \lambda_0 r$ and $\lambda_1 < \lambda_0$, there exists necessarily τ_ε which is the largest number in $(a, t_\varepsilon]$ such that

$$\phi^\varepsilon(t) = d(u(t)-\xi^\varepsilon(t)) \geq \lambda_1 r \quad \text{for} \quad t \in (a, \tau_\varepsilon). \qquad (3.22)$$

As the function $p \mapsto p(2r-p)$ is increasing in $[0,r]$, the inequality (3.11) yields

$$|\dot{\xi}^\varepsilon(t)|\big(\phi^\varepsilon(t)(2r-\phi^\varepsilon(t)) - U(t)\big) \le -(r-(\rho+\phi^\varepsilon(t)))\dot{U}(t), \tag{3.23}$$

while (3.22) implies

$$\phi^\varepsilon(t)(2r-\phi^\varepsilon(t)) \ge r^2\lambda_1(2-\lambda_1) \quad \text{for} \quad t \in (a,\tau_\varepsilon). \tag{3.24}$$

From (3.23) it follows that

$$|\dot{\xi}^\varepsilon(t)|\big(r^2\lambda_1(2-\lambda_1) - U(t)\big) \le -(r-(\rho+\phi^\varepsilon(t)))\dot{U}(t) \tag{3.25}$$

for a.e. $t \in (a,\tau_\varepsilon)$. Noting that $\lambda_1 < \lambda_0 \le 1/4$, we know by virtue of (3.13) that

$$U(a) \le (\lambda_0 r + R\rho)^2 = r^2(\lambda_0+\theta)^2 = r^2\lambda_1 < r^2\lambda_1(2-\lambda_1). \tag{3.26}$$

Hence, there exists σ_ε, the largest number in $(a,\tau_\varepsilon]$, such that $r^2\lambda_1(2-\lambda_1) - U(t) > 0$ for $t \in (a,\sigma_\varepsilon)$. Therefore, using (3.12), we have for $t \in (a,\sigma_\varepsilon)$ that

$$r-(\rho+\phi^\varepsilon(t)) > r-2\rho-U^{1/2}(t) > r-2\rho-r\sqrt{\lambda_1(2-\lambda_1)} > r-2\rho-\frac{1}{\sqrt{2}}r > 0,$$

and (3.25) implies $\dot{U}(t) < 0$ for $t \in (a,\sigma_\varepsilon)$. Hence $U(t)$ is decreasing in (a,σ_ε), and it follows from (3.26) that $\sigma_\varepsilon = \tau_\varepsilon$. From (3.25) we then obtain a counterpart of (3.17) in the form

$$\frac{\dot{U}(t)}{r^2\lambda_1(2-\lambda_1)-U(t)} \le -\frac{1}{r-\rho}|\dot{\xi}^\varepsilon(t)| \le 0 \tag{3.27}$$

for a. e. $t \in (a,\tau_\varepsilon)$. Integrating over (a,τ_ε) and using the relation $r^2\lambda_1(2-\lambda_1) - U(a) \ge r^2\lambda_1(2-\lambda_1) - \lambda_1 r^2 = r^2\lambda_1(1-\lambda_1)$ we thus get

$$\begin{aligned}\int_a^{\tau_\varepsilon} |\dot{\xi}^\varepsilon(t)|\,\mathrm{d}t &\le \int_a^{\tau_\varepsilon} \frac{(\rho-r)\dot{U}(t)}{r^2\lambda_1(2-\lambda_1)-U(t)}\,\mathrm{d}t = (r-\rho)\log\left(\frac{r^2\lambda_1(2-\lambda_1)-U(\tau_\varepsilon)}{r^2\lambda_1(2-\lambda_1)-U(a)}\right)\\ &\le (r-\rho)\log\left(\frac{2-\lambda_1}{1-\lambda_1}\right) \le (r-\rho)\log\frac{7}{3}.\end{aligned} \tag{3.28}$$

We claim that $\tau_\varepsilon < a+\ell$. Indeed, assuming that $\tau_\varepsilon = a+\ell$, since $\lambda_1 > \mu_*$ (with μ_* from (3.2)), then the inequality

$$\varepsilon|\dot{\xi}^\varepsilon(t)| = f(d(u(t)-\xi^\varepsilon(t))) \ge f(\lambda_1 r) > f(\mu_* r)$$

holds for a. e. $t \in (a,a+\ell)$, which together with (3.28) yields

$$r-\rho \ge \int_a^{a+\ell} |\dot{\xi}^\varepsilon(t)|\,\mathrm{d}t > \frac{\ell}{\varepsilon}f(\mu_* r),$$

in contradiction with the fact that $\varepsilon < \varepsilon_\ell$, with ε_ℓ as in (3.3). Furthermore, recalling the identity (2.2) and using the continuity of f, we can find a constant $C_r > 0$ independent of ε such that $|\dot{\xi}^\varepsilon(t)| \le C_r/\varepsilon$ in $[a, t_\varepsilon)$. By the choice of τ_ε in (3.22), for any $\kappa > 0$, we can find a continuity point $a_1 \in (\tau_\varepsilon, \tau_\varepsilon + \kappa\varepsilon)$ of u such that $d(u(a_1) - \xi^\varepsilon(a_1)) < \lambda_1 r < \mu_{i+1} r$, thus (3.28) yields

$$\int_a^{a_1} |\dot{\xi}^\varepsilon(t)| \,\mathrm{d}t \le (r - \rho) \log \frac{7}{3} + \kappa C_r.$$

Choosing $\kappa > 0$ sufficiently small we obtain the assertion (ii).

To conclude the proof, it remains to show that, also in Case (ii), i. e., when $\lambda_0 > \lambda^*$, we have $t_\varepsilon = a + \ell$ and the inequality (3.7) also holds in $(a_1, a + \ell]$. Noting that (3.7) holds true in $(a, \tau_\varepsilon]$, it follows that $\phi^\varepsilon(\tau_\varepsilon) < \lambda_1 r + (R + 2)\rho$, thus using (3.8) we obtain

$$d(u(\tau_\varepsilon+) - \xi^\varepsilon(\tau_\varepsilon)) \le \phi^\varepsilon(\tau_\varepsilon) + |u(\tau_\varepsilon+) - u(\tau_\varepsilon)| < \frac{r}{4} + (R + 4)\rho < r$$

which ensures in particular that $\tau_\varepsilon < t_\varepsilon$ (see Corollary 2.3). Choosing a_1 sufficiently close to τ_ε, we may assume that

$$d(u(a_1) - \xi^\varepsilon(a_1)) < \mu_{i+1} r, \quad d(u(t) - \xi^\varepsilon(t)) < \mu_{i+1} r + (R + 2)\rho \;\text{ for }\; t \in [a, a_1]. \tag{3.29}$$

We now repeat the procedure in the interval $[a_1, a + \ell]$. Firstly, note that the values of ξ^ε in the interval $[a_1, t_\varepsilon)$ remain unchanged if we replace $u(t)$ with $\bar{u}(t) = u(t)$ for $t \le a + \ell$, $\bar{u}(t) = u(a + \ell)$ for $t \ge a + \ell$, so that

$$|\bar{u}(t) - \bar{u}(s)| < 2\rho \;\text{ for }\; a_1 < s < t \le a_1 + \ell.$$

In the interval $[a_1, a + \ell]$, we use Hypothesis 1.7 to find $x_1^* \in Z$ such that $|Q(u(a_1) - \xi^\varepsilon(a_1)) - x_1^*| \le R\rho$ and $B_{3\rho}(x_1^*) \subset Z$. We proceed as above and distinguish Case (i) if $\lambda_1 \le \lambda^*$, or Case (ii) if $\lambda^* < \lambda_1 < \mu_{i+1}$. In Case (i) we stop the algorithm as we infer that (3.19) holds and Corollary 2.3 guarantees that $t_\varepsilon = a + \ell$. In Case (ii) we find $a_2 \in (a_1, a_1 + \ell)$ such that

$$d(u(a_2) - \xi^\varepsilon(a_2)) < \mu_{i+2} r, \quad d(u(t) - \xi^\varepsilon(t)) < \mu_{i+2} r + (R + 2)\rho \;\text{ for }\; t \in [a_1, a_2]$$

in analogy to (3.29). We continue by induction over $i \in \mathbb{N}$, and since by (3.21) we have $\mu_* = \lim_{i\to\infty} \mu_i < \lambda^*$, after finitely many steps only Case (i) remains, which completes the proof. ∎

Lemma 3.1 provides us with the tools to build the solution of the viscous problem by moving forward with interval steps of length controlled by a fixed number ℓ. To determine such a number, we first need to isolate the points where jump discontinuities exceed some given value. Invoking the definition of the set $\mathcal{U}^*$ in Theorem 2.6, we find a division $0 = \hat{t}_0 < \hat{t}_1 < \cdots < \hat{t}_N = T$ of the interval $[0, T]$ such that the implication

$$|u^*(t+) - u^*(t)| \ge \frac{\rho}{3} \implies \exists\, i \in \{0, \dots, N\} : \; t = \hat{t}_i \tag{3.30}$$

holds for all $t \in [0,T]$. Indeed, since u^* is a left-continuous regulated function, it cannot have infinitely many jumps of size exceeding $\rho/3$, see e.g. [16, Proposition 2.4]. We claim that there exists $\hat{\ell} > 0$ such that

$$\forall i \in \{1,\dots,N\}\ \forall u \in \mathcal{U}^*\ \forall a \in [\hat{t}_{i-1},\hat{t}_i)\ \forall t \in (a, a+\hat{\ell}] \cap (a,\hat{t}_i] : |u(t) - u(a+)| \leq \rho. \quad (3.31)$$

Indeed, if (3.31) does not hold, then there exists $i \in \{1,\dots,N\}$ such that

$$\forall n \in \mathbb{N}\ \exists u_n \in \mathcal{U}^*\ \exists s_n \in \left(\hat{t}_{i-1},\hat{t}_i\right)\ \exists t_n \in (s_n,\hat{t}_i),\ t_n - s_n < \frac{1}{n} : |u_n(t_n) - u_n(s_n)| > \rho, \quad (3.32)$$

hence

$$|u^*(t_n) - u^*(s_n)| \geq |u_n(t_n) - u_n(s_n)| - \frac{\rho}{2} \geq \frac{\rho}{2},$$

which contradicts the hypothesis (3.30).

We refer again to the sequence μ_i defined in (3.1) for θ as in (3.4). Recall that its limit μ_* is smaller than the critical value $\frac{1}{2}\sqrt{\rho/r}$ by virtue of (3.21). Hence, we find $K \in \mathbb{N}$ such that

$$\mu_k r < \frac{1}{2}\sqrt{\rho r} \quad \text{for} \quad k \geq K, \quad (3.33)$$

and put

$$\ell = \min\left\{\hat{\ell}, \frac{1}{K}(\hat{t}_i - \hat{t}_{i-1}),\ i = 1,\dots,N\right\} \quad (3.34)$$

with $\hat{\ell}$ from (3.31). Before we proceed with the analytic proof of Theorem 2.6, it might be helpful to say a few words about the idea behind it. To benefit from the estimate in (3.31), on each subinterval $[\hat{t}_{i-1},\hat{t}_i]$ we construct the solution 'piece by piece' in intervals of length at most $\ell \leq \hat{\ell}$ by applying Lemma 3.1. For that, one should observe that the occurrences of Cases (i) and (ii) listed in Lemma 3.1 show a certain pattern classified in terms of the function

$$\lambda^*(t) := \frac{1}{r} d(u(t+) - \xi^\varepsilon(t)) \quad (3.35)$$

which characterizes the distance of $u - \xi^\varepsilon$ from the boundary of Z and which has to be kept between 0 and 1. The strategy is the following, see Figure 1.

1. In each interval $[\hat{t}_{i-1},\hat{t}_i]$ we start with $a = \hat{t}_{i-1}$ and check whether $\lambda^*(a) \leq \frac{1}{2}\sqrt{\frac{\rho}{r}}$ (Case (i)) or $\lambda^*(a) > \frac{1}{2}\sqrt{\frac{\rho}{r}}$ (Case (ii)).

2. If $\lambda^*(a) > \frac{1}{2}\sqrt{\frac{\rho}{r}}$, then by a repeated argument of Case (ii) of Lemma 3.1, we find a sequence $a = a_0 < a_1 < a_2 \dots$ such that $a_j - a_{j-1} < \ell$ and $\lambda^*(a_j) < \lambda^*(a_{j-1})$, and we show that after K steps at most we pull $\lambda^*(t)$ down below the critical value $\frac{1}{2}\sqrt{\frac{\rho}{r}}$. We call this phase of the proof the *active regime*.

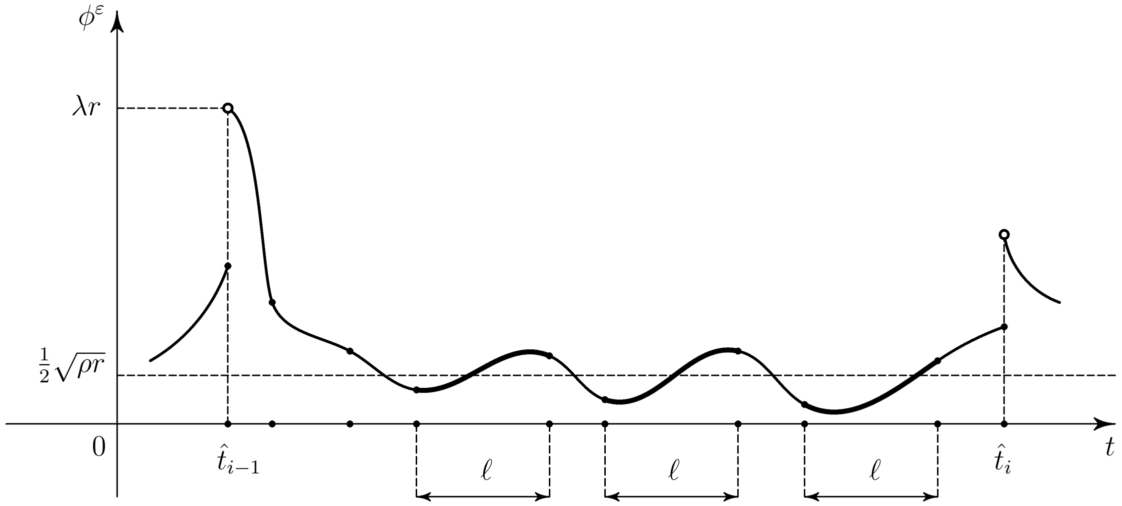

Figure 1: Illustration to the proof of Theorem 2.6: Bold lines represent the passive phase.

3. As soon as $\lambda^*(\bar{t}) \le \frac{1}{2}\sqrt{\frac{\rho}{r}}$ at some point $\bar{t} \in [\hat{t}_{i-1}, \hat{t}_i - \ell]$, we see that $\lambda^*(t)$ remains small in $[\bar{t}, \bar{t}+\ell]$ according to Case (i) of Lemma 3.1. We say that we are in the *passive regime*.

4. At the point $\bar{t}+\ell$ we have again the alternative $\lambda^*(\bar{t}+\ell) \le \frac{1}{2}\sqrt{\frac{\rho}{r}}$ (Case (i)) or $\lambda^*(\bar{t}+\ell) > \frac{1}{2}\sqrt{\frac{\rho}{r}}$ (Case(ii)).

5. In Case (i), we extend the solution as before to the interval $[\bar{t}+\ell, \bar{t}+2\ell]$.

6. In Case (ii), we prove that $\lambda^*(\bar{t}+\ell)$ is 'not too far' from the critical value $\frac{1}{2}\sqrt{\frac{\rho}{r}}$, so that after just one iteration of Case (ii) we come back to Case (i) again.

7. Altogether, by definition of ℓ in (3.34), at most $3K$ iterations are sufficient for reaching the endpoint $\hat{t}_i$.

8. We continue the same procedure in the interval $[\hat{t}_i, \hat{t}_{i+1}]$, and after finitely many steps we prove the existence of a global solution and we get the desired bound for the variation on the whole interval $[0, T]$.

Let us carry out this program in detail.

Proof of Theorem 2.6. With $\varepsilon_0 = \varepsilon_\ell$ as in (3.3) and K as in (3.33), we choose $\varepsilon < \varepsilon_\ell$ and an arbitrary $u \in \mathcal{U}^*$. Let $0 = \hat{t}_0 < \cdots < \hat{t}_N = T$ be a division as in (3.30). We proceed by induction and fix an arbitrary interval $[\hat{t}_{i-1}, \hat{t}_i]$ for $i \in \{1, \ldots, N\}$, assuming that $d(u(\hat{t}_{i-1}+) - \xi^\varepsilon(\hat{t}_{i-1})) \le \frac{r}{4}$.

Consider first the case that $d(u(\hat{t}_{i-1}+) - \xi^\varepsilon(\hat{t}_{i-1})) = \lambda r$ with $\frac{1}{4} \ge \lambda > \frac{1}{2}\sqrt{\frac{\rho}{r}}$. By Case (ii) of Lemma 3.1, there exists $a_1 \in (\hat{t}_{i-1}, \hat{t}_{i-1}+\ell)$ such that $d(u(a_1) - \xi^\varepsilon(a_1)) = \lambda_1 r$ with

$\lambda_1 \le \mu_1$ in terms of the sequence $\{\mu_i\}$ defined by (3.1). If now $\lambda_1 > \frac{1}{2}\sqrt{\frac{\rho}{r}}$, we repeat the procedure and find $a_2 \in (a_1, a_1 + \ell)$ such that $d(u(a_2) - \xi^\varepsilon(a_2)) = \lambda_2 r$ with $\lambda_2 \le \mu_2$. By induction, after $m \le K$ steps, we find $\bar t = a_m$ such that $d(u(\bar t) - \xi^\varepsilon(\bar t)) = \lambda_m r$ with $\lambda_m \le \frac{1}{2}\sqrt{\frac{\rho}{r}}$ and finish the active phase.

Starting from the point $\bar t$, we apply Case (i) of Lemma 3.1 and check that for $t \in [\bar t, \bar t + \ell] \cap [\hat t_{i-1}, \hat t_i]$ we have

$$d(u(t) - \xi^\varepsilon(t)) \le \bar\lambda r \quad \text{with} \quad \bar\lambda = \frac{1}{2}\sqrt{\frac{\rho}{r}} + (R+2)\frac{\rho}{r}.$$

If $\bar t + \ell < \hat t_i$, either $\lambda^*(\bar t + \ell) \le \frac{1}{2}\sqrt{\frac{\rho}{r}}$ or we have Case (ii). In both situations we check that there exist $\bar b \in [\bar t + \ell, \bar t + 2\ell)$ such that $d(u(\bar b) - \xi^\varepsilon(\bar b)) = \bar\mu r$ with

$$\bar\mu < \frac{1}{r^2}(\bar\lambda r + R\rho)^2 < \frac{1}{2}\sqrt{\frac{\rho}{r}},$$

meaning that we stay in/return to the passive regime in the interval $[b, b+\ell]$.

Alternating possibly Case (i) and Case (ii) we fill successively the whole interval $[\hat t_{i-1}, \hat t_i]$ keeping $d(u(t) - \xi^\varepsilon(t))$ far away from the critical value r. The end point $\hat t_i$ is either achieved during the passive regime or, in the least favorable case, we have the situation illustrated in Figure 1. More precisely, we reach a point t^* at the end of the passive phase such that $t^* < \hat t_i < t^* + \ell$ with $\lambda^*(t^*) \ge \frac{1}{2}\sqrt{\frac{\rho}{r}}$. Applying Lemma 3.1 in the interval $[t^*, t^* + \ell]$ with u replaced by the truncation $\bar u(t) = u(t)$ for $t \le \hat t_i$, $\bar u(t) = u(\hat t_i)$ for $t \in (\hat t_i, t^* + \ell]$, we guarantee the existence of solution in $[t^*, \hat t_i]$, while (3.7) implies

$$d(u(\hat t_i) - \xi^\varepsilon(\hat t_i)) \le \lambda^*(t^*) r + (R+2)\rho \le \frac{1}{2}\sqrt{\rho r} + 2(R+2)\rho. \tag{3.36}$$

where the last inequality follows from the fact t^* is the end of the passive phase, therefore

$$\lambda^*(t^*) r = d(u(t^*) - \xi^\varepsilon(t^*)) \le \frac{1}{2}\sqrt{\rho r} + (R+2)\rho.$$

By (2.17) and (3.36) we have

$$\begin{aligned} d(u(\hat t_i +) - \xi^\varepsilon(\hat t_i)) &< \frac{r}{6} + \frac{1}{2}\sqrt{\rho r} + 2(R+2)\rho = r\left(\frac{1}{6} + \frac{1}{2}\sqrt{\frac{\rho}{r}} + 2(R+2)\frac{\rho}{r}\right) \\ &< r\left(\frac{1}{6} + \frac{1}{4(R+6)} + \frac{1}{2(R+6)}\right) \le \frac{r}{4}, \end{aligned}$$

which allow us to repeat the above procedure in the subsequent subinterval and by induction over $i = 1, \dots, N$ we construct a global solution to (2.2) which, thanks to (3.7), satisfies

$$d(u(t) - \xi^\varepsilon(t)) \le \frac{r}{4} + (R+2)\rho \le \frac{r}{3}$$

for all $t \in [0,T]$.

In each interval $[\hat{t}_{i-1}, \hat{t}_i]$, there are at most K intervals corresponding to the passive phase, and at most $2K$ intervals corresponding to the active phase. The total variation of ξ^ε over each of these intervals is smaller or equal to $r-\rho$ according to Lemma 3.1. Hence,

$$\operatorname*{Var}_{[0,T]} \xi^\varepsilon \le 3KN(r-\rho)$$

which completes the proof of Theorem 2.6. ■

Let us mention an immediate consequence of Theorem 2.6.

Corollary 3.2. *Let Hypothesis 1.7 hold, and let $u^* \in \mathcal{U}$ and $\xi_0^\varepsilon \in X$ be given such that $d(u^*(0+) - \xi_0^\varepsilon) < (r-\rho)/4$. Assume that $\{u_n : n \in \mathbb{N}\} \subset G_L(0,T;X)$ is a sequence such that*

$$\lim_{n\to\infty} |u_n - u^*|_{[0,T]} = 0,$$

and let ξ_n^ε be the solutions of (2.2) *corresponding to inputs u_n and initial conditions $\xi_n^\varepsilon(0)$ such that $d(u_n(0+) - \xi_n^\varepsilon(0)) < r/4$ for all ε and n. Then there exist $\varepsilon_0 > 0$, $n_0 \in \mathbb{N}$, and a constant $V_0 > 0$ such that*

$$\operatorname*{Var}_{[0,T]} \xi_n^\varepsilon \le V_0 \quad \text{for } \; n \ge n_0 \; \text{ and } \; \varepsilon \in (0,\varepsilon_0).$$

4 Proof of Theorem 2.7

We start with an elementary inequality. Let f be as in (2.2). Then for every $v, w \in X$ we have

$$\left\langle \frac{f(|v|)}{|v|} v - \frac{f(|w|)}{|w|} w, v-w \right\rangle \ge \frac{1}{2}\left(\frac{f(|v|)}{|v|} + \frac{f(|w|)}{|w|}\right)|v-w|^2 \tag{4.1}$$

with the convention $\frac{f(s)}{s} = f'(0+)$ for $s = 0$.

Indeed, for $|v| > 0, |w| > 0$ we have

$$\begin{aligned}\left\langle \frac{f(|v|)}{|v|} v - \frac{f(|w|)}{|w|} w, v-w \right\rangle &= |v|f(|v|) + |w|f(|w|) - \left(\frac{f(|v|)}{|v|} + \frac{f(|w|)}{|w|}\right)\langle v, w\rangle \\ &= |v|f(|v|) + |w|f(|w|) + \frac{1}{2}\left(\frac{f(|v|)}{|v|} + \frac{f(|w|)}{|w|}\right)(|v-w|^2 - |v|^2 - |w|^2),\end{aligned}$$

while the convexity of f ensures that

$$\begin{aligned} 2(|v|f(|v|) + |w|f(|w|)) &- \left(\frac{f(|v|)}{|v|} + \frac{f(|w|)}{|w|}\right)(|v|^2 + |w|^2) \\ &= |v|f(|v|) + |w|f(|w|) - \frac{f(|v|)}{|v|}|w|^2 - \frac{f(|w|)}{|w|}|v|^2 \\ &= \left(\frac{f(|v|)}{|v|} - \frac{f(|w|)}{|w|}\right)(|v|^2 - |w|^2) \ge 0 \end{aligned}$$

hence,

$$|v|f(|v|) + |w|f(|w|) - \frac{1}{2}\left(\frac{f(|v|)}{|v|} + \frac{f(|w|)}{|w|}\right)(|v|^2 + |w|^2) \geq 0,$$

and (4.1) follows. The case $v = 0$ or $w = 0$ can be obtained from the previous computation by taking the limit as v or w tends to 0.

We now proceed to the proof of the Hölder type continuity of the solution mapping $u \mapsto \xi^\varepsilon$ associated with Problem 2.1.

Proof of Theorem 2.7. Given $u_1, u_2 \in \mathcal{U}^*$, let $\varepsilon \in (0, \varepsilon_0)$ be arbitrarily chosen, with ε_0 from Theorem 2.6, and let $\xi_1^\varepsilon, \xi_2^\varepsilon$ be the solutions of (2.2) corresponding to inputs u_1, u_2 and initial conditions $\xi_1^\varepsilon(0), \xi_2^\varepsilon(0)$. As in (2.14), we can write the following variational inequalities

$$\left\langle \dot\xi_1^\varepsilon, u_1 - \xi_1^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon - z_1 \right\rangle + \frac{|\dot\xi_1^\varepsilon|}{2r}\left|u_1 - \xi_1^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon - z_1\right|^2 \geq 0, \tag{4.2}$$

$$\left\langle \dot\xi_2^\varepsilon, u_2 - \xi_2^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon - z_2 \right\rangle + \frac{|\dot\xi_2^\varepsilon|}{2r}\left|u_2 - \xi_2^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon - z_2\right|^2 \geq 0 \tag{4.3}$$

for all $z_1, z_2 \in Z$. We choose here in particular $z_2 = u_1 - \xi_1^\varepsilon - (f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)/|\dot\xi_1^\varepsilon|)\dot\xi_1^\varepsilon$, $z_1 = u_2 - \xi_2^\varepsilon - (f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)/|\dot\xi_2^\varepsilon|)\dot\xi_2^\varepsilon$. Summing up (4.2)–(4.3) and using the triangle inequality together with the classical inequality $(a+b)^2 \leq (1+\delta)a^2 + (1+(1/\delta))b^2$ for all $a, b, \delta > 0$ we obtain

$$\begin{aligned}
&\left\langle \dot\xi_2^\varepsilon - \dot\xi_1^\varepsilon, \xi_2^\varepsilon - \xi_1^\varepsilon \right\rangle + \left\langle \frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon, \xi_2^\varepsilon - \xi_1^\varepsilon \right\rangle \\
&\leq \left\langle \dot\xi_2^\varepsilon - \dot\xi_1^\varepsilon, u_2 - u_1 \right\rangle + \frac{|\dot\xi_2^\varepsilon| + |\dot\xi_1^\varepsilon|}{2r}\left|(u_2 - u_1) - (\xi_2^\varepsilon - \xi_1^\varepsilon) - \frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon + \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon\right|^2 \\
&\leq \left\langle \dot\xi_2^\varepsilon - \dot\xi_1^\varepsilon, u_2 - u_1 \right\rangle + \left(1 + \frac{1}{\delta}\right)\frac{|\dot\xi_2^\varepsilon| + |\dot\xi_1^\varepsilon|}{2r}\left|(u_2 - u_1) - (\xi_2^\varepsilon - \xi_1^\varepsilon)\right|^2 \\
&\quad + \frac{(1+\delta)(|\dot\xi_2^\varepsilon| + |\dot\xi_1^\varepsilon|)}{2r}\left|\frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon\right|^2
\end{aligned} \tag{4.4}$$

for all $\delta > 0$. We claim that for $0 < \delta \leq 1/2$ we have

$$\begin{aligned}
&\left\langle \frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon, \dot\xi_2^\varepsilon - \dot\xi_1^\varepsilon \right\rangle \\
&\qquad \geq \frac{(1+\delta)(|\dot\xi_2^\varepsilon| + |\dot\xi_1^\varepsilon|)}{2r}\left|\frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon\right|^2
\end{aligned} \tag{4.5}$$

a. e. in $(0,T)$. In order to prove this claim let us denote

$$L_\varepsilon = \left\langle \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon, \dot\xi_1^\varepsilon - \dot\xi_2^\varepsilon \right\rangle, \tag{4.6}$$

$$R_\varepsilon = \frac{|\dot\xi_1^\varepsilon| + |\dot\xi_2^\varepsilon|}{2r}\left| \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon \right|^2, \tag{4.7}$$

and let us check that we are in the situation of (4.1) with

$$v = \frac{f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)}{|\dot\xi_1^\varepsilon|}\dot\xi_1^\varepsilon, \quad w = \frac{f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)}{|\dot\xi_2^\varepsilon|}\dot\xi_2^\varepsilon.$$

Indeed, noting that $|v| = f^{-1}(\varepsilon|\dot\xi_1^\varepsilon|)$, $\varepsilon\dot\xi_1^\varepsilon = \frac{f(|v|)}{|v|}v$, $|w| = f^{-1}(\varepsilon|\dot\xi_2^\varepsilon|)$, $\varepsilon\dot\xi_2^\varepsilon = \frac{f(|w|)}{|w|}w$, we can write (4.6) and (4.7) in the form

$$L_\varepsilon = \frac{1}{\varepsilon}\left\langle \frac{f(|v|)}{|v|}v - \frac{f(|w|)}{|w|}w, v - w \right\rangle, \tag{4.8}$$

$$R_\varepsilon = \frac{1}{2r\varepsilon}(f(|v|) + f(|w|))|v-w|^2. \tag{4.9}$$

By Theorem 2.6 and recalling (2.15), we know that

$$\begin{aligned} |v(t)| + |w(t)| &= f^{-1}(\varepsilon|\dot\xi_1^\varepsilon(t)|) + f^{-1}(\varepsilon|\dot\xi_2^\varepsilon(t)|) \\ &= d(u_1(t) - \xi_1^\varepsilon(t)) + d(u_2(t) - \xi_2^\varepsilon(t)) \le \frac{2r}{3} \end{aligned}$$

for $t \in [0,T]$. Hence, using (4.1),

$$R_\varepsilon \le \frac{1}{3\varepsilon}\frac{f(|v|) + f(|w|)}{|v| + |w|}|v-w|^2 \le \frac{1}{3\varepsilon}\left(\frac{f(|v|)}{|v|} + \frac{f(|w|)}{|w|}\right)|v-w|^2 \le \frac{2}{3}L_\varepsilon. \tag{4.10}$$

This shows that (4.5) holds for every $0 < \delta \le 1/2$. Now choosing $\delta = 1/2$ we can reduce (4.4) to

$$\begin{aligned} \left\langle \dot\xi_2^\varepsilon - \dot\xi_1^\varepsilon, \xi_2^\varepsilon - \xi_1^\varepsilon \right\rangle &\le \left\langle \dot\xi_2^\varepsilon - \dot\xi_1^\varepsilon, u_2 - u_1 \right\rangle + \frac{3}{2r}(|\dot\xi_2^\varepsilon| + |\dot\xi_1^\varepsilon|)\,|(u_2 - u_1) - (\xi_2^\varepsilon - \xi_1^\varepsilon)|^2 \\ &\le \max\left\{1, \frac{3}{r}\right\}(|\dot\xi_2^\varepsilon| + |\dot\xi_1^\varepsilon|)\left(|u_2 - u_1| + |u_2 - u_1|^2 + |\xi_2^\varepsilon - \xi_1^\varepsilon|^2\right). \end{aligned} \tag{4.11}$$

This is an inequality of the form

$$\dot y(t) \le \alpha(t) y(t) + \gamma(t), \tag{4.12}$$

where $\alpha(t) = 2\max\{1, r/3\}(|\dot{\xi}_2^\varepsilon(t)| + |\dot{\xi}_1^\varepsilon(t)|)$, $\gamma(t) = \alpha(t)(|u_2(t) - u_1(t)| + |u_2(t) - u_1(t)|^2)$. We have $\alpha, \gamma \in L^1(0,T)$. Putting $A(t) = \int_0^t \alpha(\tau)\,\mathrm{d}\tau$, by Theorem 2.6 we have $A(t) \le 4\max\{1, r/3\}V_0$, and from (4.12) we deduce that

$$y(t) \le \mathrm{e}^{A(t)} y(0) + \int_0^t \mathrm{e}^{A(t)-A(\tau)} \gamma(\tau)\,\mathrm{d}\tau. \tag{4.13}$$

In terms of (4.11) this yields

$$\sup_{t\in[0,T]} |\xi_2^\varepsilon(t) - \xi_1^\varepsilon(t)|^2 \le C\left(|\xi_2^\varepsilon(0) - \xi_1^\varepsilon(0)|^2 + \sup_{t\in[0,T]} |u_2(t) - u_1(t)| + \sup_{t\in[0,T]} |u_2(t) - u_1(t)|^2\right)$$

with a constant $C > 0$ independent of ε, which we wanted to prove. ■

5 Explicit solutions for piecewise constant inputs

Using Lemma 1.5, we can find the solution to (2.2) in closed form in every interval where the input u is constant. The result reads as follows.

Proposition 5.1. *Let $0 \le t_* < t^* \le T$ be arbitrary, and assume that there exists $\bar{u} \in X$ such that $u(t) = \bar{u}$ for each $t \in (t_*, t^*)$. Assume furthermore that $\xi^\varepsilon : [0, t_*] \to X$ satisfies (2.4) for a.e. $t \in [0, t_*]$. If $d(\bar{u} - \xi^\varepsilon(t_*)) =: d_* \in [0, r)$, then the solution ξ^ε can be extended to $[0, t^*]$, and for $t \in [t_*, t^*]$ we have*

$$\xi^\varepsilon(t) = \begin{cases} \xi^\varepsilon(t_*) & \text{if } d_* = 0, \\ \xi^\varepsilon(t_*) + (1 - \alpha(t)) D(\bar{u} - \xi^\varepsilon(t_*)) & \text{if } d_* > 0, \end{cases} \tag{5.1}$$

where $\alpha : [t_*, t^*] \to (0, \infty)$ is the solution to the ODE

$$\varepsilon\dot{\alpha}(t) + \frac{1}{d_*} f(d_*\alpha(t)) = 0, \quad \alpha(t_*) = 1, \tag{5.2}$$

and the inequality $d(u(t) - \xi^\varepsilon(t)) \le d_*$ holds for all $t \in [t_*, t^*)$.

Proof. We directly check that the function ξ^ε defined by (5.1) is a solution to (2.4). The case $d_* = 0$ is trivial. For $d_* > 0$ we have

$$\varepsilon\dot{\xi}^\varepsilon(t) = -\varepsilon\dot{\alpha}(t) D(\bar{u} - \xi^\varepsilon(t_*))$$

for all $t \in (t_*, t^*)$. On the other hand, the function α is decreasing, therefore $0 < \alpha(t) \le 1$ and we can apply Lemma 1.5 so that

$$D(\bar{u} - \xi^\varepsilon(t)) = D(\bar{u} - \xi^\varepsilon(t_*) + (\alpha(t) - 1) D(\bar{u} - \xi^\varepsilon(t_*))) = \alpha(t) D(\bar{u} - \xi^\varepsilon(t_*)). \tag{5.3}$$

Hence, $d(\bar{u} - \xi^\varepsilon(t)) = d_*\alpha(t)$ for all $t \in (t_*, t^*)$, and (2.2) follows from (5.2). ■

Corollary 5.2. *Let $u \in G_L(0,T;X)$ be a step function of the form*

$$u(t) = u_0 \chi_{\{0\}}(t) + \sum_{j=1}^{m} u_j \chi_{(t_{j-1},t_j]}(t) \tag{5.4}$$

corresponding to a division $0 = t_0 < \cdots < t_m = T$ of $[0,T]$, where χ_S denotes the characteristic function of a set $S \subset [0,T]$, that is, $\chi_S(t) = 1$ if $t \in S$, $\chi_S(t) = 0$ if $t \notin S$. Assume that the given elements $u_0, u_1, \ldots u_m$ from X satisfy the condition

$$r^* := \max\{|u_j - u_{j-1}|; j = 1, \ldots, m\} < r. \tag{5.5}$$

Let the initial condition ξ_0^ε be such that $d(u_0 - \xi_0^\varepsilon) = 0$. Then there exists $\varepsilon_0 > 0$ such that for every $\varepsilon \in (0, \varepsilon_0)$ the global solution of $\xi^\varepsilon \in W^{1,\infty}(0,T;X)$ to Problem 2.1 exists, and in each interval $[t_{j-1}, t_j]$ is given by a formula of the form (5.1)*. Moreover, $d(u(t) - \xi^\varepsilon(t)) \le (r + r^*)/2$ for all $t \in [0,T]$.*

Proof. The statement will follow from Proposition 5.1 by an induction argument over $j = 1, \ldots, m$. We first check that

$$d_1 := d(u_1 - \xi^\varepsilon(0)) \le |(u_1 - \xi^\varepsilon(0)) - (u_0 - \xi^\varepsilon(0))| = |u_1 - u_0| \le r^*, \tag{5.6}$$

where we have used the hypothesis that $u_0 - \xi^\varepsilon(0) \in Z$. Furthermore, as induction step, assume that for some $j = 1, \ldots, m-1$ the solution is available on $[0, t_{j-1}]$ and

$$d_j := d(u_j - \xi^\varepsilon(t_{j-1})) \le \frac{r + r^*}{2}. \tag{5.7}$$

Then the solution can be extended to $[0, t_j]$ according to Proposition 5.1 by the formula

$$\xi^\varepsilon(t) = \begin{cases} \xi^\varepsilon(t_{j-1}) & \text{if } d_j = 0, \\ \xi^\varepsilon(t_{j-1}) + (1 - \alpha_j(t))\, D(u_j - \xi^\varepsilon(t_{j-1})) & \text{if } d_j > 0, \end{cases} \tag{5.8}$$

where the function $\alpha_j : [t_{j-1}, t_j] \to (0, \infty)$ is defined as the solution to the differential equation as a counterpart to (5.2)

$$\varepsilon \dot\alpha_j(t) + \frac{1}{d_j} f(d_j \alpha_j(t)) = 0, \quad \alpha_j(t_{j-1}) = 1, \tag{5.9}$$

and by Proposition 5.1 we have $|D(u(t) - \xi^\varepsilon(t))| \le d_j$ for all $t \in [t_{j-1}, t_j)$.

The induction step will be complete if we prove that

$$d(u_{j+1} - \xi^\varepsilon(t_j)) \le \frac{r + r^*}{2}. \tag{5.10}$$

holds provided ε is sufficiently small. We argue as in (5.6) in the case $d_j = 0$. For $d_j > 0$ we have by (5.9) that

$$\int_{d_j\alpha_j(t_j)}^{d_j} \frac{\mathrm{d}s}{f(s)} = -\int_{t_{j-1}}^{t_j} \frac{d_j\dot{\alpha}_j(t)}{f(d_j\alpha_j(t))}\,\mathrm{d}t = \frac{1}{\varepsilon}(t_j - t_{j-1}). \tag{5.11}$$

For $\sigma \in (0, r]$ put

$$\hat{F}(\sigma) = \int_{\sigma}^{r} \frac{\mathrm{d}s}{f(s)}. \tag{5.12}$$

Then $\hat{F}$ is decreasing in $(0, r]$, $\hat{F}(r) = 0$, and, since f is convex, $\hat{F}(0+) = +\infty$. Put

$$\hat{F}^* = \hat{F}\left(\frac{r - r^*}{2}\right), \quad \varepsilon_0 = \frac{1}{\hat{F}^*} \min_{j=1,\dots,m} (t_j - t_{j-1}). \tag{5.13}$$

By virtue of (5.11) we have for $\varepsilon < \varepsilon_0$ that

$$\hat{F}(d_j\alpha_j(t_j)) = \frac{1}{\varepsilon}(t_j - t_{j-1}) + \hat{F}(d_j) > \hat{F}^*, \tag{5.14}$$

hence $d_j\alpha_j(t_j) < (r - r^*)/2$. From (5.3) it follows that

$$d(u_j - \xi^\varepsilon(t_j)) = \alpha(t_j) d(u_j - \xi^\varepsilon(t_{j-1})) = \alpha(t_j) d_j < \frac{r - r^*}{2}. \tag{5.15}$$

Let $z_j \in Z$ be such that $|u_j - \xi^\varepsilon(t_j) - z_j| = d(u_j - \xi^\varepsilon(t_j))$. Then

$$d(u_{j+1} - \xi^\varepsilon(t_j)) \le |u_{j+1} - \xi^\varepsilon(t_j) - z_j| \le |u_j - \xi^\varepsilon(t_j) - z_j| + |u_{j+1} - u_j| < \frac{r - r^*}{2} + r^* = \frac{r + r^*}{2},$$

so that (5.10) holds and the induction step is complete. ■

6 Proof of Theorem 2.8

In [24], we have proved the existence of a unique solution to (0.4) in the case that the input u is right-continuous. The conversion of the result to left-continuous inputs is easy. Consider $u \in G_L(0, T; X)$ and an initial condition $x_0 \in Z$. We look for $\xi \in BV_L(0, T; X)$ such that

$$\int_0^T \langle u(t+) - \xi(t+) - z(t), \mathrm{d}\xi(t)\rangle + \frac{1}{2r}\int_0^T |u(t+) - \xi(t+) - z(t)|^2 \,\mathrm{d}V(\xi(t)) \ge 0 \tag{6.1}$$

for all $z \in G(0, T; Z)$, $\xi(0) = u(0) - x_0$. To this end consider the function

$$\bar{u}(t) = u(t+) \quad \text{for} \quad t \in [0, T],$$

with the convention that $u(T+) = u(T)$. Then $\bar{u}$ is right-continuous, and by [24, Theorem 4.2] there exists a unique $\bar{\xi} \in BV_R(0,T;X)$ such that

$$\int_0^T \left\langle \bar{u}(t) - \bar{\xi}(t) - z(t),\, \mathrm{d}\bar{\xi}(t) \right\rangle + \frac{1}{2r} \int_0^T |\bar{u}(t) - \bar{\xi}(t) - z(t)|^2 \, \mathrm{d}V(\bar{\xi}(t)) \geq 0 \tag{6.2}$$

for all $z \in G(0,T;Z)$, $\bar{\xi}(0) = \bar{u}(0) - \bar{x}_0$, where

$$\bar{x}_0 = Q(x_0 + \bar{u}(0) - u(0)). \tag{6.3}$$

We claim that ξ can be constructed as follows

Lemma 6.1. *The variational inequality* (6.1) *is satisfied for*

$$\xi(t) = \begin{cases} \bar{\xi}(t-) & \text{for } t \in (0,T], \\ u(0) - x_0 & \text{for } t = 0. \end{cases} \tag{6.4}$$

The proof relies on the following elementary result.

Lemma 6.2. *Let $v \in G(0,T;X)$ and $w \in BV(0,T;X)$ be given such that the set $A = \{t \in [0,T] : w(t) \neq 0\}$ is countable. Then*

$$\int_0^T \langle v(t),\, \mathrm{d}w(t) \rangle = \langle v(T), w(T) \rangle - \langle v(0), w(0) \rangle .$$

This is indeed obvious if the set A is finite. The general case is obtained by passing to the limit following the same argument used in the proof of [26, Lemma 6.3.16], whose result concerns the case of real-valued functions.

Proof of Lemma 6.1. We have by definition that $u(t+) = \bar{u}(t)$, $\xi(t+) = \bar{\xi}(t)$ for all $t \in [0,T)$. Besides, noting that $\bar{u}(T-) = \bar{u}(T)$, it follows that $\bar{\xi}$ as well as $V(\bar{\xi})$ are left-continuous in T (see [24, Lemma 5.1]). Let A be the countable set of all $t \in [0,T]$ such that $\xi(t) \neq \bar{\xi}(t)$. For all $t \notin A$ (i.e. if $\xi(t) = \bar{\xi}(t)$) we also have

$$V(\xi)(t) = \bar{V}(\bar{\xi})(t),$$

where $\bar{V}(\bar{\xi})(t) = V(\bar{\xi})(t) + |\bar{\xi}(0) - \xi(0)|$. It follows from (6.2) that

$$\int_0^T \left\langle \bar{u}(t) - \bar{\xi}(t) - z(t),\, \mathrm{d}\bar{\xi}(t) \right\rangle + \frac{1}{2r} \int_0^T |\bar{u}(t) - \bar{\xi}(t) - z(t)|^2 \, \mathrm{d}\bar{V}(\bar{\xi})(t) \geq 0 \tag{6.5}$$

for all $z \in G(0,T;Z)$. Hence, using (6.5) and Lemma 6.2 we obtain for any arbitrary function $z \in G(0,T;Z)$ that

$$\int_0^T \langle u(t+) - \xi(t+) - z(t),\, \mathrm{d}\xi(t)\rangle + \frac{1}{2r}\int_0^T |u(t+) - \xi(t+) - z(t)|^2 \,\mathrm{d}V(\xi(t))$$
$$= \int_0^T \big\langle \bar u(t) - \bar\xi(t) - z(t),\, \mathrm{d}\xi(t)\big\rangle + \frac{1}{2r}\int_0^T |\bar u(t) - \bar\xi(t) - z(t)|^2 \,\mathrm{d}V(\xi)(t)$$
$$\geq \int_0^T \big\langle \bar u(t) - \bar\xi(t) - z(t),\, \mathrm{d}(\xi - \bar\xi)(t)\big\rangle + \frac{1}{2r}\int_0^T |\bar u(t) - \bar\xi(t) - z(t)|^2 \,\mathrm{d}(V(\xi) - \bar V(\bar\xi))(t)$$
$$= -\big\langle \bar u(0) - \bar\xi(0) - z(0), (\xi - \bar\xi)(0)\big\rangle - |\bar u(0) - \frac{1}{2r}\bar\xi(0) - z(0)|^2 (V(\xi) - \bar V(\bar\xi))(0),$$
$$= \big\langle \bar u(0) - \bar\xi(0) - z(0), \bar\xi(0) - \xi(0)\big\rangle + \frac{1}{2r}|\bar u(0) - \bar\xi(0) - z(0)|^2 |\bar\xi(0) - \xi(0)|, \tag{6.6}$$

where we have used the fact that $\xi(T) = \bar\xi(T)$, $V(\xi)(T) = \bar V(\bar\xi)(T)$.

By definition (6.3) of $\bar x_0$ we have

$$\langle x_0 - \bar x_0 + \bar u(0) - u(0), \bar x_0 - z\rangle + \frac{1}{2r}|\bar x_0 - z|^2 |x_0 - \bar x_0 + \bar u(0) - u(0)| \geq 0$$

for all $z \in Z$, while by construction $\bar x_0 - z = \bar u(0) - \bar\xi(0) - z$ and $x_0 - \bar x_0 + \bar u(0) - u(0) = \bar\xi(0) - \xi(0)$. We thus conclude from (6.6) that (6.1) holds, and Lemma 6.1 is proved. ■

In Corollary 5.2 we have derived an explicit formula for the solution of (2.2) if the input u is a left-continuous step function. Likewise, for such particular inputs, the solution of the Kurzweil variational inequality (6.1) is again described by step function which can be constructed via an iterative process as the one presented in [24]; a type of catching up algorithm. These are the ingredients of the proof of the following result.

Lemma 6.3. *Let the hypotheses of Corollary 5.2 hold, and let $u(t)$ be a step function given by (5.4) for $t \in [0,T]$. Given $x_0 \in Z$, put $\xi_0 = u_0 - x_0$ and let $\xi_0^\varepsilon \in X$ for $\varepsilon > 0$ be such that $\xi_0^\varepsilon \to \xi_0$ as $\varepsilon \to 0$. If $\xi^\varepsilon \in W^{1,\infty}(0,T;X)$ is the unique solution of (2.2) such that $\xi^\varepsilon(0) = \xi_0^\varepsilon$, then*

$$\lim_{\varepsilon\to 0} \xi^\varepsilon(t) = \xi(t) \qquad \forall t \in [0,T].$$

where $\xi \in BV_L(0,T;X)$ is the solution of (6.1) given by

$$\xi(t) = \xi_0 \chi_{\{0\}}(t) + \sum_{j=1}^m \xi_j \chi_{(t_{j-1},t_j]}(t) \tag{6.7}$$

with $\xi_0 = u_0 - x_0$, and

$$\xi_j = \xi_{j-1} + D(u_j - \xi_{j-1}) \quad \textit{for } j = 1, \dots, m. \tag{6.8}$$

Proof. Let $\omega > 0$ be given. We claim that for every $t \in [0,T]$ there exists $\bar{\varepsilon}(t) > 0$ such that

$$|\xi^{\varepsilon}(t) - \xi(t)| < \omega \quad \text{for } \varepsilon < \bar{\varepsilon}(t). \tag{6.9}$$

The statement is obvious for $t = 0$. In order to prove it in $(0,T]$, firstly observe that according to Corollary 5.2, on each subinterval $(t_{j-1}, t_j]$ the function ξ^{ε} is given by (5.8). Therefore, with the notation from (5.7), for $t \in (t_{j-1}, t_j]$ we have

$$|\xi^{\varepsilon}(t) - \xi(t)| \le |\xi^{\varepsilon}(t_{j-1}) - \xi(t_{j-1})| + |D(u_j - \xi^{\varepsilon}(t_{j-1})) - D(u_j - \xi(t_{j-1}))| + d_j\alpha_j(t), \tag{6.10}$$

where α_j is the function satisfying (5.9). By Corollary 5.2 we know that $d(u_j - \xi^{\varepsilon}(t_{j-1})) \le (r + r^*)/2$. For $j = 1$, we get

$$d(u_1 - \xi(t_0)) \le |u_1 - \xi_0 - x_0| = |u_1 - u_0| \le r^*,$$

while for $j \ge 2$, the identity (6.8) gives $\xi_{j-1} = u_{j-1} - Q(u_{j-1} - \xi_{j-2})$ and we get

$$d(u_j - \xi(t_{j-1})) \le |u_j - \xi_{j-1} - Q(u_{j-1} - \xi_{j-2})| = |u_j - u_{j-1}| \le r^*.$$

where the last inequality holds thanks to (5.5). In either case, we have $d(u_j - \xi(t_{j-1})) \le r^* < (r + r^*)/2$. Applying Lemma 1.4 we thus obtain

$$|D(u_j - \xi^{\varepsilon}(t_{j-1})) - D(u_j - \xi(t_{j-1}))| \le K|\xi^{\varepsilon}(t_{j-1}) - \xi(t_{j-1})|$$

with $K = 1 + \sqrt{3}/\kappa$, $\kappa = \sqrt{\frac{2r}{r+r^*} - 1}$; consequently, we deduce from (6.10) the following estimate for the solutions ξ^{ε} and ξ

$$|\xi^{\varepsilon}(t) - \xi(t)| \le (1+K)|\xi^{\varepsilon}(t_{j-1}) - \xi(t_{j-1})| + d_j\alpha_j(t) \quad \text{for } t \in (t_{j-1}, t_j]. \tag{6.11}$$

Let us fix an arbitrary $t^* \in (t_{k-1}, t_k]$ for some $k \in \{1, \dots, m\}$, and prove the existence of $\bar{\varepsilon}(t^*) > 0$ so that (6.9) holds. Put

$$\tau^* = \min\{t^* - t_{k-1}, \min\{t_j - t_{j-1} : j = 1, \dots, k-1\}\}. \tag{6.12}$$

Considering $\hat{F}$ as in (5.12), for each $j \in \{1, \dots, k\}$ such that $d_j > 0$ we have $\hat{F}(d_j) > 0$ and similarly as in (5.11) and (5.14) we infer that

$$\hat{F}(d_j\alpha_j(t)) > \frac{1}{\varepsilon}(t - t_{j-1}) \quad \text{for } t \in (t_{j-1}, t_j]$$

This together with the definition of τ^* thus ensures that

$$d_j\alpha(t_j) \le \hat{F}^{-1}\left(\frac{\tau^*}{\varepsilon}\right), \quad d_k\alpha_k(t^*) \le \hat{F}^{-1}\left(\frac{\tau^*}{\varepsilon}\right).$$

Finally, we conclude from (6.11) that

$$|\xi^\varepsilon(t^*) - \xi(t^*)| \le (1+K)^m |\xi_0^\varepsilon - \xi_0| + \frac{(1+K)^m}{K} \hat{F}^{-1}\left(\frac{\tau^*}{\varepsilon}\right), \tag{6.13}$$

and (6.9) follows. ■

We now are ready to prove that, for regulated inputs, the limit of the solutions ξ^ε to (2.2), as $\varepsilon \to 0$, satisfies the variational inequality (6.1).

Proof of Theorem 2.8. Consider an arbitrary $u \in \mathcal{U}$. We find a sequence $\{u_n : n \in \mathbb{N}\}$ of step functions such that $|u_n - u|_{[0,T]} \to 0$ as $n \to \infty$, $u_n(0) = u(0)$. By Theorem 2.6 and Corollary 3.2 for $u^* = u$, there exist $\varepsilon_0 > 0$ and $n_0 \in \mathbb{N}$ such that for $\varepsilon \in (0, \varepsilon_0)$ and $n \ge n_0$, the Problem 2.1 corresponding to inputs u and u_n has a solution which we denote by ξ^ε and ξ_n^ε, respectively (with $\xi^\varepsilon(0) = \xi_n^\varepsilon(0) = \xi_0^\varepsilon$). Moreover, we can find a constant V_0 independent of n such that

$$\operatorname*{Var}_{[0,T]} \xi_n^\varepsilon \le V_0, \quad \operatorname*{Var}_{[0,T]} \xi^\varepsilon \le V_0 \quad \text{for } \varepsilon \in (0, \varepsilon_0). \tag{6.14}$$

Let ξ and ξ_n be the solutions to (0.4) associated with the inputs u and u_n, respectively. Given $\omega > 0$, by Theorem 2.7 we find $n_1 \in \mathbb{N}$, $n_1 > n_0$, such that

$$|\xi_n^\varepsilon - \xi^\varepsilon|_{[0,T]} \le \frac{\omega}{4} \quad \text{for } n \ge n_1, \ \varepsilon \in (0, \varepsilon_0). \tag{6.15}$$

Similarly, by [24, Theorem 4.5], we can find $n_2 \in \mathbb{N}$, $n_2 > n_0$, such that

$$|\xi_n - \xi|_{[0,T]} \le \frac{\omega}{4} \quad \text{for } n \ge n_2. \tag{6.16}$$

The convergence $\xi^\varepsilon(t) \to \xi(t)$ is obvious for $t = 0$. Let now $t \in (0, T]$ be arbitrary. For $\varepsilon \in (0, \varepsilon_0)$ and $n = \max\{n_1, n_2\}$ we have

$$|\xi^\varepsilon(t) - \xi(t)| \le |\xi^\varepsilon(t) - \xi_n^\varepsilon(t)| + |\xi_n^\varepsilon(t) - \xi_n(t)| + |\xi_n(t) - \xi(t)| \le \frac{\omega}{2} + |\xi_n^\varepsilon(t) - \xi_n(t)|.$$

We refer to Lemma 6.3 and find $\varepsilon = \varepsilon(t) > 0$ sufficiently small such that $|\xi_n^\varepsilon(t) - \xi_n(t)| < \omega/2$. Since $\omega > 0$ is arbitrary, we obtain the assertion. ■

7 Proof of Theorem 2.9

Let us start with the following variant of the Young inequality.

Lemma 7.1. *Let $h : \mathbb{R}_+ \to \mathbb{R}_+$ be a concave strictly increasing function, $h(0) = 0$, and let $H(x) = \int_0^x h(s)\,\mathrm{d}s$. Then for all positive numbers x, y, σ we have*

$$x\, h(y) \le \sigma \left(H(y) + H\left(\frac{x}{\sigma}\right)\right). \tag{7.1}$$

Proof. It is enough to prove that (7.1) holds for continuously differentiable functions h. The general case is then obtained by approximating the function h uniformly by smooth concave increasing functions and passing to the limit.

Put $\beta(s) = \sigma s(h^{-1})'(s)$ for $s \geq 0$. The function h^{-1} is convex, hence β is increasing. From the classical Young inequality it follows that

$$x\, h(y) \leq \int_0^x \beta^{-1}(s)\, \mathrm{d}s + \int_0^{h(y)} \beta(s)\, \mathrm{d}s.$$

By convexity of h^{-1} we have $\beta(z) \geq \sigma h^{-1}(z)$ for all $z \geq 0$, hence $\beta^{-1}(s) \leq h(s/\sigma)$ for all $s \geq 0$, which yields

$$\int_0^x \beta^{-1}(s)\, \mathrm{d}s \leq \int_0^x h\left(\frac{s}{\sigma}\right)\, \mathrm{d}s = \sigma H\left(\frac{x}{\sigma}\right).$$

In the second integral we substitute $s = h(z)$ and obtain

$$\int_0^{h(y)} \beta(s)\, \mathrm{d}s = \sigma \int_0^{h(y)} s(h^{-1})'(s)\, \mathrm{d}s = \sigma \int_0^y h(z)(h^{-1})'(h(z))h'(z)\, \mathrm{d}z = \sigma H(y),$$

which completes the proof. ■

Proof of Theorem 2.9. Let $u \in C([0,T];X)$ be given. We choose a sequence $\{u_n : n \in \mathbb{N}\} \subset W^{1,2}(0,T;X)$ such that

$$\lim_{n\to\infty} |u_n - u|_{[0,T]} = 0.$$

By [24, Corollary 5.3], for $n \in \mathbb{N}$ the solution ξ_n to (0.4) associated with u_n with initial condition $\xi_n(0) = u_n(0) - x_0$ belongs to $W^{1,2}(0,T;X)$ and satisfies almost everywhere the variational inequality

$$u_n - \xi_n \in Z, \quad \left\langle \dot{\xi}_n, u_n - \xi_n - z \right\rangle + \frac{|\dot{\xi}_n|}{2r}|u_n - \xi_n - z|^2 \geq 0 \ \text{ a. e.} \tag{7.2}$$

for every $z \in Z$. Applying Theorem 2.6 and Corollary 3.2 for $u^* = u$, we know that the solution ξ^ε to the Problem 2.1 with inputs u_n and initial conditions $\xi_n^\varepsilon(0) = u_n(0) - x_0$ exists for $\varepsilon \in (0, \varepsilon_0)$ and $n \geq n_0$. Besides, it satisfies

$$\left\langle \dot{\xi}_n^\varepsilon, u_n - \xi_n^\varepsilon - \frac{f^{-1}(\varepsilon|\dot{\xi}_n^\varepsilon|)}{|\dot{\xi}_n^\varepsilon|}\dot{\xi}_n^\varepsilon - z \right\rangle + \frac{|\dot{\xi}_n^\varepsilon|}{2r}\left| u_n - \xi_n^\varepsilon - \frac{f^{-1}(\varepsilon|\dot{\xi}_n^\varepsilon|)}{|\dot{\xi}_n^\varepsilon|}\dot{\xi}_n^\varepsilon - z \right|^2 \geq 0 \tag{7.3}$$

for every $z \in Z$. Choosing in (7.2) $z = u_n(t \pm s) - \xi_n(t \pm s)$ for $s > 0$, dividing by s and letting s tend to 0 we see that the quadratic term vanishes in the limit, and we get

$$\left\langle \dot{\xi}_n(t), \dot{u}_n(t) - \dot{\xi}_n(t) \right\rangle = 0, \quad \text{hence} \ \ |\dot{\xi}_n(t)| \leq |\dot{u}_n(t)| \tag{7.4}$$

for a. e. $t \in (0,T)$. Moreover, putting $z = \xi_n - u_n$ in (7.3), $z = u_n - \xi_n^\varepsilon - \frac{f^{-1}(\varepsilon|\dot\xi_n^\varepsilon|)}{|\dot\xi_n^\varepsilon|}\dot\xi_n^\varepsilon$ in (7.2), summing the inequalities up we get

$$\left\langle \frac{f^{-1}(\varepsilon|\dot\xi_n^\varepsilon|)}{|\dot\xi_n^\varepsilon|}\dot\xi_n^\varepsilon, \dot\xi_n^\varepsilon - \dot\xi_n \right\rangle + \left\langle \dot\xi_n^\varepsilon - \dot\xi_n, \xi_n^\varepsilon - \xi_n \right\rangle \leq \frac{|\dot\xi_n^\varepsilon| + |\dot\xi_n|}{2r}\left| \xi_n^\varepsilon - \xi_n + \frac{f^{-1}(\varepsilon|\dot\xi_n^\varepsilon|)}{|\dot\xi_n^\varepsilon|}\dot\xi_n^\varepsilon \right|^2,$$

wherefrom the classical inequality $(a+b)^2 \leq (1+\delta)a^2 + (1+(1/\delta))b^2$ for $a, b, \delta > 0$ yields

$$\left\langle \frac{f^{-1}(\varepsilon|\dot\xi_n^\varepsilon|)}{|\dot\xi_n^\varepsilon|}\dot\xi_n^\varepsilon, \dot\xi_n^\varepsilon - \dot\xi_n \right\rangle + \left\langle \dot\xi_n^\varepsilon - \dot\xi_n, \xi_n^\varepsilon - \xi_n \right\rangle \leq \left(1 + \frac{1}{\delta}\right)\frac{|\dot\xi_n^\varepsilon| + |\dot\xi_n|}{2r}|\xi_n^\varepsilon - \xi_n|^2 + \frac{(1+\delta)(|\dot\xi_n^\varepsilon| + |\dot\xi_n|)}{2r}\left(f^{-1}(\varepsilon|\dot\xi_n^\varepsilon|)\right)^2. \tag{7.5}$$

We define functions $H : [0,\infty) \to \mathbb{R}$ and $\hat H : X \to \mathbb{R}$ by the formula

$$H(x) = \int_0^x f^{-1}(s)\,\mathrm{d}s \quad \text{for} \quad x \geq 0, \qquad \hat H(v) = H(|v|) \quad \text{for} \quad v \in X.$$

Both H and $\hat H$ are convex in their respective domains of definition. Hence,

$$\hat H(v) - \hat H(w) \leq \left\langle \nabla \hat H(v), v - w \right\rangle = \left\langle \frac{f^{-1}(|v|)}{|v|}v, v - w \right\rangle$$

for all $v, w \in X$. In particular, by the convexity argument, the first term on the left-hand side of (7.5) can be estimated from below as follows:

$$\left\langle \frac{f^{-1}(\varepsilon|\dot\xi_n^\varepsilon|)}{|\dot\xi_n^\varepsilon|}\dot\xi_n^\varepsilon, \dot\xi_n^\varepsilon - \dot\xi_n \right\rangle \geq \frac{1}{\varepsilon}\left(H(\varepsilon|\dot\xi_n^\varepsilon|) - H(\varepsilon|\dot\xi_n|)\right) \tag{7.6}$$

The last term on the right-hand side of (7.5) has to be estimated from above. By virtue of Theorem 2.6 and recalling (2.15) we have

$$f^{-1}(\varepsilon|\dot\xi_n^\varepsilon(t)|) = d(u(t) - \xi_n^\varepsilon(t)) \leq \frac{r}{3}. \tag{7.7}$$

Then

$$\frac{(1+\delta)|\dot\xi_n^\varepsilon|}{2r}\left(f^{-1}(\varepsilon|\dot\xi_n^\varepsilon|)\right)^2 \leq \frac{(1+\delta)|\dot\xi_n^\varepsilon|}{6}f^{-1}(\varepsilon|\dot\xi_n^\varepsilon|) = \frac{(1+\delta)}{6\varepsilon}\varepsilon|\dot\xi_n^\varepsilon|f^{-1}(\varepsilon|\dot\xi_n^\varepsilon|). \tag{7.8}$$

By Lemma 7.1 with $h = f^{-1}$, $x = y$, and $\sigma = 1$ we have $xf^{-1}(x) \leq 2H(x)$, hence

$$\frac{(1+\delta)}{6\varepsilon}\varepsilon|\dot{\xi}_n^\varepsilon| f^{-1}(\varepsilon|\dot{\xi}_n^\varepsilon|) \leq \frac{1+\delta}{3\varepsilon} H(\varepsilon|\dot{\xi}_n^\varepsilon|). \tag{7.9}$$

Using (7.7) again, similarly we derive the following estimate

$$\frac{(1+\delta)|\dot{\xi}_n|}{2r}\left(f^{-1}(\varepsilon|\dot{\xi}_n^\varepsilon|)\right)^2 \leq \frac{(1+\delta)}{6\varepsilon}\varepsilon|\dot{\xi}_n| f^{-1}(\varepsilon|\dot{\xi}_n^\varepsilon|). \tag{7.10}$$

We are now again in the situation of Lemma 7.1 with $h = f^{-1}$, $\sigma = \delta$, $x = \varepsilon|\dot{\xi}_n|$, and $y = \varepsilon|\dot{\xi}_n^\varepsilon|$, which together with (7.4) yields that

$$\frac{(1+\delta)}{6\varepsilon}\varepsilon|\dot{\xi}_n| f^{-1}(\varepsilon|\dot{\xi}_n^\varepsilon|) \leq \frac{\delta(1+\delta)}{6\varepsilon}\left(H(\varepsilon|\dot{\xi}_n^\varepsilon|) + H\left(\frac{\varepsilon|\dot{u}_n|}{\delta}\right)\right). \tag{7.11}$$

Choosing $\delta = 1$ we obtain

$$\frac{1+\delta}{3} + \frac{\delta(1+\delta)}{6} = 1,$$

hence, by combining (7.6) and (7.8)–(7.11) with the inequality (7.5) we get

$$\begin{aligned}
&\frac{1}{\varepsilon}\left(H(\varepsilon|\dot{\xi}_n^\varepsilon|) - H(\varepsilon|\dot{\xi}_n|)\right) + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\xi_n^\varepsilon - \xi_n\right|^2 \\
&\qquad \leq \frac{1}{r}(|\dot{\xi}_n^\varepsilon| + |\dot{\xi}_n|)\left|\xi_n^\varepsilon - \xi_n\right|^2 + \frac{1}{\varepsilon}H(\varepsilon|\dot{\xi}_n^\varepsilon|) + \frac{1}{3\varepsilon}H(\varepsilon|\dot{u}_n|)
\end{aligned}$$

which can be reduced using also (7.4) to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left|\xi_n^\varepsilon - \xi_n\right|^2 \leq \tilde{C}\left((|\dot{\xi}_n^\varepsilon| + |\dot{u}_n|)\left|\xi_n^\varepsilon - \xi_n\right|^2 + \frac{1}{\varepsilon}H\left(\varepsilon|\dot{u}_n|\right)\right) \tag{7.12}$$

with a constant $\tilde{C}$ independent of ε and n. This is an inequality of the form (4.12) with $\alpha(t) = \tilde{C}(|\dot{\xi^\varepsilon}_n(t)| + |\dot{u}_n(t)|)$, and $\gamma(t) = \frac{\tilde{C}}{\varepsilon}H\left(\varepsilon|\dot{u}_n(t)|\right)$. By construction $\xi_n^\varepsilon(0) = \xi_n(0)$, thus the Gronwall inequality (4.13) gives

$$\sup_{t\in[0,T]}\left|\xi_n^\varepsilon(t) - \xi_n(t)\right|^2 \leq C_n\frac{1}{\varepsilon}\int_0^T H\left(\varepsilon|\dot{u}_n(t)|\right)\,\mathrm{d}t, \tag{7.13}$$

where $C_n = \tilde{C}^2(V_0 + \mathrm{Var}_{[0,T]}\,u_n) \geq \tilde{C}\exp\left(\int_0^T \alpha(t)\,\mathrm{d}t\right)$, with constant $V_0 > 0$ from Theorem 2.8.

The assertion of Theorem 2.9 now follows from the triangle inequality. We have indeed

$$|\xi^\varepsilon - \xi|_{[0,T]} \leq |\xi^\varepsilon - \xi_n^\varepsilon|_{[0,T]} + |\xi_n^\varepsilon - \xi_n|_{[0,T]} + |\xi_n - \xi|_{[0,T]}.$$

Hence by Theorem 2.7 and (7.13) we obtain

$$|\xi^\varepsilon - \xi|_{[0,T]} \le \tilde{L}\Bigg(|\xi^\varepsilon(0) - \xi^\varepsilon_n(0)| + |u-u_n|^{1/2}_{[0,T]} + |u-u_n|_{[0,T]}\Bigg) + \left(\frac{C_n}{\varepsilon}\int_0^T H\left(\varepsilon|\dot{u}_n(t)|\right)\,\mathrm{d}t\right)^{1/2} + |\xi_n - \xi|_{[0,T]}. \tag{7.14}$$

Recalling the value of the initial conditions, note that

$$|\xi^\varepsilon(0) - \xi^\varepsilon_n(0)| \le |\xi^\varepsilon(0) - \xi(0)| + |\xi(0) - \xi^\varepsilon_n(0)| = |\xi^\varepsilon_0 - \xi_0| + |u(0) - u_n(0)|,$$

and consequently (7.14) becomes

$$|\xi^\varepsilon - \xi|_{[0,T]} \le |\xi_n - \xi|_{[0,T]} + \tilde{L}\Bigg(|u(0) - u_n(0)| + |u-u_n|^{1/2}_{[0,T]} + |u-u_n|_{[0,T]}\Bigg) + \tilde{L}|\xi^\varepsilon_0 - \xi_0| + \left(\frac{C_n}{\varepsilon}\int_0^T H\left(\varepsilon|\dot{u}_n(t)|\right)\,\mathrm{d}t\right)^{1/2} \tag{7.15}$$

Let $\kappa > 0$ be arbitrarily small. By [24, Theorem 4.5], the functions ξ_n converge uniformly to ξ. We thus can choose $n_0 \in \mathbb{N}$ such that

$$|\xi_n - \xi|_{[0,T]} + |u_n - u|^{1/2}_{[0,T]} + 2|u_n - u|_{[0,T]} < \frac{\kappa}{2(\tilde{L}+1)} \quad \text{for } \ n \ge n_0 \tag{7.16}$$

To estimate the integral term, we first notice that H is a convex function, hence, $H(s) \le s f^{-1}(s)$ for $s \ge 0$, so that

$$\frac{1}{\varepsilon}\int_0^T H\left(\varepsilon|\dot{u}_n(t)|\right)\,\mathrm{d}t \le \int_0^T |\dot{u}_n(t)| f^{-1}\left(\varepsilon|\dot{u}_n(t)|\right)\,\mathrm{d}t. \tag{7.17}$$

Denoting $\beta_1 := \mathrm{Var}_{[0,T]}\, u_{n_0}$ and $\beta_2 := \int_0^T |\dot{u}_{n_0}(t)|^2\,\mathrm{d}t$, we find $\gamma > 0$ such that

$$s \in [0,\gamma] \implies f^{-1}(s) \le \frac{\kappa^2}{32 C_{n_0}(\beta_1 + 1),}$$

and define the sets

$$A_\varepsilon(\gamma) = \left\{ t \in (0,T) : |\dot{u}_{n_0}(t)| > \frac{\gamma}{\varepsilon} \right\}, \quad B_\varepsilon(\gamma) = (0,T) \setminus A_\varepsilon(\gamma).$$

In view of this, we get

$$\int_{B_\varepsilon(\gamma)} |\dot{u}_{n_0}(t)| f^{-1}\left(\varepsilon|\dot{u}_{n_0}(t)|\right)\,\mathrm{d}t \le \frac{\kappa^2}{32 C_{n_0}(\beta_1+1)}\int_0^T |\dot{u}_{n_0}(t)|\,\mathrm{d}t \le \frac{\kappa^2}{32 C_{n_0}} \tag{7.18}$$

On the other hand, since f^{-1} is concave, there exists a constant $C^* > 0$ such that $f^{-1}(s) \leq C^*(1+s)$ for all $s \geq 0$. Therefore

$$\begin{aligned}\int_{A_\varepsilon(\gamma)} |\dot{u}_{n_0}(t)| f^{-1}\left(\varepsilon|\dot{u}_{n_0}(t)|\right) \mathrm{d}t &\leq C^* \int_{A_\varepsilon(\gamma)} |\dot{u}_{n_0}(t)| \, \mathrm{d}t + \varepsilon C^* \int_{A_\varepsilon(\gamma)} |\dot{u}_{n_0}(t)|^2 \, \mathrm{d}t \\ &\leq C^* \frac{\varepsilon}{\gamma} \int_{A_\varepsilon(\gamma)} |\dot{u}_{n_0}(t)|^2 \, \mathrm{d}t + \varepsilon C^* \beta_2\end{aligned}$$

that is,

$$\int_{A_\varepsilon(\gamma)} |\dot{u}_{n_0}(t)| f^{-1}\left(\varepsilon|\dot{u}_{n_0}(t)|\right) \mathrm{d}t \leq \varepsilon C^* \beta_2 \left(\frac{1}{\gamma} + 1\right). \tag{7.19}$$

From (7.17)–(7.18) we thus conclude that

$$\frac{C_{n_0}}{\varepsilon} \int_0^T H\left(\varepsilon|\dot{u}_{n_0}(t)|\right) \mathrm{d}t \leq C_{n_0} \varepsilon \beta_2 C^* \frac{(1+\gamma)}{\gamma} + \frac{\kappa^2}{32}$$

We can therefore choose $\varepsilon_1 > 0$ sufficiently small such that

$$\tilde{L}|\xi_0^\varepsilon - \xi_0| + \left(\frac{C_{n_0}}{\varepsilon} \int_0^T H\left(\varepsilon|\dot{u}_n(t)|\right) \mathrm{d}t\right)^{1/2} \leq \frac{\kappa}{2} \quad \text{for } \varepsilon \in (0, \varepsilon_1). \tag{7.20}$$

Using (7.15), (7.16), and (7.20) we see that $|\xi^\varepsilon - \xi|_{[0,T]} < \kappa$ for $\varepsilon \in (0, \varepsilon_1)$. The parameter κ can be chosen arbitrarily small, and Theorem 2.9 is thus proved. ■

Acknowledgments
The authors are thankful to the referees for their comments that helped to improve the manuscript. The research leading to this results have been supported by GAČR Grant No. 20-14736S and by the European Regional Development Fund, Project No. CZ.02.1.01/0.0/0.0/16_019/0000778. The Institute of Mathematics of the Czech Academy of Sciences is supported by RVO:67985840.

Data Availability
Not applicable.

Statements and declarations
The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

[1] S. Adly, F. Nacry, L. Thibault, Discontinuous sweeping process with prox-regular sets. *ESAIM: COCV* **23** (2017), 1293–1329.

[2] G. Aumann, *Reelle Funktionen.* Springer-Verlag, Berlin-Göttingen-Heidelberg, 1954.

[3] H. Benabdellah, Existence of solutions to the nonconvex sweeping process. *J. Differential Equations* **164** (2000), 286–295.

[4] F. Bernicot, J. Venel, Existence of solutions for second-order differential inclusions involving proximal normal cones. *J. Math. Pures Appl.* **98** (2012), 257–294.

[5] F. Bernicot, J. Venel, Sweeping process by prox-regular sets in Riemannian Hilbert manifolds. *J. Differential Equations* **259** (2015), 4086–4121.

[6] N. Bourbaki, *Functions of a Real Variable: Elementary Theory*. Springer-Verlag, Berlin, 2004.

[7] M. Brokate, P. Krejčí, Optimal control of ODE systems involving a rate independent variational inequality. *Discrete Cont. Dyn. Systems B* **18** (2013), 331–348.

[8] N. Chemetov, M. D. P. Monteiro Marques, Non-convex quasi-variational differential inclusions. *Set-Valued Anal.* **15** (2007), 209–221.

[9] F. H. Clarke, R. J. Stern, P. R. Wolenski, Proximal smoothness and the lower-C^2 property. *J. Convex Analysis* **2** (1995), 117–144.

[10] G. Colombo, V. V. Goncharov, The sweeping process without convexity. *Set-Valued Anal.* **7** (1999), 357–374.

[11] G. Colombo, C. Kozaily, Existence and uniqueness of solutions for an integral perturbation of Moreau's sweeping process. *J. Convex Anal.* **27** (1) (2020), 227–236.

[12] G. Colombo, M. D. P. Monteiro Marques, Sweeping by a continuous prox-regular set. *J. Differential Equations* **187** (2003), 46–62.

[13] M. R. de Pinho, M. M. A. Ferreira, G. V. Smirnov, *Optimal Control Involving Sweeping Processes Set-Valued and Variational Analysis*, 2019, **27**, 523–548.

[14] J. F. Edmond, L. Thibault, BV solutions of nonconvex sweeping process differential inclusions with perturbation. *J. Differential Equations* **226** (2006), 135–179.

[15] H. Federer, Curvature measures. *Trans. Amer. Math. Soc.* **93** (1959), 418–491.

[16] D. Fraňková, Regulated functions with values in Banach space. *Math. Bohemica* **144** (2019), 437–456.

[17] C. Gavioli, P. Krejčí, Control and controllability of PDEs with hysteresis. *Appl. Math. Optim.* **84** (1) (2021), 829–847.

[18] A. L. Gurson, Continuum theory of ductile rupture by void nucleation and growth: Part I-Yield criteria and flow rules for porous ductile media. *Journal of Engineering Materials and Technology* **99** (1977), 2–15.

[19] P. Hartman, *Ordinary Differential Equations*. John Wiley and Sons, Inc., Baltimore, 1973.

[20] A. Jourani, E. Vilches, Moreau-Yosida regularization of state-dependent sweeping processes with nonregular sets. *J. Optim. Theory Appl.* **173** (1) (2017), 91–116.

[21] B. Kaltenbacher, P. Krejčí: Analysis of an optimization problem for a piezoelectric energy harvester. *Arch. Appl. Mech.* **89** (2019), 1103–1122.

[22] P. Krejčí, M. Liero, Rate independent Kurzweil processes. *Appl. Math.* **54** (2009), 117–145.

[23] P. Krejčí, G. A. Monteiro, What is the best viscous approximation to a rate-independent process? *Journal of Convex Analysis* **27** (3) (2020), 1015–1032.

[24] P. Krejčí, G. A. Monteiro, V. Recupero, Non-convex sweeping processes in the space of regulated functions. *Communications on Pure and Applied Analysis* **21** (9) (2022), 2999–3029.

[25] M. Kunze, M. D. P. Monteiro Marques, Yosida-Moreau regularization of sweeping processes with unbounded variation. *J. Differential Equations* **130** (1996), 292–306.

[26] G. A. Monteiro, A. Slavík, M. Tvrdý, *Kurzweil–Stieltjes Integral: Theory and Applications.* World Scientific Publishing, New Jersey, 2019.

[27] J.-J. Moreau, Rafle par un convexe variable I. *Sém. d'Anal. Convexe*, Montpellier **1** (1971), Exposé No. 15.

[28] J.-J. Moreau, Evolution problem associated with a moving convex set in a Hilbert space. *J. Differential Equations* **26** (1977), 347–374.

[29] F. Nacry, L. Thibault, BV prox-regular sweeping process with bounded truncated variation. *Optimization*, 2020 Optimization **69** (7-8) (2020), 1391–1437.

[30] R. A. Poliquin, R. T. Rockafellar, L. Thibault, Local differentiability of distance functions. *Trans. Amer. Math. Soc.* **352** (2000), 5231–5249.

[31] L. Thibault, Sweeping process with regular and nonregular sets. *J. Differential Equations* **193** (2003), 1–26.

[32] H. C. Schnabel, Zur Wohlgestelltheit des Gurson-Modells. PhD Thesis, TU München, 2006.

[33] M. Valadier, Quelques problèmes d'entrainement unilatéral en dimension finie. *Sém. Anal. Convexe*, Montpellier (1988), Exposé No. 8.

[34] J.-P. Vial, Strong and weak convexity of sets and functions. *Math. Ops. Res.* **8** (1983), 231–259.