ESAIM: COCV 32 (2026) 40
https://doi.org/10.1051/cocv/2026022

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

DIRECTIONAL DIFFERENTIABILITY FOR SOLUTION OPERATORS OF SWEEPING PROCESSES WITH CONVEX POLYHEDRAL ADMISSIBLE SETS

Martin Brokate[1] and Constantin Christof[2,*]

Abstract. We study directional differentiability properties of solution operators of rate-independent evolution variational inequalities with full-dimensional convex polyhedral admissible sets. It is shown that, if the space of continuous functions of bounded variation is used as the domain of definition, then the most prototypical examples of such solution operators – the vector play and stop – are Hadamard directionally differentiable in a pointwise manner *if and only if* the admissible set is non-obtuse. We further prove that, in those cases where they exist, the directional derivatives of the vector play and stop are uniquely characterized by a system of projection identities and variational inequalities and that directional differentiability cannot be expected in the obtuse case even if the solution operator is restricted to the space of Lipschitz continuous functions. Our results can be used, for example, to formulate Bouligand stationarity conditions for optimal control problems involving sweeping processes.

Mathematics Subject Classification. 34C55, 47J40, 49J40, 49J52, 49K40, 90C31.

Received March 22, 2025. Accepted March 9, 2026.

1. Introduction

This paper is concerned with directional differentiability properties of solution maps $(u, x_0) \mapsto x$ of sweeping processes of the form

$$x(0) = x_0, \qquad x(t) \in Z - u(t) \ \forall t \in [0, T], \qquad \dot{x}(t) \in -\mathcal{N}_{Z-u(t)}(x(t)) \text{ for a.a. } t \in (0, T), \tag{1.1}$$

involving a full-dimensional convex polyhedron $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, an initial value $x_0 \in Z - u(0)$, a terminal time $T > 0$, a forcing term $u\colon [0, T] \to \mathbb{R}^d$, and the outward normal cone $\mathcal{N}_{Z-u(t)}(x(t))$ of the set $Z - u(t)$ at $x(t)$. We study the problem (1.1) in its evolution variational inequality (EVI) formulation

$$y(0) = y_0, \qquad y(t) \in Z \ \forall t \in [0, T], \qquad \langle \dot{y}(t) - \dot{u}(t), v - y(t)\rangle \geq 0 \ \forall v \in Z \text{ for a.a. } t \in (0, T), \tag{1.2}$$

Keywords and phrases: Hysteresis, evolution variational inequality, sweeping process, rate-independence, play operator, stop operator, directional differentiability, optimal control, Bouligand stationarity, strong stationarity.

[1] Technische Universität München, CIT, Department of Mathematics, Boltzmannstraße 3, 85748 Garching, Germany; Weierstrass Institute for Applied Analysis and Stochastics, Anton-Wilhelm-Amo-Straße 39, 10117 Berlin, Germany; Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 16629 Praha 6, Czech Republic.
[2] Technische Universität Darmstadt, Department of Mathematics, Dolivostraße 15, 64293 Darmstadt, Germany.

* Corresponding author: christof@mathematik.tu-darmstadt.de

that is obtained by introducing the variable $y := x + u$, by defining $y_0 := x_0 + u(0) \in Z$, and by rewriting the normal-cone inclusion in (1.1) in terms of the Euclidean scalar product $\langle \cdot, \cdot \rangle \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. More precisely, we consider the following weak formulation of (1.2) that is not only sensible for absolutely continuous inputs u but also for all u in the space $CBV([0,T];\mathbb{R}^d)$ of continuous functions of bounded variation with values in $\mathbb{R}^d$:

$$y \in CBV([0,T];\mathbb{R}^d), \qquad y(0) = y_0, \qquad y(t) \in Z \ \forall t \in [0,T], \qquad \int_0^T \langle v - y, \mathrm{d}(y-u)\rangle \geq 0 \ \forall v \in C([0,T];Z). \tag{V}$$

Here, $C([0,T];Z)$ denotes the set of all continuous functions $v \colon [0,T] \to \mathbb{R}^d$ with values in Z and the integral has to be understood in the sense of Kurzweil-Stieltjes (see Sect. 2.4). Recall that, in the context of the EVI-formulations (1.2) and (V), the solution mapping $\mathcal{S} \colon (u, y_0) \mapsto y$ is also referred to as the *stop operator*. Its twin, the function $\mathcal{P}(u, y_0) := u - \mathcal{S}(u, y_0)$, is called the *play operator*. The main result of this paper characterizes precisely for which convex polyhedral admissible sets Z the solution map $\mathcal{S} \colon (u, y_0) \mapsto y$ of (V) possesses a pointwise (or, more precisely, $BV([0,T];\mathbb{R}^d)$-weak-star) directional derivative.

1.1. Main result

The main result of this paper is the following theorem:

Theorem 1.1. *Let $T > 0$ be given and let $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a full-dimensional convex polyhedron. Then the variational inequality (V) possesses a unique solution $y \in CBV([0,T];\mathbb{R}^d)$ for all $(u, y_0) \in CBV([0,T];\mathbb{R}^d) \times Z$. The associated solution operator $\mathcal{S} \colon CBV([0,T];\mathbb{R}^d) \times Z \to CBV([0,T];\mathbb{R}^d)$, $(u, y_0) \mapsto y$, is globally Lipschitz continuous in the sense that there exists a constant $L > 0$ satisfying*

$$\|\mathcal{S}(u,y_0) - \mathcal{S}(\tilde{u},\tilde{y}_0)\|_{CBV([0,T];\mathbb{R}^d)} \leq L\left(\|u - \tilde{u}\|_{CBV([0,T];\mathbb{R}^d)} + |y_0 - \tilde{y}_0|\right) \ \forall (u,y_0), (\tilde{u},\tilde{y}_0) \in CBV([0,T];\mathbb{R}^d) \times Z. \tag{1.3}$$

Further, the following statements hold regarding the directional differentiability properties of $\mathcal{S}$.

(I) If Z is non-obtuse (see Def. 2.12), then the map $\mathcal{S}$ is directionally differentiable in the sense that, for all $(u, y_0) \in CBV([0,T];\mathbb{R}^d) \times Z$ and all $(h, h_0) \in CBV([0,T];\mathbb{R}^d) \times \mathbb{R}^d$ satisfying $y_0 + \tau_0 h_0 \in Z$ for some $\tau_0 > 0$, there exists a unique $\delta := \mathcal{S}'((u,y_0);(h,h_0)) \in BV([0,T];\mathbb{R}^d)$ satisfying

$$\lim_{(0,\tau_0) \ni \tau \to 0} \frac{\mathcal{S}(u + \tau h, y_0 + \tau h_0)(t) - \mathcal{S}(u,y_0)(t)}{\tau} = \delta(t) \qquad \forall t \in [0,T]. \tag{1.4}$$

This directional derivative δ of $\mathcal{S}$ at (u, y_0) in direction (h, h_0) is uniquely characterized by the system

$$\begin{aligned}
&\delta \in BV([0,T];\mathbb{R}^d), \qquad \delta_+ \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y,u), \qquad \delta(0) = h_0,\\
&\delta(t) = \pi_{\mathcal{Z}(y(t))}(\delta(t-)) \ \forall t \in [0,T], \qquad \delta(t+) = \pi_{V^+(t)}(\delta(t)) \ \forall t \in [0,T],\\
&\int_{s_1}^{s_2} \langle z, \mathrm{d}(\delta - h)\rangle + \int_{s_1}^{s_2} \langle \delta, \mathrm{d}h\rangle - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 \geq 0 \quad \forall 0 \leq s_1 < s_2 \leq T \quad \forall z \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y,u).
\end{aligned} \tag{1.5}$$

Here, δ_+ denotes the right-limit function of δ; y is defined by $y := \mathcal{S}(u, y_0)$; $\mathcal{K}^{\mathrm{crit}}_{G_r}(y,u)$ is given by

$$\mathcal{K}^{\mathrm{crit}}_{G_r}(y,u) := \left\{ z \in G_r([0,T];\mathbb{R}^d) \;\middle|\; z(t) \in \mathcal{Z}(y(t)) \ \forall t \in [0,T] \text{ and } \int_0^T \langle z, \mathrm{d}(y-u)\rangle = 0 \right\};$$

the symbol $G_r([0,T];\mathbb{R}^d)$ denotes the space of right-continuous regulated functions $v \colon [0,T] \to \mathbb{R}^d$; $\mathcal{Z}(x)$ is the tangent cone of Z at $x \in Z$; $\delta(t-)$ and $\delta(t+)$ denote a left and right limit, respectively; π denotes a Euclidean projection; and $V^+(t)$ is a suitably defined time-dependent subspace (see Def. 3.10).

(II) If Z is not non-obtuse, then the map $\mathcal{S}$ is not directionally differentiable in the pointwise sense of (I), not even if its domain of definition $CBV([0,T];\mathbb{R}^d) \times Z$ is replaced by $W^{1,\infty}((0,T);\mathbb{R}^d) \times Z$.

For details on the unique solvability part of Theorem 1.1 and the Lipschitz estimate (1.3), we refer the reader to Proposition 2.33 and Theorem 2.37. (These two results rely heavily on the preliminary work of [1–4].) The proofs of the statements (I) and (II), which are the main contributions of this paper, can be found in Sections 3 and 4; see in particular Theorems 3.32 and 4.1. Before we discuss the role and background of Theorem 1.1 in more detail, we give some remarks on its mathematical scope.

Remark 1.2.

- The assumption that Z is full-dimensional (*i.e.*, has a nonempty interior) can be made without loss of generality (w.l.o.g.) in Theorem 1.1. If Z is nonempty, then full-dimensionality can always be achieved by reformulating (V) as an EVI in the affine hull of Z. For the same reason, we may assume w.l.o.g. that the underlying vector space is $\mathbb{R}^d$. Indeed, if we start with a problem in an arbitrary Hilbert space with a polyhedral admissible set Z, then we can always reformulate (V) as a problem in the finite-dimensional linear hull of the normal vectors of Z and consider this space instead; see [4], Remark 4.5.
- Due to the Lipschitz continuity of the function $\mathcal{S}$, the difference quotients in (I) also converge weakly-star in $BV([0,T];\mathbb{R}^d)$. In particular, the map $\mathcal{S}$ is directionally differentiable as a function from $CBV([0,T];\mathbb{R}^d) \times Z$ to $L^p((0,T);\mathbb{R}^d)$ for all $1 \leq p < \infty$ in (I). The Lipschitz continuity of $\mathcal{S}$ also implies that $\mathcal{S}$ is pointwisely Hadamard directionally differentiable in (I); see [5], Proposition 2.49.
- Due to the evolutionary nature of (V), the nonexistence of a pointwise directional derivative in (II) implies that $\mathcal{S}$ cannot be directionally differentiable as a map from $W^{1,\infty}((0,T);\mathbb{R}^d) \times Z$ to $L^1((0,T);\mathbb{R}^d)$ (or, more generally, from $W^{1,\infty}((0,T);\mathbb{R}^d) \times Z$ to any sensible function space); see Theorem 4.1. Note that it could still be the case here that $\mathcal{S}$ is directionally differentiable when restricted to a suitable subset of $W^{1,\infty}((0,T);\mathbb{R}^d) \times Z$, *e.g.*, to $\{u \in W^{1,\infty}((0,T);\mathbb{R}^d) \mid \dot{u} \in BV([0,T];\mathbb{R}^d)\} \times Z$; *cf.* Section 4.

1.2. Background and relations to prior work

Sweeping processes of the type (1.1) have first been investigated by Moreau in the nineteen-seventies to study problems arising in elastoplasticity; see [6–8]. On the most fundamental level, they can be understood as models for the dynamical behavior of an infinitesimally small body that is contained in a hollow shape which moves through the Euclidean space $\mathbb{R}^d$ and which forces the body to move along when it comes into contact with the shape's boundary (similarly to a coin on a table under an upside-down drinking glass). In this interpretation, the function $x\colon [0,T] \to \mathbb{R}^d$ in (1.1) describes the position of the body, Z is the shape, $u\colon [0,T] \to \mathbb{R}^d$ models the motion of Z, the condition $x(t) \in Z - u(t)$ reflects that the body cannot leave the set $Z - u(t)$, and the inclusion $\dot{x}(t) \in -\mathcal{N}_{Z-u(t)}(x(t))$ expresses that the body is "swept" into an inward normal direction when in contact with the boundary $\partial(Z - u(t))$. Since their initial introduction, sweeping processes have been used to model phenomena in various application areas (mechanics [3, 9], electrical engineering [10], and crowd control [11] to name a few). They moreover often serve as prototypical examples for the class of energetic rate-independent systems [12]. Due to this prominent role, problems of the type (1.1) have been subject to extensive research during the last fifty years; see, *e.g.*, [1–4, 8, 9, 12–14] and the references therein.

Despite the considerable attention that sweeping processes have received in the past, the question of whether their solution mappings are directionally differentiable or not is still severely underinvestigated. At least to the best of our knowledge, the only contributions on this topic currently found in the literature are [15], [16], and [17], which all deal with the EVI (1.2) in the case $d = 1$ and $Z = [-r, r]$, $r > 0$. In the first of these papers, [15], it is shown that the scalar stop operator $\mathcal{S}\colon (u, y_0) \mapsto y$ possesses a pointwise Hadamard directional derivative when considered as a function from $C([0,T]) \times Z$ to $C([0,T])$; see [15], Proposition 5.3, Corollary 5.4. The proof of this result relies heavily on the fact that, in the one-dimensional setting, it is possible to derive a semi-explicit solution formula for the play operator $\mathcal{P}$ in terms of the cumulated maximum function; see [15], Lemma 5.1. Based on the same observation, Newton- and Bouligand differentiability properties of the scalar stop and play are investigated in [16]. The third paper, [17], finally establishes an auxiliary variational inequality that uniquely characterizes the derivatives of the scalar play and stop into directions $(h, h_0) \in CBV([0,T]) \times \mathbb{R}$; see [17], Theorem 2.1. At least to the best of our knowledge, directional differentiability results for vectorial sweeping processes like (1.2) can currently not be found in the literature. One of the reasons for the absence of such results is that, although simple at first glance, the structure of (1.2) makes establishing the (directional) differentiability of the mappings $\mathcal{S}$ and $\mathcal{P}$ rather difficult. The discontinuity of the normal-cone mapping in (1.1), for example, causes pointwise limits of difference quotients of $\mathcal{S}$ and $\mathcal{P}$ to be typically neither left- nor right-continuous and, thus, creates serious regularity problems; see Sections 3 and 4 and Example 3.36.

It should be noted that this lack of directional differentiability results for solution operators is not limited to sweeping processes but also extends to the larger class of "true" evolution variational inequalities (where with "true" we mean that the EVI cannot be recast as a PDE by means of a viscous regularization or similar tricks; *cf.* [18]). Even in this wider field, we are only aware of two additional contributions on the directional differentiability properties of solution maps. The first one, [19], is concerned with the directional differentiability of the solution operator of the scalar parabolic Signorini problem. The main idea of this paper is to interpret the EVI at hand as a problem in a fractional Sobolev space and to then employ an argumentation that has originally been developed in [20, 21] for elliptic obstacle problems. The second paper, [22], is concerned with general parabolic obstacle-type evolution variational inequalities. It establishes by means of pointwise-a.e. convexity properties that solution operators of this type of EVI are directionally differentiable as functions into the Lebesgue spaces and demonstrates that the arguments of [19] fail for all obstacle-type EVIs that involve pointwise constraints on a set of positive measure; see [22], Section 3, Theorem 4.1.

That so little is currently known about the differential sensitivity analysis of solution maps of EVIs is rather unsatisfying – in particular since, for elliptic variational inequalities, differentiability results have been established in a multitude of papers and the subject is nowadays almost completely understood. Compare, *e.g.*, with the seminal works [20, 21, 23] on this topic, with [24–28], and with the results in [29], Theorem 4.3, [30], Theorem 1.4.1, and [31], Theorem 4.1, which characterize precisely by means of the concept of second-order epi-differentiability in which situations the solution map of an elliptic variational inequality of the first or the second kind is directionally differentiable and in which situations it is not.

We would like to emphasize at this point that the question of whether directional differentiability results can be derived for problems like (1.2) is far from academic. Results on this topic are of fundamental importance when optimal control problems governed by systems involving the play and stop operator are considered (*e.g.*, PDE-systems involving hysteresis) as they form the basis for the formulation of first-order necessary optimality conditions and the design of derivative-based solution algorithms. In the field of optimal control of sweeping processes, the lack of knowledge about the differentiability properties of solution maps has caused essentially all authors concerned with the matter to resort to approximation or regularization approaches that replace the original problem formulation with a sequence of discretized and/or regularized problems. Compare, *e.g.*, with [13, 32–37] in this context, which employ discrete approximations, and with [38–43], which utilize various mollification techniques. Again, this state of affairs is rather unsatisfactory, in particular when compared with the field of optimal control of elliptic variational inequalities for which the available directional differentiability results allow to derive so-called "strong stationarity"-systems that are equivalent to the notion of Bouligand stationarity (*i.e.*, the condition that the directional derivatives of the reduced objective function are nonnegative in all admissible directions) and thus do not suffer from the loss of information that regularization and discretization approaches are typically subject to; see [44–47]. That it is possible to obtain stationarity systems of the same strength for evolution variational inequalities once directional differentiability results are available has recently been demonstrated in [48] for the one-dimensional play and stop based on the findings of [15–17]; see [48], Theorem 7.1 and Section 5. To our best knowledge, the result in [48] is the first on strong stationarity conditions in the field of optimal control of true EVIs; *cf.* the comments in [48], Section 1.1. In particular, strong stationarity conditions for problems governed by sweeping processes with vector-valued in- and output functions seem to be unknown at the moment.

The principal aim of the present paper is to shed some further light on the directional differentiability properties of solution maps of sweeping processes and evolution variational inequalities in general. By doing so, we also hope to lay the foundation for the study of stationarity conditions for optimal control problems governed by EVIs that do not rely on an approximation or regularization procedure and go beyond the one-dimensional setting of [48]. Theorem 1.1 serves these two goals because it answers conclusively the question under which conditions the vector play and stop are directionally differentiable when the admissible set Z is a convex polyhedron and the input space is chosen as $CBV([0,T];\mathbb{R}^d)$. At least to the best of our knowledge, Theorem 1.1 is the first result on the directional differentiability properties of solution maps of vectorial sweeping processes (and, more generally, vector-valued EVIs). It is also of broader interest because it highlights some quite peculiar effects that emerge in the differential sensitivity analysis of evolution variational inequalities but are absent in the study of elliptic problems. Some of the main takeaways are the following points.

- Theorem 1.1 shows that, even when we restrict the attention to polyhedra, it is very well possible that some admissible sets behave completely differently than others in an EVI like (1.2). In particular, whether the solution mapping of an EVI is directionally differentiable or not may hinge on geometric properties – in the situation of Theorem 1.1 the interior angles of Z – that are completely irrelevant for questions of well-posedness and Lipschitz stability; *cf.* Proposition 2.33 and Theorem 2.37.
- Theorem 1.1 demonstrates that solution maps of EVIs may fail to be directionally differentiable even if the admissible set and the considered input space are very well behaved. In particular, directional non-differentiability may also occur when the solution mapping of an EVI like (1.2) is considered on the space $H^1((0,T);\mathbb{R}^d)$, which is, to our knowledge, the most popular choice for the control space in the field of optimal control of rate-independent EVIs; see [41], Section 4, [48], Section 5.
- The system (1.5) and the fact that the directional derivatives δ in case (I) of Theorem 1.1 can be discontinuous from both the left and the right at some times $t \in (0,T)$ (see Ex. 3.36) indicate that, for the evolution variational inequality (1.2), it is in general not possible to write down a "classical" auxiliary EVI which characterizes the directional derivatives of the solution map $\mathcal{S}$. A unique characterization of the directional derivatives is apparently only possible in (I) by means of the combination of the "very weak" variational inequality in (1.5) and the separate jump conditions $\delta(t) = \pi_{\mathcal{Z}(y(t))}(\delta(t-))$ and $\delta(t+) = \pi_{V^+(t)}(\delta(t))$. This is a major difference to both the results for the one-dimensional play and stop in [17] (for which the directional derivatives of the solution map satisfy $\delta(t) \in \{\delta(t-), \delta(t+)\}$ for all $t \in [0,T]$) and elliptic variational inequalities; see [30], Theorem 1.4.1 and [31], Theorem 4.1.

We would like to point out that Theorem 1.1 contains the directional differentiability and characterization results of [15, 17] as a special case (modulo some minor differences in the regularity assumptions). For more details on this topic, see Section 3.7. As the argumentation that we use for the derivation of Theorem 1.1 is quite different from that in [15, 17], this also means that our analysis provides alternative proofs of the theorems in these two papers. Regarding the distinction between the cases (I) and (II) in Theorem 1.1, it should be noted that non-obtuse polyhedra also play a special role in the context of the question of whether different techniques for extending the solution mapping $\mathcal{S}\colon (u, y_0) \mapsto y$ of (1.2) to the set $BV([0,T];\mathbb{R}^d) \times Z$ give rise to the same operator; see [49]. If this is just a coincidence or has some deeper mathematical meaning is an open question.

1.3. Structure of the remainder of the paper

We conclude this introduction with an overview of the content and structure of the remainder of the paper. Section 2 is concerned with preliminaries. Here, we introduce basic notation and recall results from the theory of metric projections, convex polyhedra, Kurzweil-Stieltjes integration, and sweeping processes that are needed for the derivation of Theorem 1.1. This section also contains the proofs of the unique solvability of (V) and the Lipschitz estimate (1.3); see Proposition 2.33 and Theorem 2.37. As we aim to keep our analysis as self-contained as possible, we are very thorough in our discussion of preliminaries. Section 3 contains the proof of statement (I) of Theorem 1.1. The basic idea of the analysis in this section is to consider for a fixed choice of the tuple (u, y_0) and the direction (h, h_0) the set $\mathcal{D}$ of all $BV([0,T];\mathbb{R}^d)$-weak-star limits of difference quotients of the mapping $\mathcal{S}$ and to then derive in various steps more and more information about the elements of $\mathcal{D}$ until it becomes clear that this set is a singleton – similarly to the approaches for elliptic problems in [30, 31]. In combination with (1.3) and the Helly selection theorem, the directional differentiability of $\mathcal{S}$ then follows from a simple contradiction argument. A main difficulty in this argumentation (and also a major difference to the elliptic setting) is that, in the case of the EVI (V), one has to identify precisely how the elements of the set $\mathcal{D}$ jump at their points of discontinuity. We solve this problem by exploiting that EVIs (V) with a non-obtuse convex polyhedral admissible set Z can be studied by means of a vector-valued variant of Stampacchia's lemma; *cf.* [50], Theorem 5.8.2 and [51], Lemma 6.3. For a more detailed overview of the proof of statement (I), see Section 3.1. At the end of Section 3, we also demonstrate that part (I) of Theorem 1.1 indeed implies the results on the existence and characterization of directional derivatives of the scalar play and stop in [15, 17]. (Note that, in the case $d = 1$, all convex polyhedra $Z \subset \mathbb{R}^d$ are non-obtuse.) In Section 4, we prove statement (II) of Theorem 1.1 by explicitly constructing – for each convex polyhedron that is not non-obtuse – a situation in which the difference quotients of $\mathcal{S}$ diverge. Section 5 of the paper finally contains some remarks on how Theorem 1.1 can be used in the field of optimal control of sweeping processes. Here, we demonstrate how Bouligand stationarity conditions for optimization problems governed by EVIs can be established based on the directional differentiability results in Theorem 1.1 and sketch some ideas on how the system (1.5) and the analysis of Section 3 can be used to derive strong stationarity conditions for this problem class.

2. Preliminaries and preparatory results

This section contains preliminary results that are needed for the proof of Theorem 1.1. In Sections 2.1 and 2.2, we begin with basic notation and some results on metric projections. Section 2.3 is then concerned with preliminaries on convex polyhedra and their description. The results on non-obtuse and not non-obtuse polyhedra in Section 2.3 may also be of independent interest – in particular as the property of non-obtuseness is also crucial in areas other than the differential sensitivity analysis of EVIs; see [49]. In Section 2.4, we collect preliminaries on the theory of Kurzweil-Stieltjes integration. For a detailed introduction to this type of integral, we refer to [52]. Section 2.5 finally contains some preliminaries on the vector play and stop. Here, we in particular address the unique solvability part of Theorem 1.1 and the Lipschitz estimate (1.3).

2.1. Basic notation

In this paper, we write $\mathbb{R}$, $\mathbb{N}$, and $\mathbb{N}_0$ for the real numbers, the natural numbers, and the nonnegative integers, respectively. We further denote norms and inner products defined on a real vector space X by $\|\cdot\|_X$ and $\langle\cdot,\cdot\rangle_X$, respectively. For the topological dual space of a normed space $(X, \|\cdot\|_X)$, we use the symbol X^*, and for a dual pairing, the brackets $\langle\cdot,\cdot\rangle_{X^*,X}$. A closed ball of radius $r > 0$ in a normed space $(X, \|\cdot\|_X)$ centered at $x \in X$ is denoted by $B_r^X(x)$. In the special case $X = \mathbb{R}^d$, $d \in \mathbb{N}$, we write $|\cdot|$ for the Euclidean norm, $\langle\cdot,\cdot\rangle$ for the Euclidean scalar product, $B_r(x)$ for the ball $B_r^{\mathbb{R}^d}(x)$ defined with respect to (w.r.t.) $|\cdot|$, and $S^{d-1} := \{x \in \mathbb{R}^d \mid |x| = 1\}$ for the unit sphere. If D is a subset of a normed space, then we denote by ∂D the boundary of D, by $\mathrm{int}(D)$ the interior of D, by $\mathrm{cl}(D)$ the closure of D, by $\mathrm{conv}(D)$ the convex hull of D, and by $\mathrm{span}(D)$ the linear hull of D. For the empty set, we use the convention $\mathrm{span}(\emptyset) := \{0\}$. Analogously, we also interpret sums over empty index sets as zero. Recall that a set $D \subset X$ is called a cone if $sx \in D$ holds for all $x \in D$ and all $s \in (0,\infty)$. For a cone D in a normed space $(X, \|\cdot\|_X)$, we define $D^\circ := \{z^* \in X^* \mid \langle z^*, z\rangle_{X^*,X} \le 0 \ \forall z \in D\}$ to be the polar cone of D. In the case $X = \mathbb{R}^d$, we always identify X^* with X in this definition, *i.e.*, we use the convention $D^\circ := \left\{z^* \in \mathbb{R}^d \mid \langle z^*, z\rangle \le 0 \ \forall z \in D\right\}$. If D is a closed convex nonempty subset of a normed space $(X, \|\cdot\|_X)$ and $x \in D$, then we denote by $K_{\mathrm{rad}}(x; D) := \{z \in X \mid z = s(y - x), s > 0, y \in D\}$ the radial cone of D at x, by $K_{\tan}(x; D) := \mathrm{cl}(K_{\mathrm{rad}}(x; D))$ the tangent cone of D at x, and by $\mathcal{N}_D(x) := K_{\tan}(x; D)^\circ$ the (outward) normal cone of D at x. If V is a closed subspace of a Hilbert space, then we write $V^\perp$ for the orthogonal complement of V. For the cardinality of a set D, we use the symbol $|D|$. The modes of strong and weak convergence in a normed space are denoted by $\to$ and $\rightharpoonup$, respectively. If $F\colon X \to Y$ is a mapping and $\emptyset \ne D \subset X$, then we write $F|_D$ for the restriction of F to D and $\mathbb{1}_D\colon X \to \{0,1\}$ for the $\{0,1\}$-indicator function of D. For set-valued mappings, we use the notation $F\colon X \rightrightarrows Y$. Directional, Gâteaux, and Fréchet derivatives are denoted by a prime in this paper. If F is multivariate, then the symbol ∂_k indicates that we take the partial Fréchet derivative w.r.t. the k-th component of the argument of F. For weak/strong derivatives that are taken w.r.t. a time-variable, we also use a dot instead of a prime. Gradients are denoted by ∇.

If $-\infty < a < b < \infty$ and $d \in \mathbb{N}$ are given, then we use the standard symbols $L^p((a,b);\mathbb{R}^d)$, $H^1((a,b);\mathbb{R}^d)$, and $W^{1,p}((a,b);\mathbb{R}^d)$, $1 \leq p \leq \infty$, for the vectorial Lebesgue and first-order Sobolev spaces, respectively, endowed with their canonical norms, *i.e.*,

$$\|v\|_{L^p((a,b);\mathbb{R}^d)} := \left(\int_a^b |v(t)|^p \mathrm{d}t\right)^{1/p} \qquad \forall 1 \leq p < \infty, \qquad \|v\|_{L^\infty((a,b);\mathbb{R}^d)} := \operatorname*{ess\,sup}_{t\in[a,b]} |v(t)|, \qquad \textit{etc.}$$

The space of smooth, real-valued functions on $\mathbb{R}$ is denoted by $C^\infty(\mathbb{R})$ and the space of all functions $v \in C^\infty(\mathbb{R})$ that possess a compact support $\operatorname{supp}(v)$ by $C_c^\infty(\mathbb{R})$. We further write $C^\infty([a,b];\mathbb{R}^d)$ for the space of all functions $v\colon [a,b] \to \mathbb{R}^d$ whose components can be written as restrictions of elements of $C^\infty(\mathbb{R})$ to the interval $[a,b]$. For the space of all regulated functions $v\colon [a,b] \to \mathbb{R}^d$, we use the notation $G([a,b];\mathbb{R}^d)$. Recall that a function is called regulated if it is the uniform limit of a sequence of step functions or, equivalently, if it possesses left and right limits at all points $t \in [a,b]$; see [52], Definition 4.1.1, Theorem 4.1.5. The space $G([a,b];\mathbb{R}^d)$ is a Banach space when endowed with the supremum norm $\|v\|_\infty := \|v\|_{G([a,b];\mathbb{R}^d)} := \sup_{t\in[a,b]} |v(t)|$; see [52], Theorem 4.2.1. When we want to emphasize which interval we refer to, we also write $\|\cdot\|_{\infty,[a,b]}$ instead of $\|\cdot\|_\infty$. For the left and right limit of a function $v \in G([a,b];\mathbb{R}^d)$ at $t \in [a,b]$, we use the notation $v(t-)$ and $v(t+)$, respectively, with the usual conventions at the interval endpoints a and b, *i.e.*,

$$v(t-) := \lim_{(a,t)\ni s\to t} v(s) \;\; \forall t \in (a,b], \qquad v(a-) := v(a), \qquad v(t+) := \lim_{(t,b)\ni s\to t} v(s) \;\; \forall t \in [a,b), \qquad v(b+) := v(b).$$

We further define $v_\pm(t) := v(t\pm)$ for all $t \in [a,b]$. If the underlying interval is important, then we highlight this by writing $v_{-,[a,b]}$ and $v_{+,[a,b]}$. Note that, for $a < s_1 < s_2 < b$, one typically has $v_{-,[s_1,s_2]} \neq (v_{-,[a,b]})|_{[s_1,s_2]}$ and $v_{+,[s_1,s_2]} \neq (v_{+,[a,b]})|_{[s_1,s_2]}$ due to the different conventions at the interval endpoints. Recall that the set of points of discontinuity $[a,b] \setminus \{t \in [a,b] \mid v(t) = v(t+) = v(t-)\}$ of a regulated function $v\colon [a,b] \to \mathbb{R}^d$ is at most countable; see [52], Theorem 4.1.8. For every $v \in G([a,b];\mathbb{R}^d)$, we further have $v_-, v_+ \in G([a,b];\mathbb{R}^d)$ and

$$\begin{aligned} &v_-(t-) = v(t-) \;\forall t \in [a,b], \qquad && v_-(t+) = v(t+) \;\forall t \in [a,b), \\ &v_+(t-) = v(t-) \;\forall t \in (a,b], \qquad && v_+(t+) = v(t+) \;\forall t \in [a,b]; \end{aligned} \tag{2.1}$$

see [52], Corollary 4.1.9. We define $G_{rl}([a,b];\mathbb{R}^d) := \{v \in G([a,b];\mathbb{R}^d) \mid v(t) \in \{v(t-), v(t+)\} \;\forall t \in [a,b]\}$, $G_r([a,b];\mathbb{R}^d) := \{v \in G([a,b];\mathbb{R}^d) \mid v = v_+\}$, and $G_l([a,b];\mathbb{R}^d) := \{v \in G([a,b];\mathbb{R}^d) \mid v = v_-\}$. Note that $G_r([a,b];\mathbb{R}^d)$ and $G_l([a,b];\mathbb{R}^d)$ are closed subspaces of $G([a,b];\mathbb{R}^d)$ and, thus, Banach. The same is true for the space $C([a,b];\mathbb{R}^d)$ of continuous functions $v\colon [a,b] \to \mathbb{R}^d$. For arbitrary $v\colon [a,b] \to \mathbb{R}^d$, we denote by

$$\operatorname{var}(v;[a,b]) := \sup_{a=t_0\leq t_1\leq \ldots\leq t_N=b, N\in\mathbb{N}} \sum_{j=1,\ldots,N} |v(t_j) - v(t_{j-1})| \in [0,\infty]$$

the variation of v on $[a,b]$. Here, the supremum is taken over all partitions $a = t_0 \leq \ldots \leq t_N = b$, $N \in \mathbb{N}$, of the interval $[a,b]$. Note that this definition implies $\operatorname{var}(v;[a,b]) = \|\dot v\|_{L^1((a,b);\mathbb{R}^d)}$ for all $v \in W^{1,1}((a,b);\mathbb{R}^d)$. The space of all functions $v\colon [a,b] \to \mathbb{R}^d$ with finite variation is denoted by $BV([a,b];\mathbb{R}^d)$. Recall that $BV([a,b];\mathbb{R}^d)$ is a Banach space when equipped with the norm $\|v\|_{BV([a,b];\mathbb{R}^d)} := |v(a)| + \operatorname{var}(v;[a,b])$ and a subspace of $G([a,b];\mathbb{R}^d)$; see [52], Theorem 2.2.2. Further, it holds $v_+, v_- \in BV([a,b];\mathbb{R}^d)$ for all $v \in BV([a,b];\mathbb{R}^d)$. We define $BV_{rl}([a,b];\mathbb{R}^d) := BV([a,b];\mathbb{R}^d) \cap G_{rl}([a,b];\mathbb{R}^d)$, $BV_r([a,b];\mathbb{R}^d) := BV([a,b];\mathbb{R}^d) \cap G_r([a,b];\mathbb{R}^d)$, $BV_l([a,b];\mathbb{R}^d) := BV([a,b];\mathbb{R}^d) \cap G_l([a,b];\mathbb{R}^d)$, and $CBV([a,b];\mathbb{R}^d) := C([a,b];\mathbb{R}^d) \cap BV([a,b];\mathbb{R}^d)$. Note that the latter three of these sets are closed subspaces of $BV([a,b];\mathbb{R}^d)$ and, thus, Banach spaces. In what follows, we endow the space $CBV([a,b];\mathbb{R}^d)$ with the norm $\|v\|_{CBV([a,b];\mathbb{R}^d)} := \|v\|_{C([a,b];\mathbb{R}^d)} + \mathrm{var}(v;[a,b]) = \|v\|_\infty + \mathrm{var}(v;[a,b])$, which is equivalent to $\|\cdot\|_{BV([a,b];\mathbb{R}^d)}$. If $K\colon [a,b] \rightrightarrows \mathbb{R}^d$ is a set-valued function, then we define $G([a,b];K) := \{v \in G([a,b];\mathbb{R}^d) \mid v(t) \in K(t) \text{ for all } t \in [a,b]\}$. Analogous shorthand notations are also used for $C([a,b];\mathbb{R}^d)$, $BV([a,b];\mathbb{R}^d)$, *etc.* Sets $K \subset \mathbb{R}^d$ are interpreted as constant functions in this context. If $d = 1$ holds, then we simply write $G([a,b])$, $L^p(a,b)$, *etc.* instead of $G([a,b];\mathbb{R}^d)$, $L^p((a,b);\mathbb{R}^d)$, *etc.*

2.2. Preliminaries on projections

In what follows, we recall basic results on Euclidean projections and distance functions that are needed for the derivation of Theorem 1.1.

Definition 2.1 (projection and distance function)**.** Let $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a nonempty, convex, and closed set.

i) We denote by π_Z the Euclidean projection onto Z, *i.e.*,

$$\pi_Z\colon \mathbb{R}^d \to \mathbb{R}^d, \qquad \pi_Z(x) := \arg\min \left\{ \frac{1}{2}|v - x|^2 \;\middle|\; v \in Z \right\}. \tag{2.2}$$

ii) We denote by dist_Z the Euclidean distance function to Z, *i.e.*,

$$\mathrm{dist}_Z\colon \mathbb{R}^d \to [0,\infty), \qquad \mathrm{dist}_Z(x) := |\pi_Z(x) - x| = \min_{v \in Z} |v - x|.$$

Proposition 2.2 (properties of π_Z and dist_Z)**.** *Let $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be nonempty, convex, and closed.*

i) For every $x \in \mathbb{R}^d$, the projection $\pi_Z(x)$ is uniquely characterized by the variational inequality

$$\pi_Z(x) \in Z, \qquad \langle \pi_Z(x) - x, v - \pi_Z(x)\rangle \geq 0 \quad \forall v \in Z. \tag{2.3}$$

ii) The projection $\pi_Z\colon \mathbb{R}^d \to \mathbb{R}^d$ is one-Lipschitz, i.e.,

$$|\pi_Z(x_1) - \pi_Z(x_2)| \leq |x_1 - x_2| \qquad \forall x_1, x_2 \in \mathbb{R}^d. \tag{2.4}$$

iii) The function $F\colon \mathbb{R}^d \to \mathbb{R}$, $x \mapsto \mathrm{dist}_Z(x)^2$, is continuously Fréchet differentiable. Its gradient $\nabla F(x)$ w.r.t. the Euclidean scalar product at $x \in \mathbb{R}^d$ is given by

$$\nabla F(x) = 2(x - \pi_Z(x)). \tag{2.5}$$

iv) If Z is additionally a cone, then it holds

$$\begin{gathered} x = \pi_Z(x) + \pi_{Z^\circ}(x) \quad \text{and} \quad \langle \pi_Z(x), \pi_{Z^\circ}(x)\rangle = 0 \quad \forall x \in \mathbb{R}^d, \\ \pi_Z(sx) = s\pi_Z(x) \quad \forall x \in \mathbb{R}^d \quad \forall s \in [0,\infty), \end{gathered} \tag{2.6}$$

and the function $F\colon \mathbb{R}^d \to \mathbb{R}$, $x \mapsto \mathrm{dist}_Z(x)^2$, in iii) satisfies

$$\nabla F(x) = 2\pi_{Z^\circ}(x) \quad \forall x \in \mathbb{R}^d. \tag{2.7}$$

Proof. All of the presented facts are well known. We include proofs of iii) and iv) for the convenience of the reader. We start with iii): Suppose that $x, h \in \mathbb{R}^d$ are given. Due to (2.3), we have

$$\left\langle \pi_Z(x+\tau h) - x - \tau h, \frac{\pi_Z(x) - \pi_Z(x+\tau h)}{\tau} \right\rangle \geq 0 \quad \text{and} \quad \left\langle \pi_Z(x) - x, \frac{\pi_Z(x+\tau h) - \pi_Z(x)}{\tau} \right\rangle \geq 0 \tag{2.8}$$

for all $\tau > 0$. From (2.4), it follows further that the difference quotients

$$\delta_\tau := \frac{\pi_Z(x+\tau h) - \pi_Z(x)}{\tau}$$

are bounded by $|h|$ for all $\tau > 0$. By compactness, this implies $\delta_k := \delta_{\tau_k} \to \delta$ for some sequence $0 < \tau_k \to 0$ and some $\delta \in \mathbb{R}^d$. By passing to the limit $k \to \infty$ along $\{\tau_k\}$ in (2.8), we obtain that $0 \leq \langle \pi_Z(x) - x, \delta \rangle \leq 0$, *i.e.*, $\langle \pi_Z(x) - x, \delta \rangle = 0$. Due to the definition of dist_Z and the identity $|a|^2 - |b|^2 = \langle a + b, a - b \rangle$ for $a, b \in \mathbb{R}^d$, we furthermore have

$$\begin{aligned}
\operatorname{dist}_Z(x+\tau h)^2 - \operatorname{dist}_Z(x)^2 &= |x + \tau h - \pi_Z(x+\tau h)|^2 - |x - \pi_Z(x)|^2 \\
&= \langle x + \tau h - \pi_Z(x+\tau h) + x - \pi_Z(x), x + \tau h - \pi_Z(x+\tau h) - x + \pi_Z(x) \rangle \\
&= \langle x + \tau h - \pi_Z(x+\tau h) + x - \pi_Z(x), \tau h - \pi_Z(x+\tau h) + \pi_Z(x) \rangle \qquad \forall \tau > 0.
\end{aligned}$$

By dividing by $\tau > 0$ and passing to the limit $k \to \infty$ along $\{\tau_k\}$ in the above, we obtain that

$$\lim_{k\to\infty} \frac{\operatorname{dist}_Z(x+\tau_k h)^2 - \operatorname{dist}_Z(x)^2}{\tau_k} = 2 \langle x - \pi_Z(x), h - \delta \rangle = 2 \langle x - \pi_Z(x), h \rangle . \tag{2.9}$$

Since the right-hand side of (2.9) does not depend on the choice of $\{\tau_k\}$, it follows that $\mathbb{R}^d \ni x \mapsto \operatorname{dist}_Z(x)^2 \in \mathbb{R}$ is Gâteaux differentiable. As the Gâteaux derivative depends continuously on x by the continuity of π_Z, the continuous Fréchet differentiability now follows from the mean value theorem. This completes the proof of iii). To establish iv), suppose that $x \in \mathbb{R}^d$ is given. From the cone property of Z and (2.3), it follows that

$$\langle \pi_Z(x) - x, \pi_Z(x) \rangle = 0 \leq \langle \pi_Z(x) - x, v \rangle \quad \forall v \in Z. \tag{2.10}$$

The above implies $x - \pi_Z(x) \in Z^\circ$. For all $w \in Z^\circ$, we further have $\langle \pi_Z(x), w \rangle \leq 0$ and, thus,

$$0 \geq \langle \pi_Z(x), w \rangle + \langle \pi_Z(x) - x, \pi_Z(x) \rangle = \langle \pi_Z(x), w + \pi_Z(x) - x \rangle \qquad \forall w \in Z^\circ.$$

Rewriting this inequality gives

$$x - \pi_Z(x) \in Z^\circ, \qquad \langle (x - \pi_Z(x)) - x, w - (x - \pi_Z(x)) \rangle \geq 0 \qquad \forall w \in Z^\circ.$$

Thus, $x - \pi_Z(x) = \pi_{Z^\circ}(x)$ holds as claimed. To see that $\pi_Z(sx) = s\pi_Z(x)$ holds for all $x \in \mathbb{R}^d$ and $s \in [0, \infty)$, it suffices to multiply (2.3) by s^2 and to exploit the cone property and i). This establishes (2.6). As (2.7) is a straightforward consequence of (2.5) and (2.6), iv) follows and the proof of the proposition is complete. ☐

2.3. Preliminaries on convex polyhedra

Convex polyhedra in $\mathbb{R}^d$, $d \in \mathbb{N}$, can be constructed as intersections of half-spaces of the form $H = \{x \in \mathbb{R}^d \mid \langle \nu, x \rangle \leq \alpha\}$, where $(\nu, \alpha) \in \mathbb{R}^d \times \mathbb{R}$ is a tuple with $|\nu| = 1$, *i.e.*, $\nu \in S^{d-1}$.

Definition 2.3 (convex polyhedron)**.** A set $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, is called a convex polyhedron if $Z = \mathbb{R}^d$ holds or if there exist an index set $I = \{1, ..., n\}$, $n \in \mathbb{N}$, and tuples $(\nu_i, \alpha_i) \in S^{d-1} \times \mathbb{R}$, $i \in I$, such that Z satisfies

$$Z = \bigcap_{i \in I} H_i, \qquad H_i := \left\{ x \in \mathbb{R}^d \mid \langle \nu_i, x \rangle \leq \alpha_i \right\}.$$

In the latter case, we call the n-tuple $\{(\nu_i, \alpha_i)\}_{i \in I} \in (S^{d-1} \times \mathbb{R})^n$ a description of Z. A polyhedron Z is called full-dimensional if it has a nonempty interior.

Remark 2.4 (conventions for the case $Z = \mathbb{R}^d$)**.** In what follows, some care has to be taken regarding the edge case $Z = \mathbb{R}^d$ in Definition 2.3. Henceforth, we use the convention that the polyhedron $Z = \mathbb{R}^d$ possesses precisely one description, namely the 0-tuple (a.k.a. the empty tuple). We further define the index set I belonging to this description to be the empty set and we use the convention that $\{(\nu_i, \alpha_i)\}_{i \in I}$ denotes the 0-tuple whenever $I = \emptyset$. Note that this means in particular that $Z = \mathbb{R}^d$ possesses one and only one description.

Note that every convex polyhedron $Z \neq \{0\}$ in $\mathbb{R}^d$, $d \in \mathbb{N}$, containing the origin is a full-dimensional polyhedron in the sense of Definition 2.3 when considered as a subset of its linear hull. The same is true for the case $Z = \{0\}$ if $\{0\}$ is understood as a full-dimensional polyhedron in $\mathbb{R}^0 := \{0\}$.

Definition 2.5 (set of active/inactive indices, regular description)**.** Given a convex polyhedron $Z \subset \mathbb{R}^d$ with a description $\{(\nu_i, \alpha_i)\}_{i \in I}$ and $x \in Z$, we denote by $\mathcal{A}(x) := \{i \in I \mid \langle \nu_i, x \rangle = \alpha_i\}$ the set of active indices of $\{(\nu_i, \alpha_i)\}_{i \in I}$ at x and by $\mathcal{I}(x) := \{i \in I \mid \langle \nu_i, x \rangle < \alpha_i\}$ the set of inactive indices of $\{(\nu_i, \alpha_i)\}_{i \in I}$ at x. A description is called regular if, for every $x \in Z$, the vectors in the set $\{\nu_i \mid i \in \mathcal{A}(x)\}$ are linearly independent.

The octahedron in $\mathbb{R}^3$ (or, more generally, the closed unit ball w.r.t. the ℓ^1-norm in $\mathbb{R}^d$, $d \geq 3$) is an example of a polyhedron that does not possess a regular description since four facets meet at corner points.

Descriptions are typically not unique. As they are defined as tuples, we can reorder a given description, for example, to generate a new one with different maps $\mathcal{A}, \mathcal{I} \colon Z \rightrightarrows \mathbb{N}$. Further, a description of a polyhedron Z may contain tuples (ν_j, α_j) that are redundant in the sense that their removal does not change Z, *i.e.*,

$$Z = \bigcap_{i \in I} \{x \in \mathbb{R}^d \mid \langle \nu_i, x \rangle \leq \alpha_i\} = \bigcap_{i \in I, i \neq j} \{x \in \mathbb{R}^d \mid \langle \nu_i, x \rangle \leq \alpha_i\}.$$

The following well-known result shows that, for a full-dimensional convex polyhedron, it is possible to overcome these uniqueness and redundancy issues (at least in a certain sense) by considering *standard descriptions.*

Proposition 2.6 (standard descriptions)**.** *Let $Z \subset \mathbb{R}^d$ be a full-dimensional convex polyhedron. Then there exists a description $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i \in I}$ of Z that contains no redundant half-spaces. This description is unique up to permutations of the tuples (ν_i, α_i) and satisfies*

$$\forall i \in I \;\exists x_i \in Z\colon \quad \mathcal{A}(x_i) = \{i\} \quad (\textit{that is, } \; \langle \nu_i, x_i \rangle = \alpha_i \textit{ and } \langle \nu_j, x_i \rangle < \alpha_j \; \forall j \neq i). \tag{2.11}$$

Henceforth, we call a description $\mathcal{H}$ with the above properties a standard description *of the polyhedron Z.*

Proof. This is a standard result from linear programming and polyhedral theory; see [53], Theorem 1.6 or [54], Theorem 12.1.5. Note that the 0-tuple is the unique standard description of $Z = \mathbb{R}^d$ by our conventions. ☐

Corollary 2.7 (standard descriptions of cones)**.** *Let $Z \subset \mathbb{R}^d$ be a full-dimensional convex polyhedral cone and let $\mathcal{H}$ be a standard description of Z. Then $\mathcal{H}$ has the form $\mathcal{H} = \{(\nu_i, 0)\}_{i \in I}$.*

Proof. Let $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i \in I}$ be a standard description of Z and let $i \in I$ be arbitrary. Since $0 \in Z$, we have $\alpha_i \geq \langle \nu_i, 0 \rangle = 0$. For x_i from (2.11), it further holds $\langle \nu_i, 2x_i \rangle = 2\alpha_i \leq \alpha_i$ as $2x_i \in Z$. Thus $\alpha_i = 0$. ☐

Next, we define the *linearization cone*, which provides a local approximation for a convex polyhedron.

Definition 2.8 (linearization cone)**.** Let $Z \subset \mathbb{R}^d$ be a nonempty convex polyhedron with description $\{(\nu_i, \alpha_i)\}_{i \in I}$. Then, for every $x \in Z$, we define $\mathcal{Z}(x) := \{z \in \mathbb{R}^d \mid \langle \nu_i, z \rangle \leq 0 \ \forall i \in \mathcal{A}(x)\}$.

Note that we have $Z - x \subset \mathcal{Z}(x)$ for all $x \in Z$ in Definition 2.8 since $\langle \nu_j, y - x \rangle = \langle \nu_j, y \rangle - \alpha_j \leq 0$ holds for all $y \in Z$ and all $j \in \mathcal{A}(x)$. It will be shown immediately below in Lemma 2.10iii) that the cone $\mathcal{Z}(x)$ is independent of the chosen description of Z.

Lemma 2.9 (full-dimensionality and standard descriptions of the linearization cone)**.** *Let $Z \subset \mathbb{R}^d$ be a full-dimensional convex polyhedron. Let $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i \in I}$ be a standard description of Z and let $x \in Z$ be given. Then $\mathcal{Z}(x)$ is full-dimensional and $\{(\nu_i, 0)\}_{i \in \mathcal{A}(x)}$ is a standard description of $\mathcal{Z}(x)$.*

Proof. Since Z is full-dimensional, so is $\mathcal{Z}(x) \supset Z - x$. It remains to prove the statement about the standard description. To this end, we assume w.l.o.g. that $0 \in \operatorname{int}(Z)$. (This can always be achieved by a translation.) Note that this inclusion implies $0 = \langle \nu_i, 0 \rangle < \alpha_i$ for all $i \in I$. Consider now some $i \in \mathcal{A}(x)$. We have to show that the half-space $H_i = \{z \in \mathbb{R}^d \mid \langle \nu_i, z \rangle \leq 0\}$ is not redundant for $\mathcal{Z}(x)$. We choose $x_i \in Z$ according to (2.11) and set $z := (1 + \varepsilon)x_i - x$ with $\varepsilon > 0$. Then $z \notin H_i$ since $\langle \nu_i, z \rangle = \varepsilon \alpha_i > 0$. On the other hand, for $j \in \mathcal{A}(x) \setminus \{i\}$ we have $\langle \nu_j, x_i - x \rangle < 0$, so $\langle \nu_j, z \rangle = \varepsilon \langle \nu_j, x_i \rangle + \langle \nu_j, x_i - x \rangle \leq 0$ if ε is small enough. This shows that H_i is indeed nonredundant for $\mathcal{Z}(x)$ and completes the proof. □

Lemma 2.10 (properties of the linearization cone)**.** *Let $Z \subset \mathbb{R}^d$ be a nonempty convex polyhedron with a description $\{(\nu_i, \alpha_i)\}_{i \in I}$. Then the following statements are true.*

i) For every $x \in Z$, there exists $\varepsilon > 0$ such that

$$w + sz \in Z \qquad \forall w \in B_\varepsilon(x) \cap Z \quad \forall s \in [0, \varepsilon] \quad \forall z \in B_1(0) \cap \mathcal{Z}(x).$$

ii) For every $x \in \mathbb{R}^d$, there exists $\varepsilon > 0$ such that

$$\pi_Z(x + z) = \pi_Z(x) + \pi_{\mathcal{Z}(\pi_Z(x))}(x + z - \pi_Z(x)) \qquad \forall z \in B_\varepsilon(0).$$

iii) For every $x \in Z$, it holds $\mathcal{Z}(x) = K_{\mathrm{rad}}(x; Z) = K_{\tan}(x; Z)$. In particular, if Z is a nonempty convex polyhedral cone, then $\mathcal{Z}(0) = K_{rad}(0; Z) = Z$.

iv) If $\{(\nu_i, \alpha_i)\}_{i \in I}$ is regular, then, for every point $x \in Z$, there exists a unique collection of vectors $\{e_i\}_{i \in \mathcal{A}(x)} \subset \mathcal{Z}(x)$ (empty in the case $\mathcal{A}(x) = \emptyset$) satisfying

$$e_i \in \operatorname{span}(\{\nu_k \mid k \in \mathcal{A}(x)\}), \quad |e_i| = 1, \quad \text{and} \quad \langle \nu_i, e_i \rangle < 0 = \langle \nu_j, e_i \rangle \quad \forall i, j \in \mathcal{A}(x), i \neq j. \tag{2.12}$$

The vectors e_i, $i \in \mathcal{A}(x)$, form a basis of $\operatorname{span}(\{\nu_k \mid k \in \mathcal{A}(x)\})$, and it holds

$$z \in \mathcal{Z}(x) \quad \Leftrightarrow \quad z = z_1 + z_2 \text{ with } z_1 \in \operatorname{span}(\{\nu_i \mid i \in \mathcal{A}(x)\})^\perp, z_2 = \sum_{i \in \mathcal{A}(x)} \beta_i e_i,\ \beta_i \geq 0 \ \forall i \in \mathcal{A}(x). \tag{2.13}$$

Proof. We begin with i): Let $x \in Z$. As the functions $F_i \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $(w, z) \mapsto \langle \nu_i, w + z \rangle$, $i \in I$, are continuous and satisfy $F_i(x, 0) < \alpha_i$ for all $i \in \mathcal{I}(x)$, we can find $\varepsilon > 0$ such that $F_i(w, z) < \alpha_i$ holds for all $w \in B_\varepsilon(x)$, $z \in B_\varepsilon(0)$, and $i \in \mathcal{I}(x)$. This implies that, for all $i \in \mathcal{I}(x)$, we have $\langle \nu_i, w + sz \rangle \leq \alpha_i$ for all $w \in B_\varepsilon(x) \cap Z$, $s \in [0, \varepsilon]$, and $z \in B_1(0) \cap \mathcal{Z}(x)$. If, on the other hand, $i \in \mathcal{A}(x)$, then we have $\langle \nu_i, w + sz \rangle = \langle \nu_i, w \rangle + s\langle \nu_i, z \rangle \leq \alpha_i + 0$ for all $w \in B_\varepsilon(x) \cap Z$, $s \in [0, \varepsilon]$, and $z \in B_1(0) \cap \mathcal{Z}(x)$ by the definitions of Z and $\mathcal{Z}(x)$. This establishes i).

To prove ii), suppose that $x \in \mathbb{R}^d$ is given. From the Lipschitz continuity of π_Z, we obtain that there exists $\varepsilon > 0$ such that $\mathcal{A}(\pi_Z(x+z)) \subset \mathcal{A}(\pi_Z(x))$ holds for all $z \in B_\varepsilon(0)$. In combination with (2.2), this yields

$$\pi_Z(x+z) = \arg\min \left\{ \frac{1}{2}|v - x - z|^2 \;\middle|\; \langle \nu_i, v\rangle \leq \alpha_i \ \forall i \in \mathcal{A}(\pi_Z(x)) \right\} \qquad \forall z \in B_\varepsilon(0),$$

which, by means of the variable transformation $w = v - \pi_Z(x)$, can be rewritten as

$$\begin{aligned}\pi_Z(x+z) &= \pi_Z(x) + \arg\min \left\{ \frac{1}{2}|w + \pi_Z(x) - x - z|^2 \;\middle|\; \langle \nu_i, w\rangle \leq 0 \ \forall i \in \mathcal{A}(\pi_Z(x)) \right\} \\ &= \pi_Z(x) + \pi_{\mathcal{Z}(\pi_Z(x))}(x + z - \pi_Z(x)) \qquad \forall z \in B_\varepsilon(0).\end{aligned}$$

This establishes ii).

Next, we prove iii): Let $x \in Z$ be given. Since $\mathcal{Z}(x)$ is closed, it suffices to prove that $\mathcal{Z}(x) = K_{\text{rad}}(x; Z)$. If $z = s(y - x) \in K_{\text{rad}}(x; Z)$ with $y \in Z$ and $s > 0$, then $\langle \nu_i, z\rangle = s(\langle \nu_i, y\rangle - \alpha_i) \leq 0$ for all $i \in \mathcal{A}(x)$. Conversely, if $0 \neq z \in \mathcal{Z}(x)$, then from i) we obtain that $x + sz = x + s|z|(z/|z|) \in Z$ if $s > 0$ is small enough.

It remains to prove iv): Let $x \in Z$ be given. We assume w.l.o.g. that $\mathcal{A}(x) \neq \emptyset$. (For the degenerate case $\mathcal{A}(x) = \emptyset$, the assertion is true due to the conventions $\text{span}(\emptyset) = \{0\}$, $\sum_{i\in\emptyset} \beta_i e_i = 0$, *etc.*) Fix $i \in \mathcal{A}(x)$. We set $V := \text{span}(\{\nu_j \mid j \in \mathcal{A}(x)\})$ and $W := \text{span}(\{\nu_j \mid j \in \mathcal{A}(x) \setminus \{i\}\})$. Since the vectors ν_j, $j \in \mathcal{A}(x)$, are linearly independent by our assumptions, we have $\dim(W) = \dim(V) - 1$ and $0 \neq e := \nu_i - \pi_W(\nu_i) \in W^\perp \cap V$. Moreover, $\dim(W^\perp) = d + 1 - \dim(V)$ and $\dim(W^\perp + V) = d$ since $W^\perp \supset V^\perp$. It follows that $\dim(W^\perp \cap V) = 1$ and $\langle \nu_i, e\rangle = \langle e, e\rangle \neq 0$. Thus, $W^\perp \cap V$ contains precisely one vector e_i – a scalar multiple of e – satisfying (2.12). By repeating this construction for all other indices in $\mathcal{A}(x)$, the uniqueness and existence of vectors e_i, $i \in \mathcal{A}(x)$, with the properties in (2.12) follow. That the vectors e_i, $i \in \mathcal{A}(x)$, are linearly independent and, consequently, form a basis of V follows straightforwardly from (2.12). It remains to prove (2.13). The implication "$\Leftarrow$" in this equivalence is again a direct consequence of (2.12). To prove "$\Rightarrow$", let $z \in \mathcal{Z}(x)$ be given. Then there exist unique $z_1 \in V^\perp$ and $z_2 \in V$ satisfying $z = z_1 + z_2$. Since the vectors e_i, $i \in \mathcal{A}(x)$, form a basis of V, there exist unique $\beta_i \in \mathbb{R}$, $i \in \mathcal{A}(x)$, satisfying $z_2 = \sum_{i\in\mathcal{A}(x)} \beta_i e_i$. From $z \in \mathcal{Z}(x)$ and (2.12), it follows that $0 \geq \langle \nu_j, z\rangle = \langle \nu_j, z_2\rangle = \beta_j \langle \nu_j, e_j\rangle$ holds for all $j \in \mathcal{A}(x)$, where $\langle \nu_j, e_j\rangle < 0$. Thus, $\beta_j \geq 0$ for all $j \in \mathcal{A}(x)$ and (2.13) is proved. This establishes iv) and completes the proof of the lemma. □

Lemma 2.11 (formulas for polar and normal cones)**.**

i) Let Z be a convex polyhedral cone with a description $\{(\nu_i, 0)\}_{i\in I}$. Then it holds

$$Z^\circ = \left\{ \sum_{i\in I} \beta_i \nu_i \;\middle|\; \beta_i \geq 0 \ \forall i \in I \right\}.$$

ii) Let Z be a nonempty convex polyhedron with a description $\{(\nu_i, \alpha_i)\}_{i\in I}$. Then, for every $x \in Z$, it holds

$$\mathcal{N}_Z(x) = \mathcal{Z}(x)^\circ = \left\{ \sum_{i\in\mathcal{A}(x)} \beta_i \nu_i \;\middle|\; \beta_i \geq 0 \ \forall i \in \mathcal{A}(x) \right\}.$$

Proof. This is [55], Lemma 3.1. □

Note that the right-hand sides of the formulas in Lemma 2.11 are understood as $\{0\}$ in the case that the summation index set is empty (by the convention for the empty sum). We are now in the position to introduce:

Definition 2.12 (non-obtuse polyhedron)**.** A full-dimensional convex polyhedron $Z \subset \mathbb{R}^d$ is called *non-obtuse* if $\langle \nu_i, \nu_j\rangle \leq 0$ for all $i, j \in I$ with $i \neq j$ holds for one (and thus all) of its standard descriptions $\{(\nu_i, \alpha_i)\}_{i\in I}$.

The following two results collect properties of non-obtuse polyhedra that are essential for the analysis of the jumps of the directional derivatives of the vectorial stop operator in Sections 3.3 and 3.4.

Proposition 2.13 (regularity of non-obtuse polyhedra)**.** *Let $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a full-dimensional convex non-obtuse polyhedron and let $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i \in I}$ be a standard description of Z. Then $\mathcal{H}$ is regular.*

Proof. Let $x \in Z$ be arbitrary and assume that $\beta_i \in \mathbb{R}$, $i \in \mathcal{A}(x)$, are numbers such that $\sum_{i \in \mathcal{A}(x)} \beta_i \nu_i = 0$ holds. Choose $\tilde{x} \in \operatorname{int}(Z)$. Then it holds $\langle \nu_i, \tilde{x} \rangle < \alpha_i$ for all $i \in I$ and $h := \tilde{x} - x \in \mathbb{R}^d$ satisfies $\langle \nu_i, h \rangle < 0$ for all $i \in \mathcal{A}(x)$. Since Z is non-obtuse, we further obtain

$$\begin{aligned} 0 = \left| \sum_{i \in \mathcal{A}(x)} \beta_i \nu_i \right|^2 &= \left| \sum_{i \in \mathcal{A}(x), \beta_i > 0} \beta_i \nu_i \right|^2 + 2 \sum_{i \in \mathcal{A}(x), \beta_i > 0} \sum_{j \in \mathcal{A}(x), \beta_j < 0} \beta_i \beta_j \langle \nu_i, \nu_j \rangle + \left| \sum_{i \in \mathcal{A}(x), \beta_i < 0} \beta_i \nu_i \right|^2 \\ &\geq \left| \sum_{i \in \mathcal{A}(x), \beta_i > 0} \beta_i \nu_i \right|^2 + \left| \sum_{i \in \mathcal{A}(x), \beta_i < 0} \beta_i \nu_i \right|^2 . \end{aligned}$$

The above implies

$$0 = \left\langle h, \sum_{i \in \mathcal{A}(x), \beta_i > 0} \beta_i \nu_i \right\rangle = \sum_{i \in \mathcal{A}(x), \beta_i > 0} \beta_i \langle h, \nu_i \rangle \quad \text{and} \quad 0 = \left\langle h, \sum_{i \in \mathcal{A}(x), \beta_i < 0} \beta_i \nu_i \right\rangle = \sum_{i \in \mathcal{A}(x), \beta_i < 0} \beta_i \langle h, \nu_i \rangle .$$

As $\langle h, \nu_i \rangle < 0$ for all $i \in \mathcal{A}(x)$, this can only be true if there is no $i \in \mathcal{A}(x)$ with $\beta_i \neq 0$. This shows that the vectors in the set $\{\nu_i \mid i \in \mathcal{A}(x)\}$ are linearly independent and that $\mathcal{H}$ is indeed regular. □

Lemma 2.14 (properties of non-obtuse polyhedra)**.** *Let $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a full-dimensional non-obtuse convex polyhedron and let $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i \in I}$ be a standard description of Z. Then the following is true.*

i) It holds

$$x, y \in Z, v \in \mathbb{R}^d, x + v \in Z \qquad \Rightarrow \qquad x + \pi_{\mathcal{Z}(y)^\circ}(v) \in Z. \tag{2.14}$$

ii) For all $j \in I$, it holds

$$x \in \mathbb{R}^d,\ \langle \nu_j, x \rangle = \alpha_j \qquad \Rightarrow \qquad \langle \nu_j, \pi_Z(x) \rangle = \alpha_j.$$

iii) If Z is additionally a cone, then it holds

$$x, y \in Z \qquad \Rightarrow \qquad \pi_{y + Z^\circ}(x) \in Z \text{ and } x - \pi_{y+Z^\circ}(x) = \pi_Z(x - y) \in Z.$$

iv) If Z is additionally a cone, $J \subset I$ an index set, and $V := \operatorname{span}(\{\nu_i \mid i \in J\})^\perp$, then it holds

$$x \in Z \qquad \Rightarrow \qquad \pi_V(x) \in Z.$$

Proof. We begin with i): Let x, y, and v be as on the left-hand side of (2.14). For $\mathcal{Z}(y) = \mathbb{R}^d$, (2.14) is trivially true. We may thus assume w.l.o.g. that $\mathcal{Z}(y) \neq \mathbb{R}^d$. In this case, we have $\mathcal{A}(y) \neq \emptyset$. Applying Proposition 2.2iv) and Lemma 2.11, we see that there exist $\beta_i \geq 0$, $i \in \mathcal{A}(y)$, such that

$$\pi_{\mathcal{Z}(y)}(v) + \pi_{\mathcal{Z}(y)^\circ}(v) = v, \qquad \langle \pi_{\mathcal{Z}(y)^\circ}(v), \pi_{\mathcal{Z}(y)}(v) \rangle = 0, \qquad \pi_{\mathcal{Z}(y)^\circ}(v) = \sum_{i \in \mathcal{A}(y)} \beta_i \nu_i. \tag{2.15}$$

It follows that $\sum_{i\in\mathcal{A}(y)} \beta_i \langle \nu_i, \pi_{\mathcal{Z}(y)}(v)\rangle = 0$. Since all summands are nonpositive,

$$\beta_j \langle \nu_j, \pi_{\mathcal{Z}(y)}(v)\rangle = 0 \qquad \forall j \in \mathcal{A}(y). \tag{2.16}$$

Consider now some $j \in I$. If $j \notin \mathcal{A}(y)$ or $j \in \mathcal{A}(y)$ and $\beta_j = 0$, then the non-obtuseness of Z, $x \in Z$, and (2.15) yield

$$\left\langle \nu_j, x + \pi_{\mathcal{Z}(y)^\circ}(v)\right\rangle = \langle \nu_j, x\rangle + \sum_{i\in\mathcal{A}(y), i\neq j} \beta_i \langle \nu_j, \nu_i\rangle \leq \langle \nu_j, x\rangle \leq \alpha_j.$$

If, on the other hand, $j \in \mathcal{A}(y)$ and $\beta_j > 0$, then it follows from (2.16) and $x + v \in Z$ that

$$\left\langle \nu_j, x + \pi_{\mathcal{Z}(y)^\circ}(v)\right\rangle = \left\langle \nu_j, x + \pi_{\mathcal{Z}(y)^\circ}(v) + \pi_{\mathcal{Z}(y)}(v)\right\rangle = \langle \nu_j, x + v\rangle \leq \alpha_j.$$

Thus, $x + \pi_{\mathcal{Z}(y)^\circ}(v) \in Z$ and the proof is complete.

Next, we prove ii): Let $x \in \mathbb{R}^d$ with $\langle \nu_j, x\rangle = \alpha_j$ for some $j \in I$. Since $x - \pi_Z(x)$ belongs to the normal cone to Z at $\pi_Z(x)$, it follows from Lemma 2.11 that there exist $\beta_i \geq 0$ such that

$$x - \pi_Z(x) = \sum_{i\in\mathcal{A}(\pi_Z(x))} \beta_i \nu_i.$$

Since $\langle \nu_j, x\rangle = \alpha_j$,

$$\alpha_j - \langle \nu_j, \pi_Z(x)\rangle = \sum_{i\in\mathcal{A}(\pi_Z(x))} \beta_i \langle \nu_j, \nu_i\rangle. \tag{2.17}$$

Suppose now that $\langle \nu_j, \pi_Z(x)\rangle < \alpha_j$. Then the left side of (2.17) is > 0, but the right side is ≤ 0 as $j \notin \mathcal{A}(\pi_Z(x))$ and Z is non-obtuse, a contradiction. Thus, $\langle \nu_j, \pi_Z(x)\rangle = \alpha_j$ and the proof of ii) is complete.

To prove iii), suppose that Z is a cone and that $x, y \in Z$ are given. We claim that $\pi_{y+Z^\circ}(x) = y + \pi_{Z^\circ}(x - y)$ holds. To see this, we note that we trivially have $y + \pi_{Z^\circ}(x - y) \in y + Z^\circ$ and that (2.3) for π_{Z° yields

$$\langle y + \pi_{Z^\circ}(x - y) - x, w - y - \pi_{Z^\circ}(x - y)\rangle = \langle \pi_{Z^\circ}(x - y) - (x - y), (w - y) - \pi_{Z^\circ}(x - y)\rangle \geq 0 \quad \forall w \in y + Z^\circ.$$

Thus, $\pi_{y+Z^\circ}(x) = y + \pi_{Z^\circ}(x - y)$ as claimed. From (2.6), we may now deduce that $x = y + \pi_{Z^\circ}(x - y) + \pi_Z(x - y) = \pi_{y+Z^\circ}(x) + \pi_Z(x - y)$. The assertion $x - \pi_{y+Z^\circ}(x) = \pi_Z(x - y) \in Z$ follows immediately from this identity. It remains to prove that $\pi_{y+Z^\circ}(x) = y + \pi_{Z^\circ}(x - y) \in Z$. Since $\mathcal{Z}(0) = Z$ by Lemma 2.10iii), setting $y = 0$ in (2.14) and replacing there x by y as well as v by $x - y$, we see that $y + \pi_{Z^\circ}(x - y) \in Z$. This establishes iii).

It remains to prove iv): By Corollary 2.7, we have $\mathcal{H} = \{(\nu_i, 0)\}_{i\in I}$. If $J = \emptyset$ holds, then we have $V = \mathbb{R}^d$ by our convention $\operatorname{span}(\emptyset) = \{0\}$ and there is nothing to prove. So let us assume that $\emptyset \neq J \subset I = \{1, ..., n\}$, $n \in \mathbb{N}$, holds; w.l.o.g. $J = \{1, ..., m\}$ for some $m \leq n$. Suppose further that $x \in Z$ is given. Then there exists $\beta := (\beta_1, ..., \beta_m)^\top \in \mathbb{R}^m$ such that $\pi_V(x) + \sum_{j=1}^m \beta_j \nu_j = x$. Define $M := (\nu_1, ..., \nu_m) \in \mathbb{R}^{d\times m}$. Then it follows from $x \in Z$ and $\langle \nu_i, \pi_V(x)\rangle = 0$ for all $i = 1, ..., m$ that $M^\top M\beta = M^\top (\pi_V(x) + M\beta) = M^\top x \in (-\infty, 0]^m$. By taking the scalar product with the vector $\tilde{\beta} := (\max(0, \beta_1), ..., \max(0, \beta_m))^\top \in \mathbb{R}^m$ in the above, by exploiting the non-obtuseness of Z, and by distinguishing cases, we obtain that

$$0 \geq \sum_{i=1}^m \sum_{j=1}^m \max(0, \beta_i)\langle \nu_i, \nu_j\rangle \beta_j \geq \sum_{i=1}^m \sum_{j=1}^m \max(0, \beta_i)\langle \nu_i, \nu_j\rangle \max(0, \beta_j) = \langle M\tilde{\beta}, M\tilde{\beta}\rangle = |M\tilde{\beta}|^2.$$

Since $\mathcal{H}$ is regular due to Proposition 2.13, the above implies $\tilde{\beta} = 0$ and, as a consequence, $\beta_i \leq 0$ for all $i = 1, ..., m$. Since Z is non-obtuse and $x \in Z$, the nonpositivity of the numbers β_i, $i = 1, ..., m$, implies

$$\langle \nu_i, \pi_V(x) \rangle = \left\langle \nu_i, x - \sum_{j=1}^{m} \beta_j \nu_j \right\rangle = \langle \nu_i, x \rangle - \sum_{j=1}^{m} \beta_j \langle \nu_i, \nu_j \rangle \leq \langle \nu_i, x \rangle \leq 0 \qquad \forall i \in I \setminus J.$$

As $\langle \nu_i, \pi_V(x) \rangle = 0$ holds for all $i \in J$ by the definition of V, this shows $\pi_V(x) \in Z$ and completes the proof. $\square$

Next, we study the properties of polyhedra that are not non-obtuse. Our main goal is to prove that such polyhedra always possess an obtuse ridge. This result is essential for the construction of the counterexamples in Section 4. We proceed in two steps, beginning with the special case of a cone.

Lemma 2.15 (existence of an obtuse ridge for a polyhedral cone that is not non-obtuse). *Let $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a full-dimensional convex polyhedral cone that is not non-obtuse. Let $\mathcal{H} = \{(\nu_i, 0)\}_{i \in I}$ be a standard description of Z. Then there exists $w \in Z$ such that $\mathcal{A}(w) = \{i_1, i_2\}$ holds for some $i_1, i_2 \in I$ satisfying $\langle \nu_{i_1}, \nu_{i_2} \rangle > 0$.*

Proof. We use induction w.r.t. the cardinality $n \in \mathbb{N}_0$ of the index set I of the standard description of Z. For $n \in \{0, 1\}$, there are no polyhedral cones that are not non-obtuse. For $n = 2$, the assertion is trivial.

Suppose now that $2 < n \in \mathbb{N}$ is given and that the assertion is proven for all cones with the properties in the statement of the lemma whose standard descriptions have at most $n - 1$ elements. Assume further that a cone $Z \subset \mathbb{R}^d$ fulfilling the assumptions of the lemma is given with a standard description $\{(\nu_i, 0)\}_{i \in I}$ satisfying $|I| = n$. As Z is not non-obtuse, there exist $i_1, i_2 \in I$ with $i_1 < i_2$ such that $\langle \nu_{i_1}, \nu_{i_2} \rangle \in (0, 1)$. Define

$$J := \{j \in I \setminus \{i_1\} \mid \exists x \in Z \cap \operatorname{span}(\{\nu_1, ..., \nu_n\}) \setminus \{0\} \text{ s.t. } \langle \nu_{i_1}, x \rangle = \langle \nu_j, x \rangle = 0\}.$$

We claim that $J \neq \emptyset$. To see this, choose $\tilde{x}$ such that the conditions in (2.11) hold for i_1. We may assume w.l.o.g. that $\tilde{x} \in \operatorname{span}(\{\nu_1, ..., \nu_n\})$ (else we replace $\tilde{x}$ by the projection $\pi_{\operatorname{span}(\{\nu_1,...,\nu_n\})}(\tilde{x})$). Choose further $y \in \operatorname{span}(\{\nu_{i_1}, \nu_{i_2}\})^\perp \cap \operatorname{span}(\{\nu_1, ..., \nu_n\}) \setminus \{0\}$. (Such a vector y exists since $\dim(\operatorname{span}(\{\nu_1, ..., \nu_n\})) \geq 3$ by our assumption $n \geq 3$ and the properties in Proposition 2.6.) Define $x_s := (1 - s)\tilde{x} + sy$, $s \in [0, 1]$. Then it holds

$$\langle \nu_{i_1}, x_s \rangle = 0 \ \ \forall s \in [0, 1], \qquad \langle \nu_{i_2}, x_1 \rangle = 0, \qquad \langle \nu_{i_2}, x_s \rangle < 0 \ \ \forall s \in [0, 1).$$

If $y = x_1 \in Z$ holds, then the above yields $i_2 \in J$ and we obtain $J \neq \emptyset$ as desired. If $y = x_1 \notin Z$ holds, then we have $\bar{s} := \sup\{s \in [0, 1] \mid x_s \in Z\} \in (0, 1)$. This implies that there exists $l \in I \setminus \{i_1\}$ such that $\langle \nu_l, x_s \rangle < 0$ holds for all $s \in [0, \bar{s})$ and $\langle \nu_l, x_{\bar{s}} \rangle = 0$. As $\langle \nu_{i_2}, x_{\bar{s}} \rangle < 0$ holds and due to the properties of $\tilde{x}$, y, and $\bar{s}$, we have $x_{\bar{s}} \in Z \cap \operatorname{span}(\{\nu_1, ..., \nu_n\}) \setminus \{0\}$. Thus, $l \in J$ and $J \neq \emptyset$ as asserted.

Next, we prove that

$$Z \cap \operatorname{span}(\{\nu_1, ..., \nu_n\}) \cap \operatorname{span}(\{\nu_{i_1}\})^\perp = \{x \in \operatorname{span}(\{\nu_1, ..., \nu_n\}) \mid \langle \nu_j, x \rangle \leq 0 = \langle \nu_{i_1}, x \rangle \, \forall j \in J\}. \tag{2.18}$$

To see that this assertion is true, we first note that the definitions of the involved sets immediately imply that "$\subset$" holds in (2.18). To see that we also have "$\supset$", we argue by contradiction. Suppose that there exists $x \in \mathbb{R}^d$ that is contained in the set on the right-hand side of (2.18) but not in the set on the left. Then it holds $x \in \operatorname{span}(\{\nu_1, ..., \nu_n\}) \setminus \{0\}$, $\langle \nu_{i_1}, x \rangle = 0$, and there exists a nonempty index set $I' \subset I \setminus (J \cup \{i_1\})$ such that $\langle \nu_i, x \rangle > 0$ for all $i \in I'$ and $\langle \nu_i, x \rangle \leq 0$ for all $i \in I \setminus I'$. We again consider a point $\tilde{x} \in \operatorname{span}(\{\nu_1, ..., \nu_n\})$ that satisfies (2.11) for the index i_1 and define $x_s := (1 - s)\tilde{x} + sx$, $s \in [0, 1]$. Then it holds $x_s \in \operatorname{span}(\{\nu_1, ..., \nu_n\})$ for all $s \in [0, 1]$ and

$$\langle \nu_i, x_0 \rangle < 0 \ \forall i \in I \setminus \{i_1\}, \quad \langle \nu_{i_1}, x_s \rangle = 0 \ \forall s \in [0, 1], \quad \langle \nu_i, x_s \rangle < 0 \ \forall s \in [0, 1) \, \forall i \in J, \quad \langle \nu_i, x_1 \rangle > 0 \ \forall i \in I'.$$

Define $\bar{s} := \sup\{s \in [0,1] \mid x_s \in Z\}$. Then we again have $\bar{s} \in (0,1)$ and for $x_{\bar{s}}$ there exists $l \in I \setminus (J \cup \{i_1\})$ such that $\langle \nu_l, x_s\rangle < 0$ holds for all $s \in [0, \bar{s})$ and $\langle \nu_l, x_{\bar{s}}\rangle = 0$. As $J \neq \emptyset$ and $\langle \nu_i, x_s\rangle < 0$ for all $s \in [0,1)$ and all $i \in J$ and due to the definition of $\bar{s}$, we have $x_{\bar{s}} \in Z \setminus \{0\}$. By the definition of J, this yields $l \in J$, which creates a contradiction. Thus, (2.18) holds as claimed.

Now we can turn our attention to the proof of the assertion of the lemma. Let us first consider the special case $J = \{i_2\}$. In this case, it follows from the definition of J that there exists $x \in Z \cap \operatorname{span}(\{\nu_1, ..., \nu_n\}) \setminus \{0\}$ such that $\langle \nu_{i_1}, x\rangle = \langle \nu_{i_2}, x\rangle = 0$ holds and it follows from $J = \{i_2\}$ that this x satisfies $\langle \nu_j, x\rangle < 0$ for all $j \in I \setminus \{i_1, i_2\}$. In this situation, the assertion of the lemma holds with $w := x$ and we are done.

Henceforth, we may thus assume that $J \setminus \{i_2\} \neq \emptyset$. Let now $\{\mu_1, ..., \mu_m\}$ be a basis of $\operatorname{span}(\{\nu_1, ..., \nu_n\})^\perp$. (If $\operatorname{span}(\{\nu_1, ..., \nu_n\}) = \mathbb{R}^d$ holds, then we set $m := 0$, $\{\mu_1, ..., \mu_m\} := \emptyset$.) We invoke Motzkin's theorem of the alternative [56], Theorem 4.2, p. 28 to obtain that precisely one of the following statements is true:

(I) There exists $h \in \mathbb{R}^d$ such that

$$\langle \nu_i, h\rangle < 0 \ \forall i \in J \setminus \{i_2\}, \qquad \langle \nu_{i_1}, h\rangle = \langle \nu_{i_2}, h\rangle = 0, \qquad \langle \mu_i, h\rangle = 0 \quad \forall i = 1, ..., m.$$

(II) There exist $\beta_i \geq 0$, $i \in J \setminus \{i_2\}$, not all zero, and numbers $\gamma_1, \gamma_2, \eta_i \in \mathbb{R}$, $i = 1, ..., m$, such that

$$\sum_{i \in J \setminus \{i_2\}} \beta_i \nu_i = \gamma_1 \nu_{i_1} + \gamma_2 \nu_{i_2} + \sum_{i=1}^m \eta_i \mu_i. \tag{2.19}$$

Let us first consider option (I). In this case, $\langle \mu_i, h\rangle = 0$ for all $i = 1, ..., m$ yields $h \in \operatorname{span}(\{\nu_1, ..., \nu_n\})$, we have $\langle \nu_{i_1}, h\rangle = \langle \nu_{i_2}, h\rangle = 0$, h is not zero since $J \setminus \{i_2\}$ is not empty, and it holds $h \in Z$ by (2.18). Due to the definition of the set J, this yields that we also have $\langle \nu_i, h\rangle < 0$ for all $i \in I \setminus (J \cup \{i_1, i_2\})$. Thus, after collecting everything, we obtain $\langle \nu_{i_1}, h\rangle = \langle \nu_{i_2}, h\rangle = 0$ and $\langle \nu_i, h\rangle < 0$ for all $i \in I \setminus \{i_1, i_2\}$. In this case, $w := h$ satisfies all of the conditions in the lemma and we are done.

It remains to study option (II). In this case, by again choosing $\tilde{x} \in \operatorname{span}(\{\nu_1, ..., \nu_n\})$ for i_1 as in (2.11) and by exploiting that the coefficients β_i are nonnegative and not all zero, we obtain that

$$0 > \sum_{i \in J \setminus \{i_2\}} \langle \beta_i \nu_i, \tilde{x}\rangle = \langle \gamma_1 \nu_{i_1} + \gamma_2 \nu_{i_2}, \tilde{x}\rangle = \gamma_2 \left\langle \nu_{i_2}, \tilde{x}\right\rangle.$$

Since $\langle \nu_{i_2}, \tilde{x}\rangle < 0$, this yields $\gamma_2 > 0$. Analogously, we also obtain $\gamma_1 > 0$. By taking the scalar product with ν_{i_1} in (2.19) and by exploiting that $\langle \nu_{i_1}, \nu_{i_2}\rangle > 0$, we now obtain

$$0 < \gamma_1 < \langle \gamma_1 \nu_{i_1} + \gamma_2 \nu_{i_2}, \nu_{i_1}\rangle = \sum_{i \in J \setminus \{i_2\}} \beta_i \langle \nu_i, \nu_{i_1}\rangle.$$

Due to $\beta_i \geq 0$, the above can only be true if there exists $k \in J \setminus \{i_2\}$ with $\langle \nu_k, \nu_{i_1}\rangle > 0$. According to the definition of J, we can find $x \in Z \cap \operatorname{span}(\{\nu_1, ..., \nu_n\}) \setminus \{0\}$ for this k such that $\langle \nu_{i_1}, x\rangle = \langle \nu_k, x\rangle = 0$. Note that there exists $j \in I$ such that $\langle \nu_j, x\rangle < 0$ for this x. Indeed, if this was not the case, then $\langle \nu_i, x\rangle = 0$ for all $i \in I$ and $x \in \operatorname{span}(\{\nu_1, ..., \nu_n\})$ would imply $x = 0$ and yield a contradiction. Consider now the cone $\mathcal{Z}(x) = \{z \in \mathbb{R}^d \mid \langle \nu_i, z\rangle \leq 0 \ \forall i \in \mathcal{A}(x)\}$. By Lemma 2.9, $\mathcal{Z}(x)$ is full-dimensional and $\{(\nu_i, 0)\}_{i \in \mathcal{A}(x)}$ is a standard description of $\mathcal{Z}(x)$. Further, we have $2 \leq |\mathcal{A}(x)| \leq |I| - 1$ by construction and $\mathcal{Z}(x)$ is not non-obtuse since $i_1, k \in \mathcal{A}(x)$. By our induction hypothesis, this yields that there exists $\tilde{w} \in \mathcal{Z}(x)$ such that $\{i \in \mathcal{A}(x) \mid \langle \nu_i, \tilde{w}\rangle = 0\} = \{j_1, j_2\}$

with ν_{j_1}, ν_{j_2} satisfying $\langle \nu_{j_1}, \nu_{j_2} \rangle > 0$. Define $w_s := x + s\tilde{w}$, $s > 0$. Then

$$\langle \nu_i, w_s \rangle = \begin{cases} \langle \nu_i, s\tilde{w} \rangle = 0 & \text{if } i \in \{j_1, j_2\}, \\ \langle \nu_i, s\tilde{w} \rangle < 0 & \text{if } i \in \mathcal{A}(x) \setminus \{j_1, j_2\}, \\ \langle \nu_i, x \rangle + s\langle \nu_i, \tilde{w} \rangle & \text{if } i \in \mathcal{I}(x), \end{cases}$$

holds for all $s > 0$. By choosing $s > 0$ small enough, we achieve that $\langle \nu_i, w_s \rangle < 0$ for all $i \in \mathcal{I}(x)$. The resulting $w_s \in Z$ then has all of the desired properties. This completes the proof. □

Next, we extend Lemma 2.15 to arbitrary polyhedra that are not non-obtuse.

Proposition 2.16 (existence of an obtuse ridge in the general case)**.** *Let $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a full-dimensional convex polyhedron that is not non-obtuse. Let $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i \in I}$ be a standard description of Z. Then there exists $w \in Z$ such that $\mathcal{A}(w) = \{i_1, i_2\}$ holds for some $i_1, i_2 \in I$ satisfying $\langle \nu_{i_1}, \nu_{i_2} \rangle > 0$.*

Proof. As Z is not non-obtuse, there exist $k, l \in I$ with $\langle \nu_k, \nu_l \rangle > 0$. From Proposition 2.6, we obtain that, for these k, l, we can find $x_k, x_l \in Z$ satisfying $\langle \nu_k, x_k \rangle = \alpha_k$, $\langle \nu_i, x_k \rangle < \alpha_i$ $\forall i \neq k$, $\langle \nu_l, x_l \rangle = \alpha_l$, $\langle \nu_i, x_l \rangle < \alpha_i$ $\forall i \neq l$. Define

$$\tilde{x} := x_k - \frac{\langle \nu_l, x_k \rangle - \alpha_l}{\langle \nu_l, \nu_k \rangle} \nu_k \in \mathbb{R}^d, \qquad h := \tilde{x} - x_l, \quad \text{and} \quad \gamma\colon [0,1] \to \mathbb{R}^d,\ \gamma(s) := x_l + sh.$$

Then we have $\gamma(0) = x_l$ and $\pi_Z(\gamma(1)) = \pi_Z(\tilde{x}) = x_k$, it holds

$$\langle \nu_l, \gamma(s) \rangle = (1-s)\langle \nu_l, x_l \rangle + s\langle \nu_l, \tilde{x} \rangle = (1-s)\alpha_l + s\left\langle \nu_l, x_k - \frac{\langle \nu_l, x_k \rangle - \alpha_l}{\langle \nu_l, \nu_k \rangle}\nu_k \right\rangle = \alpha_l \quad \forall s \in [0,1], \tag{2.20}$$

and there exists $\varepsilon \in (0,1)$ such that

$$\pi_Z(\gamma(s)) = \gamma(s) \qquad \forall s \in [0, \varepsilon]. \tag{2.21}$$

Define

$$\bar{s} := \sup\{s \in [0,1] \mid \langle \nu_l, \pi_Z(\gamma(r)) \rangle = \alpha_l\ \forall r \in [0,s]\}.$$

Then (2.20), (2.21), $\langle \nu_l, \pi_Z(\gamma(1)) \rangle = \langle \nu_l, x_k \rangle < \alpha_l$, and the continuity of π_Z imply that $0 < \varepsilon \leq \bar{s} < 1$. Consider now the point $y := \pi_Z(\gamma(\bar{s}))$. Then the definition of $\bar{s}$ implies that there exists a sequence $\{s_m\} \subset (\bar{s}, 1]$ such that $s_m \to \bar{s}$ holds and $\langle \nu_l, \pi_Z(\gamma(s_m)) \rangle < \alpha_l = \langle \nu_l, y \rangle$ for all $m \in \mathbb{N}$. In combination with Lemma 2.10ii), it follows that, for all sufficiently large m, we have

$$\langle \nu_l, \pi_Z(\gamma(s_m)) \rangle = \langle \nu_l, y + \pi_{\mathcal{Z}(y)}(\gamma(s_m) - y) \rangle = \alpha_l + \langle \nu_l, \pi_{\mathcal{Z}(y)}(\gamma(s_m) - y) \rangle < \alpha_l$$

and, thus, $\langle \nu_l, \pi_{\mathcal{Z}(y)}(\gamma(s_m) - y) \rangle < 0$. Due to (2.20) and the definition of y, however, we have $l \in \mathcal{A}(y)$ and $\langle \nu_l, \gamma(s_m) - y \rangle = \alpha_l - \alpha_l = 0$. In view of Lemma 2.14ii), this shows that $\mathcal{Z}(y)$ is not non-obtuse. (Note that $\mathcal{Z}(y)$ is full-dimensional and that $\{(\nu_i, 0)\}_{i \in \mathcal{A}(y)}$ is a standard description of $\mathcal{Z}(y)$ by Lemma 2.9.) Thus, Lemma 2.15 can be applied to the polyhedral cone $\mathcal{Z}(y)$ and it follows that there exists $\tilde{w} \in \mathcal{Z}(y)$ such that $\{i \in \mathcal{A}(y) \mid \langle \nu_i, \tilde{w} \rangle = 0\} = \{i_1, i_2\}$ holds with ν_{i_1}, ν_{i_2} satisfying $\langle \nu_{i_1}, \nu_{i_2} \rangle > 0$. Define $w_s := y + s\tilde{w}$, $s > 0$. Then it follows analogously to the last step in the proof of Lemma 2.15 that, for all small enough $s > 0$, we have

$$\langle \nu_i, w_s \rangle = \langle \nu_i, y \rangle + s\langle \nu_i, \tilde{w} \rangle \begin{cases} = \alpha_i + 0 & \text{if } i \in \{i_1, i_2\}, \\ < \alpha_i & \text{if } i \in \mathcal{A}(y) \setminus \{i_1, i_2\}, \\ < \alpha_i & \text{if } i \in \mathcal{I}(y). \end{cases}$$

The above shows that we can choose $w := w_s$ for a small enough $s > 0$ to obtain a point $w \in Z$ with all of the desired properties. This completes the proof. □

Corollary 2.17 (equivalence of local and global non-obtuseness)**.** *Let $Z \subset \mathbb{R}^d$ be a full-dimensional convex polyhedron and let $\{(\nu_i, \alpha_i)\}_{i \in I}$ be a standard description of Z. Then the following statements are equivalent.*

i) Z is non-obtuse, that is, $\langle \nu_i, \nu_j \rangle \leq 0$ for all $i, j \in I$ with $i \neq j$.
ii) Z has only non-obtuse ridges, that is, for all $x \in Z$ with $\mathcal{A}(x) = \{i_1, i_2\}$, it holds $\langle \nu_{i_1}, \nu_{i_2} \rangle \leq 0$.

2.4. Preliminaries on the Kurzweil-Stieltjes integral

A further important ingredient of our analysis is Kurzweil-Stieltjes integration theory. Let $a, b \in \mathbb{R}$, $a < b$, be given and fixed. We denote the Kurzweil-Stieltjes integral of an *integrand* $f\colon [a,b] \to \mathbb{R}$ w.r.t. an *integrator* $g\colon [a,b] \to \mathbb{R}$ by

$$\int_a^b f \,\mathrm{d}g \qquad \text{or} \qquad \int_a^b f(s) \,\mathrm{d}g(s).$$

For an in-depth introduction to the theory of Kurzweil-Stieltjes integration for scalar functions, we refer the reader to [52], Chapter 6. Let us now additionally assume that $d \in \mathbb{N}$ is given and fixed. If $f\colon [a,b] \to \mathbb{R}^d$ and $g\colon [a,b] \to \mathbb{R}^d$ are vector functions, then we define the Kurzweil-Stieltjes integral of f w.r.t. g *via*

$$\int_a^b \langle f, \mathrm{d}g \rangle := \sum_{j=1}^d \int_a^b f_j \,\mathrm{d}g_j, \tag{2.22}$$

provided that the scalar Kurzweil-Stieltjes integrals on the right-hand side of this definition all exist. Due to its componentwise definition, the vectorial version of the Kurzweil-Stieltjes integral in (2.22) inherits essentially all properties from its scalar counterpart studied in [52]. We briefly recall the most important facts and results in the following. For more details, see [52], Chapter 6.

Lemma 2.18 (existence of the integral)**.** *The Kurzweil-Stieltjes integral in (2.22) exists if $f, g \in G([a,b];\mathbb{R}^d)$ holds and at least one of the functions f and g is an element of $BV([a,b];\mathbb{R}^d)$. In this case, it yields a real number. If $f \in C([a,b];\mathbb{R}^d)$ and $g \in BV([a,b];\mathbb{R}^d)$, then the Kurzweil-Stieltjes integral coincides with the Riemann-Stieltjes integral. In particular, we have*

$$\int_a^b \langle f, \mathrm{d}g \rangle = \int_a^b \langle f(s), \dot{g}(s) \rangle \mathrm{d}s \qquad \forall f \in C([a,b];\mathbb{R}^d) \quad \forall g \in W^{1,1}((a,b);\mathbb{R}^d).$$

Proof. The assertions follow from [52], Theorems 5.6.1, 6.2.12, 6.3.11, Corollary 5.4.4 and (2.22). □

Lemma 2.19 (linearity and additivity of the integral)**.** *The Kurzweil-Stieltjes integral in (2.22) is linear in both its integrand and its integrator. Further, for all $c \in (a,b)$ and $f, g \in G([a,b];\mathbb{R}^d)$ with $f \in BV([a,b];\mathbb{R}^d)$ or*

$g \in BV([a,b];\mathbb{R}^d)$, *we have*

$$\int_a^b \langle f, \mathrm{d}g\rangle = \int_a^c \langle f, \mathrm{d}g\rangle + \int_c^b \langle f, \mathrm{d}g\rangle.$$

Proof. This follows from [52], Theorems 6.2.7, 6.2.9, 6.2.10 and (2.22). ☐

Lemma 2.20 (constant integrators and integrands)**.** *If $f, g \in G([a,b];\mathbb{R}^d)$ holds and $c \in \mathbb{R}^d$ is interpreted as a constant function, then we have*

$$\int_a^b \langle c, \mathrm{d}g\rangle = \langle c, g(b) - g(a)\rangle \qquad \text{and} \qquad \int_a^b \langle f, \mathrm{d}c\rangle = 0.$$

Proof. The assertions of this lemma follow from [52], Remark 6.3.1 and (2.22). ☐

Lemma 2.21 (singular integrators and integrands)**.** *Let $f, g \in G([a,b];\mathbb{R}^d)$ and $c \in \mathbb{R}^d$ be given. Then it holds*

$$\int_a^b \langle f(s), \mathrm{d}(c\mathbb{1}_{\{t\}}(s))\rangle = \begin{cases} -\langle f(a), c\rangle & \text{if } t = a, \\ 0 & \text{if } a < t < b, \\ \langle f(b), c\rangle & \text{if } t = b, \end{cases}$$

and

$$\int_a^b \langle c\mathbb{1}_{\{t\}}(s), \mathrm{d}g(s)\rangle = \begin{cases} \langle c, g(a+) - g(a)\rangle & \text{if } t = a, \\ \langle c, g(t+) - g(t-)\rangle & \text{if } a < t < b, \\ \langle c, g(b) - g(b-)\rangle & \text{if } t = b. \end{cases}$$

Proof. The assertions of the lemma follow from [52], Lemmas 6.3.2, 6.3.3 and (2.22). ☐

Lemma 2.22 (integrator-integrand estimate)**.** *For all $f \in G([a,b];\mathbb{R}^d)$ and $g \in BV([a,b];\mathbb{R}^d)$, it holds*

$$\left|\int_a^b \langle f, \mathrm{d}g\rangle\right| \le d\, \|f\|_\infty \operatorname{var}(g;[a,b]).$$

Proof. Using (2.22), [52], Theorem 6.3.6, and our definitions of $\|\cdot\|_\infty$ and $\operatorname{var}(\cdot;[a,b])$, we may compute that

$$\left|\int_a^b \langle f, \mathrm{d}g\rangle\right| \le \sum_{j=1}^d \left|\int_a^b f_j\, \mathrm{d}g_j\right| \le \sum_{j=1}^d \|f_j\|_\infty \operatorname{var}(g_j;[a,b]) \le d\, \|f\|_\infty \operatorname{var}(g;[a,b]).$$

This establishes the desired inequality. ☐

Lemma 2.23 (limit transitions with uniform convergence)**.** *Let $\{g_k\} \subset BV([a,b];\mathbb{R}^d)$ and $\{f_k\} \subset G([a,b];\mathbb{R}^d)$ be sequences satisfying $\sup_{k\in\mathbb{N}} \operatorname{var}(g_k;[a,b]) < \infty$, $\lim_{k\to\infty} \|g_k - g\|_\infty = 0$, and $\lim_{k\to\infty} \|f_k - f\|_\infty = 0$ for some $g \in BV([a,b];\mathbb{R}^d)$ and $f \in G([a,b];\mathbb{R}^d)$. Then it holds*

$$\lim_{k\to\infty} \int_a^b \langle f_k, \mathrm{d}g_k\rangle = \int_a^b \langle f, \mathrm{d}g\rangle.$$

Proof. The assertion of this lemma follows immediately from [52], Theorem 6.8.8 and definition (2.22). ☐

Theorem 2.24 (bounded convergence theorem)**.** *Let $g \in BV([a,b];\mathbb{R}^d)$ be given and let $\{f_k\} \subset G([a,b];\mathbb{R}^d)$ be a sequence that satisfies $\lim_{k\to\infty} f_k(s) = f(s)$ for all $s \in [a,b]$ and $\sup_{k\in\mathbb{N}} \|f_k\|_\infty < \infty$ for some $f \in G([a,b];\mathbb{R}^d)$. Then it holds*

$$\lim_{k\to\infty} \int_a^b \langle f_k, \mathrm{d}g\rangle = \int_a^b \langle f, \mathrm{d}g\rangle.$$

Proof. This result follows from [52], Theorem 6.8.13 and again the componentwise definition in (2.22). □

Lemma 2.25 (integration of limits)**.** *Let $f \in G([a,b];\mathbb{R}^d)$ and $g \in CBV([a,b];\mathbb{R}^d)$ be given. Then it holds*

$$\int_a^b \langle f, \mathrm{d}g\rangle = \int_a^b \langle f_+, \mathrm{d}g\rangle = \int_a^b \langle f_-, \mathrm{d}g\rangle.$$

Proof. The assertion of the lemma follows straightforwardly from Theorem 2.24, Lemmas 2.19 and 2.21, and the fact that regulated functions possess at most countably many points of discontinuity. □

Lemma 2.26 (integration over subintervals)**.** *Let $f \in G([a,b];\mathbb{R}^d)$, $g \in BV([a,b];\mathbb{R}^d)$, and $a \le s < t \le b$ be given. Suppose that s and t are points of continuity of g. Let $J \in \{(s,t), (s,t], [s,t), [s,t]\}$. Then it holds*

$$\int_a^b \langle \mathbb{1}_J f, \mathrm{d}g\rangle = \int_s^t \langle f, \mathrm{d}g\rangle = \int_s^t \langle f, \mathrm{d}g_+\rangle = \int_s^t \langle f, \mathrm{d}g_-\rangle. \tag{2.23}$$

Proof. The first equality in (2.23) follows from [52], Theorem 6.9.7 and (2.22). The second and third one are easy consequences of Lemmas 2.19, 2.21, and 2.23 and the fact that functions of bounded variation possess at most countably many points of discontinuity. □

Lemma 2.27 (integration by parts formula)**.** *Let $f, g \in G([a,b];\mathbb{R}^d)$ be given and suppose that at least one of the functions f and g is an element of $BV([a,b];\mathbb{R}^d)$. Then it holds*

$$\begin{aligned}\int_a^b \langle f, \mathrm{d}g\rangle = \langle f(b), g(b)\rangle - \langle f(a), g(a)\rangle - \int_a^b \langle g, \mathrm{d}f\rangle \\ + \sum_{a<s\le b} \langle f(s) - f(s-), g(s) - g(s-)\rangle - \sum_{a\le s<b} \langle f(s+) - f(s), g(s+) - g(s)\rangle.\end{aligned}$$

Proof. This result is a straightforward consequence of [52], Theorem 6.4.2 and (2.22). □

The following lemma is a useful tool in Section 3.

Lemma 2.28 (criterion for nonnegativity)**.** *Let $f \in G([a,b])$ and $g \in CBV([a,b])$ be given. Suppose that g is nondecreasing on $[a,b]$ and that the following is true:*

$$\forall s \in [a,b] \text{ with } f(s) < 0\ \exists \varepsilon_s > 0 : g = \mathrm{const} \text{ on } (s-\varepsilon_s, s+\varepsilon_s) \cap [a,b]. \tag{2.24}$$

Then it holds

$$\int_a^b f\,\mathrm{d}g \ge 0.$$

Proof. Let f and g be as in the assertion of the lemma. Choose for every $s \in [a,b]$ satisfying $f(s) < 0$ a number $\varepsilon_s \in (0,1)$ as in (2.24) and define $U := \bigcup_{s\in[a,b], f(s)<0} (s-\varepsilon_s, s+\varepsilon_s) \subset \mathbb{R}$. Then U is an open and bounded set and possesses at most countably many (necessarily open and disjoint) connected components. Further, all of these connected components are nonempty open intervals. We denote the components of U with $U_m = (\alpha_m, \beta_m) \subset \mathbb{R}$, $-\infty < \alpha_m < \beta_m < \infty$, and the associated finite/countable index set with M. Due to (2.24) and $g \in CBV([a,b])$, g is constant on $[\alpha_m, \beta_m] \cap [a,b]$ for all $m \in M$. Further, we have by construction that $\min(0,f) \in G([a,b])$ is zero on $[a,b] \setminus U$. This implies that

$$\min(0,f) = \min(0,f)\mathbb{1}_U = \sum_{m \in M} \min(0,f)\mathbb{1}_{U_m} \tag{2.25}$$

holds in $[a,b]$ and, due to Lemmas 2.20 and 2.26, that

$$\int_a^b \mathbb{1}_{U_m} \min(0,f)\mathrm{d}g = \int_{\max(a,\alpha_m)}^{\min(b,\beta_m)} \min(0,f)\mathrm{d}g = 0 \qquad \forall m \in M. \tag{2.26}$$

In combination with Lemma 2.19 and Theorem 2.24, the equations (2.25) and (2.26) yield

$$\int_a^b \min(0,f)\mathrm{d}g = \int_a^b \sum_{m \in M} \min(0,f)\mathbb{1}_{U_m}\mathrm{d}g = \sum_{m \in M} \int_a^b \min(0,f)\mathbb{1}_{U_m}\mathrm{d}g = 0.$$

As the definition of the Kurzweil-Stieltjes integral implies that the integral of a nonnegative integrand w.r.t. a nondecreasing integrator is nonnegative, see [52], Chapter 6, we may now conclude that

$$\int_a^b f\,\mathrm{d}g = \int_a^b \max(0,f)\mathrm{d}g + \int_a^b \min(0,f)\mathrm{d}g = \int_a^b \max(0,f)\mathrm{d}g \geq 0.$$

This completes the proof. □

Next, we recall Helly's selection theorem, which plays an important role in Section 3.

Theorem 2.29 (Helly selection theorem)**.** *Suppose that $\{g_k\}$ is a bounded sequence in $BV([a,b];\mathbb{R}^d)$. Then there exist a subsequence $\{g_{k_l}\}$ and a function $g \in BV([a,b];\mathbb{R}^d)$ such that*

$$\lim_{l \to \infty} g_{k_l}(s) = g(s) \quad \forall s \in [a,b] \qquad \textit{and} \qquad \operatorname{var}(g;[a,b]) \leq \sup_{k \in \mathbb{N}} \operatorname{var}(g_k;[a,b]).$$

Proof. The assertion of this theorem follows straightforwardly from [52], Theorem 2.7.4. □

We conclude this subsection by proving two convenient auxiliary results for $CBV([a,b];\mathbb{R}^d)$-functions.

Lemma 2.30 (mollification of continuous functions of bounded variation)**.** *Let $v \in CBV([a,b];\mathbb{R}^d)$ be given and let $\{\eta_\varepsilon\}_{\varepsilon>0} \subset C^\infty(\mathbb{R})$ be standard mollifiers as defined in [57], Section C.5. Define*

$$v_\varepsilon := (\eta_\varepsilon \star \hat{v})|_{[a,b]},\ \varepsilon > 0, \qquad \textit{where} \qquad \hat{v}(s) := \begin{cases} v(a) & \textit{if } s < a, \\ v(s) & \textit{if } s \in [a,b], \\ v(b) & \textit{if } s > b, \end{cases} \tag{2.27}$$

and where $\star$ denotes a (componentwise) convolution, i.e.,

$$\eta_\varepsilon \star \hat{v} \colon \mathbb{R} \to \mathbb{R}^d, \qquad (\eta_\varepsilon \star \hat{v})(t) := \int_{\mathbb{R}} \hat{v}(s)\eta_\varepsilon(t-s)\mathrm{d}s.$$

Then $v_\varepsilon \in C^\infty([a,b];\mathbb{R}^d)$ *holds for all* $\varepsilon > 0$ *and we have*

$$\|v - v_\varepsilon\|_{C([a,b];\mathbb{R}^d)} \to 0 \qquad \text{and} \qquad \operatorname{var}(v;[a,b]) \geq \operatorname{var}(v_\varepsilon;[a,b]) \to \operatorname{var}(v;[a,b]) \qquad \text{for } \varepsilon \to 0. \tag{2.28}$$

Proof. That $v_\varepsilon \in C^\infty([a,b];\mathbb{R}^d)$ holds for all $\varepsilon > 0$ and that $\|v - v_\varepsilon\|_{C([a,b];\mathbb{R}^d)} \to 0$ for $\varepsilon \to 0$ follows from [57], Theorem C.5-7(i),(iii). To see that $\operatorname{var}(v_\varepsilon;[a,b]) \leq \operatorname{var}(v;[a,b])$ holds for all $\varepsilon > 0$, suppose that a partition Δ of the form $a = t_0 \leq t_1 \leq ... \leq t_N = b$, $N \in \mathbb{N}$, of $[a,b]$ is given. The definition of v_ε implies

$$\begin{aligned}\operatorname{var}_\Delta(v_\varepsilon;[a,b]) &:= \sum_{j=1}^N |v_\varepsilon(t_j) - v_\varepsilon(t_{j-1})| = \sum_{j=1}^N \left| \int_{\mathbb{R}} \eta_\varepsilon(s)\hat{v}(t_j - s) - \eta_\varepsilon(s)\hat{v}(t_{j-1} - s)\mathrm{d}s \right| \\ &\leq \int_{\mathbb{R}} \eta_\varepsilon(s) \sum_{j=1}^N |\hat{v}(t_j - s) - \hat{v}(t_{j-1} - s)|\,\mathrm{d}s \leq \int_{\mathbb{R}} \eta_\varepsilon(s)\mathrm{d}s\,\operatorname{var}(v;[a,b]) = \operatorname{var}(v;[a,b]),\end{aligned} \tag{2.29}$$

where we used that $\hat{v}$ is constant on $(-\infty, a]$ and $[b, \infty)$, respectively. Taking the supremum over all Δ in the above yields $\operatorname{var}(v_\varepsilon;[a,b]) \leq \operatorname{var}(v;[a,b])$ for all $\varepsilon > 0$ as desired. It remains to prove that $\operatorname{var}(v_\varepsilon;[a,b])$ converges to $\operatorname{var}(v;[a,b])$ for $\varepsilon \to 0$. To this end, we note that $\operatorname{var}(\cdot;[a,b])\colon C([a,b];\mathbb{R}^d) \to [0,\infty]$ is, by definition, the pointwise supremum of the family of functions $\{\operatorname{var}_\Delta(\cdot;[a,b])\colon C([a,b];\mathbb{R}^d) \to [0,\infty) \mid \Delta \text{ is a partition of } [a,b]\}$. Here, $\operatorname{var}_\Delta(\cdot;[a,b])$ is defined as in (2.29). As the functions $\operatorname{var}_\Delta(\cdot;[a,b])\colon C([a,b];\mathbb{R}^d) \to [0,\infty)$ are continuous, this implies that $\operatorname{var}(\cdot;[a,b])\colon C([a,b];\mathbb{R}^d) \to [0,\infty]$ is a lower semicontinuous function. Consequently,

$$\operatorname{var}(v;[a,b]) \geq \limsup_{\varepsilon \to 0} \operatorname{var}(v_\varepsilon;[a,b]) \geq \liminf_{\varepsilon \to 0} \operatorname{var}(v_\varepsilon;[a,b]) \geq \operatorname{var}(v;[a,b]).$$

This completes the proof. ☐

Lemma 2.31 (fundamental theorem in $CBV([a,b];\mathbb{R}^d)$)**.** *Let* $v \in CBV([a,b];\mathbb{R}^d)$ *be given and suppose that* $F\colon \mathbb{R}^d \to \mathbb{R}$ *is continuously differentiable. Then it holds*

$$F(v(b)) - F(v(a)) = \int_a^b \langle \nabla F(v(s)), \mathrm{d}v(s)\rangle. \tag{2.30}$$

Proof. Let $\{v_\varepsilon\}$ be as in Lemma 2.30. Then, for each $\varepsilon > 0$, we have

$$F(v_\varepsilon(b)) - F(v_\varepsilon(a)) = \int_a^b \langle \nabla F(v_\varepsilon(s)), \dot{v}_\varepsilon(s)\rangle \mathrm{d}s = \int_a^b \langle \nabla F(v_\varepsilon(s)), \mathrm{d}v_\varepsilon(s)\rangle, \tag{2.31}$$

where the second identity follows from Lemma 2.18. Since $\|v - v_\varepsilon\|_{C([a,b];\mathbb{R}^d)} \to 0$ holds and since F is C^1, we have $\|\nabla F(v) - \nabla F(v_\varepsilon)\|_{C([a,b];\mathbb{R}^d)} \to 0$. Further, we know that $\operatorname{var}(v_\varepsilon;[a,b]) \leq \operatorname{var}(v;[a,b])$ for all $\varepsilon > 0$. In combination with Lemma 2.23, this allows us to pass to the limit $\varepsilon \to 0$ in (2.31) to arrive at (2.30). ☐

2.5. Preliminaries on the vector play and stop

As a final preparation for our analysis, we recall some results on the vector play and stop. Throughout this subsection, we make the following assumptions:

Assumption 2.32 (standing assumptions for Section 2.5)**.**

i) $d \in \mathbb{N}$ and $T > 0$ are given and fixed;
ii) $Z \subset \mathbb{R}^d$ is a full-dimensional convex polyhedron;
iii) $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i \in I}$ is a standard description of Z.

Proposition 2.33 (unique solvability of the EVI). *The evolution variational inequality*

$$\begin{aligned} &y \in CBV([0,T];\mathbb{R}^d), \qquad y(t) \in Z \quad \forall t \in [0,T], \qquad y(0) = y_0,\\ &\int_0^T \langle v - y, \mathrm{d}(y-u)\rangle \geq 0 \quad \forall v \in C([0,T];Z), \end{aligned} \tag{V}$$

has a unique solution $y \in CBV([0,T];\mathbb{R}^d)$ *for all* $(u, y_0) \in CBV([0,T];\mathbb{R}^d) \times Z$. *If* $u \in W^{1,1}((0,T);\mathbb{R}^d)$ *holds, then the solution* y *of* (V) *is also the unique solution of the problem*

$$\begin{aligned} &y \in W^{1,1}((0,T);\mathbb{R}^d), \qquad y(t) \in Z \quad \forall t \in [0,T], \qquad y(0) = y_0,\\ &\langle \dot y(t) - \dot u(t), v - y(t)\rangle \geq 0 \quad \forall v \in Z \quad \text{for a.a. } t \in (0,T). \end{aligned} \tag{2.32}$$

Proof. The unique solvability of (V) follows from [4], Theorem 4.1, and that the solution of (V) is uniquely characterized by (2.32) in the case $u \in W^{1,1}((0,T);\mathbb{R}^d)$ follows from [4], Proposition 4.1. (Note that the additional assumption $0 \in Z$ in [4] is unimportant since we can always satisfy it by means of a translation.) ☐

Lemma 2.34 (regulated test functions). *In the situation of Proposition 2.33, the solution* y *of* (V) *also satisfies*

$$\int_{s_1}^{s_2} \langle v - y, \mathrm{d}(y-u)\rangle \geq 0 \qquad \forall v \in G([s_1,s_2];Z) \qquad \forall 0 \leq s_1 < s_2 \leq T. \tag{2.33}$$

Proof. This lemma can be proved completely analogously to [48], Lemma 4.2; see also [4], Lemma 4.1. ☐

Definition 2.35 (play and stop operator). The solution operator $\mathcal{S}\colon CBV([0,T];\mathbb{R}^d) \times Z \to CBV([0,T];\mathbb{R}^d)$, $(u, y_0) \mapsto y$, of (V) is called the stop operator. The function $\mathcal{P}\colon CBV([0,T];\mathbb{R}^d) \times Z \to CBV([0,T];\mathbb{R}^d)$ defined by $\mathcal{P}(u,y_0) := u - \mathcal{S}(u,y_0)$ is called the play operator.

Note that, due to the above definitions, continuity and differentiability results for $\mathcal{S}$ carry over to $\mathcal{P}$ and vice versa. This allows us to focus on $\mathcal{S}$ in the remainder of this paper. Next, we study Lipschitz continuity properties of $\mathcal{S}$. For $W^{1,1}((0,T);\mathbb{R}^d)$-inputs u, a Lipschitz estimate for $\mathcal{S}$ with values in $W^{1,1}((0,T);\mathbb{R}^d)$ has been derived in [4], Theorems 7.1, 7.2. We extend this result to $CBV([0,T];\mathbb{R}^d)$ with the following lemma:

Lemma 2.36 (extension of Lipschitz estimates). *Let* X *be a normed space, let* $D \subset X$ *be a set, and let* $\mathcal{F}\colon CBV([0,T];\mathbb{R}^d) \times D \to CBV([0,T];\mathbb{R}^d)$ *be given. Suppose that there exist constants* $L_1, L_2 > 0$ *satisfying*

$$\begin{aligned} &\|\mathcal{F}(u_1,x_1) - \mathcal{F}(u_2,x_2)\|_{C([0,T];\mathbb{R}^d)}\\ &\quad \leq L_1\left(\|u_1-u_2\|_{C([0,T];\mathbb{R}^d)} + \|x_1-x_2\|_X\right) \qquad \forall (u_1,x_1),(u_2,x_2) \in CBV([0,T];\mathbb{R}^d) \times D \end{aligned} \tag{2.34}$$

and

$$\begin{aligned} &\mathrm{var}\left(\mathcal{F}(u_1,x_1) - \mathcal{F}(u_2,x_2); [0,T]\right)\\ &\quad \leq L_2\left(\|\dot u_1-\dot u_2\|_{L^1((0,T);\mathbb{R}^d)} + \|x_1-x_2\|_X\right) \qquad \forall (u_1,x_1),(u_2,x_2) \in C^\infty([0,T];\mathbb{R}^d) \times D. \end{aligned} \tag{2.35}$$

Then it holds

$$\begin{aligned} &\|\mathcal{F}(u_1,x_1) - \mathcal{F}(u_2,x_2)\|_{CBV([0,T];\mathbb{R}^d)}\\ &\quad \leq (L_1+L_2)\left(\|u_1-u_2\|_{CBV([0,T];\mathbb{R}^d)} + \|x_1-x_2\|_X\right) \qquad \forall (u_1,x_1),(u_2,x_2) \in CBV([0,T];\mathbb{R}^d) \times D. \end{aligned}$$

Proof. Let $(u_1, x_1), (u_2, x_2) \in CBV([0, T]; \mathbb{R}^d) \times D$ be given, and let $u_{1,\varepsilon}, u_{2,\varepsilon}, (u_1 - u_2)_\varepsilon \in C^\infty([0, T]; \mathbb{R}^d)$, $\varepsilon > 0$, be mollifications of u_1, u_2, and $u_1 - u_2$, respectively, as constructed in Lemma 2.30. Then it follows from (2.27), (2.28), (2.34), and the smoothness of $u_{1,\varepsilon} - u_{2,\varepsilon}$ that

$$\mathcal{F}(u_{1,\varepsilon}, x_1) - \mathcal{F}(u_{2,\varepsilon}, x_2) \to \mathcal{F}(u_1, x_1) - \mathcal{F}(u_2, x_2) \text{ for } \varepsilon \to 0 \text{ in } C([0, T]; \mathbb{R}^d) \tag{2.36}$$

and

$$\|\dot{u}_{1,\varepsilon} - \dot{u}_{2,\varepsilon}\|_{L^1((0,T);\mathbb{R}^d)} = \operatorname{var}(u_{1,\varepsilon} - u_{2,\varepsilon}; [0, T]) = \operatorname{var}((u_1 - u_2)_\varepsilon; [0, T]) \le \operatorname{var}(u_1 - u_2; [0, T]) \quad \forall \varepsilon > 0. \tag{2.37}$$

Since $C([0, T]; \mathbb{R}^d) \to [0, \infty]$, $v \mapsto \operatorname{var}(v; [0, T])$, is lower semicontinuous w.r.t. strong convergence in $C([0, T]; \mathbb{R}^d)$ (see the proof of Lem. 2.30), the results in (2.35), (2.36), and (2.37) imply

$$\begin{aligned}
\operatorname{var}(\mathcal{F}(u_1, x_1) - \mathcal{F}(u_2, x_2); [0, T]) &\le \liminf_{\varepsilon \to 0} \operatorname{var}(\mathcal{F}(u_{1,\varepsilon}, x_1) - \mathcal{F}(u_{2,\varepsilon}, x_2); [0, T]) \\
&\le L_2 \limsup_{\varepsilon \to 0} \left(\|\dot{u}_{1,\varepsilon} - \dot{u}_{2,\varepsilon}\|_{L^1((0,T);\mathbb{R}^d)} + \|x_1 - x_2\|_X \right) \\
&= L_2 \limsup_{\varepsilon \to 0} \left(\operatorname{var}((u_1 - u_2)_\varepsilon; [0, T]) + \|x_1 - x_2\|_X \right) \\
&\le L_2 \left(\operatorname{var}(u_1 - u_2; [0, T]) + \|x_1 - x_2\|_X \right).
\end{aligned}$$

By combining the above with (2.34), the assertion follows. ☐

For the stop operator $\mathcal{S}$, we now obtain:

Theorem 2.37 (CBV-Lipschitz continuity of $\mathcal{S}$)**.** *There exists a constant $L > 0$ depending only on Z such that*

$$\begin{aligned}
&\|\mathcal{S}(u, y_0) - \mathcal{S}(\tilde{u}, \tilde{y}_0)\|_{CBV([0,T];\mathbb{R}^d)} \\
&\qquad \le L \left(\|u - \tilde{u}\|_{CBV([0,T];\mathbb{R}^d)} + |y_0 - \tilde{y}_0| \right) \qquad \forall (u, y_0), (\tilde{u}, \tilde{y}_0) \in CBV([0, T]; \mathbb{R}^d) \times Z.
\end{aligned} \tag{2.38}$$

Proof. As in the proof of Lemma 2.9, we may assume w.l.o.g. that $0 \in \operatorname{int}(Z)$ holds and that $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i \in I}$ satisfies $\alpha_i > 0$ for all $i \in I$. For this situation, we obtain from [4], Theorems 7.1, 7.2 that there exist constants $L_1, L_2 > 0$ depending only on Z such that

$$\begin{aligned}
&\|\mathcal{S}(u, y_0) - \mathcal{S}(\tilde{u}, \tilde{y}_0)\|_{C([0,T];\mathbb{R}^d)} \\
&\qquad \le L_1 \left(\|u - \tilde{u}\|_{C([0,T];\mathbb{R}^d)} + |y_0 - \tilde{y}_0| \right) \qquad \forall (u, y_0), (\tilde{u}, \tilde{y}_0) \in CBV([0, T]; \mathbb{R}^d) \times Z
\end{aligned} \tag{2.39}$$

and

$$\begin{aligned}
&\left\| \dot{\mathcal{S}}(u, y_0) - \dot{\mathcal{S}}(\tilde{u}, \tilde{y}_0) \right\|_{L^1((0,T);\mathbb{R}^d)} \\
&\qquad \le L_2 \left(\left\| \dot{u} - \dot{\tilde{u}} \right\|_{L^1((0,T);\mathbb{R}^d)} + |y_0 - \tilde{y}_0| \right) \qquad \forall (u, y_0), (\tilde{u}, \tilde{y}_0) \in W^{1,1}((0, T); \mathbb{R}^d) \times Z.
\end{aligned}$$

By invoking Lemma 2.36, the assertion now follows immediately. ☐

Next, we establish a result on the local behavior of the play operator that generalizes the local monotonicity property documented for the scalar case in [48], Lemma 4.3. Recall that, for every point $x \in Z$, we have $\mathcal{A}(x) := \{i \in I \mid \langle \nu_i, x \rangle = \alpha_i\}$ and $\mathcal{Z}(x) := \{z \in \mathbb{R}^d \mid \langle \nu_i, z \rangle \le 0 \ \forall i \in \mathcal{A}(x)\}$.

Lemma 2.38 (local monotone decomposition). *Assume that $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i\in I}$ is regular. Let $t \in [0,T]$ be fixed and let $(u, y_0) \in CBV([0,T];\mathbb{R}^d) \times Z$ be given. Set $y := \mathcal{S}(u, y_0)$ and $w := \mathcal{P}(u, y_0)$. Then there exist a number $\varepsilon > 0$ and, for this ε, unique $\lambda_i \in CBV([t-\varepsilon, t+\varepsilon] \cap [0,T])$, $i \in \mathcal{A}(y(t))$, such that*

$$\mathcal{A}(y(s)) \subset \mathcal{A}(y(t)) \quad \text{and} \quad w(s) = w(\max(0, t-\varepsilon)) + \sum_{i\in\mathcal{A}(y(t))} \lambda_i(s)\nu_i \quad \forall s \in [t-\varepsilon, t+\varepsilon] \cap [0,T]. \tag{2.40}$$

Further, the functions λ_i, $i \in \mathcal{A}(y(t))$, are each nonnegative and nondecreasing on $[t-\varepsilon, t+\varepsilon] \cap [0,T]$.

Proof. Since $\mathcal{H}$ is regular and $y(t) \in Z$, we obtain from Lemma 2.10iv) that there exist unique $e_i \in \mathcal{Z}(y(t))$, $i \in \mathcal{A}(y(t))$, such that (2.12) and (2.13) hold for $x := y(t)$. From Lemma 2.10i) and the continuity of y, it follows further that there exist $\varepsilon, \tau > 0$ such that $\mathcal{A}(y(r)) \subset \mathcal{A}(y(t))$ and $y(r) + \tau z \in Z$ for all $r \in [t-\varepsilon, t+\varepsilon] \cap [0,T]$ and all $z \in B_1(0) \cap \mathcal{Z}(y(t))$. This allows us to consider test functions of the form $v\colon [t-\varepsilon, s] \cap [0,T] \to \mathbb{R}^d$, $r \mapsto y(r) + \tau z$, with arbitrary $z \in B_1(0) \cap \mathcal{Z}(y(t))$ and $s \in (t-\varepsilon, t+\varepsilon] \cap (0,T]$ in Lemma 2.34 and to obtain

$$\begin{aligned} 0 &\le \int_{\max(0,t-\varepsilon)}^{s} \langle (y(r) + \tau z) - y(r), \mathrm{d}(y(r) - u(r))\rangle = -\tau \int_{\max(0,t-\varepsilon)}^{s} \langle z, \mathrm{d}w(r)\rangle \\ &= \tau \,\langle z, w(\max(0,t-\varepsilon)) - w(s)\rangle \qquad \forall z \in B_1(0) \cap \mathcal{Z}(y(t)) \qquad \forall s \in (t-\varepsilon, t+\varepsilon] \cap (0,T] \end{aligned} \tag{2.41}$$

and, thus,

$$\langle z, w(\max(0,t-\varepsilon)) - w(s)\rangle \ge 0 \qquad \forall z \in \mathcal{Z}(y(t)) \qquad \forall s \in (t-\varepsilon, t+\varepsilon] \cap (0,T]. \tag{2.42}$$

Here, we have used Lemma 2.20. Note that (2.42) is also true for $s = \max(0, t-\varepsilon)$. Due to Lemma 2.11, the definitions of $\mathcal{Z}(y(t))$ and $\mathcal{Z}(y(t))^\circ$, and the regularity of $\mathcal{H}$, this allows us to conclude that, for each $s \in [t-\varepsilon, t+\varepsilon] \cap [0,T]$, there exist unique values $\lambda_i(s) \ge 0$, $i \in \mathcal{A}(y(t))$, such that

$$w(s) = w(\max(0,t-\varepsilon)) + \sum_{i\in\mathcal{A}(y(t))} \lambda_i(s)\nu_i \qquad \forall s \in [t-\varepsilon, t+\varepsilon] \cap [0,T]. \tag{2.43}$$

If we take the scalar product with e_j, $j \in \mathcal{A}(y(t))$, in (2.43), then it follows that

$$\lambda_j(s) = \frac{\langle e_j, w(s)\rangle - \langle e_j, w(\max(0,t-\varepsilon))\rangle}{\langle e_j, \nu_j\rangle} \qquad \forall s \in [t-\varepsilon, t+\varepsilon] \cap [0,T].$$

Thus, $\lambda_j \in CBV([t-\varepsilon, t+\varepsilon] \cap [0,T])$ for all $j \in \mathcal{A}(y(t))$ as claimed. It remains to prove that the functions λ_i are nondecreasing. To this end, we note that – completely analogously to (2.42) – we obtain that

$$\langle z, w(s_1) - w(s_2)\rangle \ge 0 \qquad \forall z \in \mathcal{Z}(y(t)) \qquad \forall \max(0, t-\varepsilon) \le s_1 < s_2 \le \min(T, t+\varepsilon).$$

Using (2.43) in the above yields

$$\left\langle z, \sum_{i\in\mathcal{A}(y(t))} (\lambda_i(s_1) - \lambda_i(s_2))\nu_i \right\rangle \ge 0 \qquad \forall z \in \mathcal{Z}(y(t)) \qquad \forall \max(0, t-\varepsilon) \le s_1 < s_2 \le \min(T, t+\varepsilon)$$

and, after choosing $z = e_j \in \mathcal{Z}(y(t))$,

$$\langle e_j, \nu_j\rangle\, (\lambda_j(s_1) - \lambda_j(s_2)) \ge 0 \qquad \forall \max(0, t-\varepsilon) \le s_1 < s_2 \le \min(T, t+\varepsilon).$$

Since $\langle e_j, \nu_j \rangle < 0$ by (2.12), it follows that $\lambda_j(s_1) \le \lambda_j(s_2)$ for all $\max(0, t-\varepsilon) \le s_1 < s_2 \le \min(T, t+\varepsilon)$ and all $j \in \mathcal{A}(y(t))$. This shows that the functions λ_j are indeed nondecreasing and completes the proof. □

Note that, if one changes ε in (2.40), then the functions λ_i are typically shifted by constant values due to the term $w(\max(0, t-\varepsilon))$. The multiplier maps λ_i in Lemma 2.38 are thus indeed only uniquely determined for fixed ε. The next lemma builds on Lemma 2.38 and is needed for the discussion of the notion of *criticality* in Section 3.2.

Lemma 2.39 (multiplier integrals and tangent vectors)**.** *Assume that $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i\in I}$ is regular. Let $t \in [0, T]$ and $(u, y_0) \in CBV([0,T]; \mathbb{R}^d) \times Z$ be given. Let y, w, ε, and λ_i, $i \in \mathcal{A}(y(t))$, be as in Lemma 2.38. Then, for all $\max(0, t-\varepsilon) \le s_1 < s_2 \le \min(T, t+\varepsilon)$ and all $z \in G([s_1, s_2]; \mathbb{R}^d)$ satisfying $z(r) \in \mathcal{Z}(y(r))$ for all $r \in [s_1, s_2]$, it holds*

$$\int_{s_1}^{s_2} \langle z(r), \nu_i \rangle \,\mathrm{d}\lambda_i(r) \le 0 \qquad \forall i \in \mathcal{A}(y(t)).$$

Proof. Let $e_i \in \mathcal{Z}(y(t))$, $i \in \mathcal{A}(y(t))$, again denote the vectors that satisfy (2.12) and (2.13) for $x := y(t)$. Due to Lemma 2.26, it suffices to prove the assertion of the lemma for the case $s_1 = \max(0, t-\varepsilon)$, $s_2 = \min(T, t+\varepsilon)$. Suppose that a function $z \in G([t-\varepsilon, t+\varepsilon] \cap [0,T]; \mathbb{R}^d)$ satisfying $z(r) \in \mathcal{Z}(y(r))$ for all $r \in [t-\varepsilon, t+\varepsilon] \cap [0,T]$ is given and fix $j \in \mathcal{A}(y(t))$. Assume further that $s \in [t-\varepsilon, t+\varepsilon] \cap [0,T]$ is a point such that $\langle z(s), \nu_j \rangle > 0$ holds. Then $z(s) \in \mathcal{Z}(y(s))$ implies $j \notin \mathcal{A}(y(s))$ and we obtain from the inclusion $\mathcal{A}(y(s)) \subset \mathcal{A}(y(t))$ in (2.40) and (2.12) that $\pm e_j \in \mathcal{Z}(y(s))$. Due to Lemma 2.10i) and the continuity of y, it follows that there exist $\tau, \gamma > 0$ such that $y(r) \pm \tau e_j \in Z$ holds for all $r \in [s-\gamma, s+\gamma] \cap [0,T]$. Analogously to (2.41), we may now deduce from Lemma 2.34, (2.40), (2.12), and Lemma 2.20 that

$$\begin{aligned}
0 &\le \int_{\max(0,t-\varepsilon,s-\gamma)}^{\min(T,t+\varepsilon,s+\gamma)} \langle (y(r) \pm \tau e_j) - y(r), \mathrm{d}(y(r) - u(r)) \rangle \\
&= -\tau \int_{\max(0,t-\varepsilon,s-\gamma)}^{\min(T,t+\varepsilon,s+\gamma)} \langle \pm e_j, \mathrm{d}w(r) \rangle \\
&= -\tau \int_{\max(0,t-\varepsilon,s-\gamma)}^{\min(T,t+\varepsilon,s+\gamma)} \langle \pm e_j, \nu_j \rangle \mathrm{d}\lambda_j(r) \\
&= \pm \langle e_j, \nu_j \rangle \tau \left(\lambda_j(\max(0, t-\varepsilon, s-\gamma)) - \lambda_j(\min(T, t+\varepsilon, s+\gamma)) \right).
\end{aligned}$$

As $\langle e_j, \nu_j \rangle \neq 0$ holds by (2.12) and since λ_j is nondecreasing by Lemma 2.38, it follows that $\lambda_j = \text{const}$ holds on $(s-\gamma, s+\gamma) \cap [\max(0, t-\varepsilon), \min(T, t+\varepsilon)]$. As $s \in [t-\varepsilon, t+\varepsilon] \cap [0,T]$ was an arbitrary point satisfying $\langle z(s), \nu_j \rangle > 0$, we may now invoke Lemma 2.28 (with $g := \lambda_j$, $f := -\langle z, \nu_j \rangle$, $a := \max(0, t-\varepsilon)$, and $b := \min(T, t+\varepsilon)$) to obtain that

$$\int_{\max(0,t-\varepsilon)}^{\min(T,t+\varepsilon)} \langle z(r), \nu_j \rangle \,\mathrm{d}\lambda_j(r) \le 0.$$

As $j \in \mathcal{A}(y(t))$ was arbitrary, this completes the proof. □

As an immediate consequence of Lemma 2.39, we obtain the following refined version of Lemma 2.34.

Lemma 2.40 (regulated tangential test functions)**.** *Assume that $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i\in I}$ is regular and let $(u, y_0) \in CBV([0,T]; \mathbb{R}^d) \times Z$ be given. Define $y := \mathcal{S}(u, y_0)$. Then it holds*

$$0 \le s_1 < s_2 \le T, \quad z \in G([s_1, s_2]; \mathbb{R}^d), \quad z(r) \in \mathcal{Z}(y(r)) \ \forall r \in [s_1, s_2] \qquad \Rightarrow \qquad \int_{s_1}^{s_2} \langle z, \mathrm{d}(y - u) \rangle \ge 0. \quad (2.44)$$

Proof. Due to Lemma 2.26, it suffices to consider the case $s_1 = 0$, $s_2 = T$. So let us assume that a function $z \in G([0,T];\mathbb{R}^d)$ satisfying $z(r) \in \mathcal{Z}(y(r))$ for all $r \in [0,T]$ is given. Choose for all $t \in [0,T]$ a number $\varepsilon_t > 0$ with the properties in Lemma 2.38. Since $[0,T]$ is compact, we can find t_m, $m = 1, ..., M$, $M \in \mathbb{N}$, such that the collection of intervals $(t_m - \varepsilon_{t_m}, t_m + \varepsilon_{t_m})$, $m = 1, ..., M$, covers $[0,T]$. Let $\{\psi_m\}_{m=1}^M$ be a smooth partition of unity on $[0,T]$ subordinate to this cover in the sense of [57], *i.e.*, a collection of functions satisfying

$$\begin{gathered}\psi_m \in C_c^\infty(\mathbb{R}), \quad 0 \le \psi_m(r) \le 1 \ \forall r \in \mathbb{R}, \quad \operatorname{supp}(\psi_m) \subset (t_m - \varepsilon_{t_m}, t_m + \varepsilon_{t_m}) \quad \forall m = 1, ..., M, \\ \sum_{m=1}^M \psi_m(r) = 1 \ \forall r \in [0,T].\end{gathered} \tag{2.45}$$

Define $z_m := \psi_m z \in G([0,T];\mathbb{R}^d)$, $m = 1, ..., M$. Then we clearly have

$$\int_0^T \langle z, \mathrm{d}(y-u)\rangle = -\int_0^T \langle z, \mathrm{d}w\rangle = -\sum_{m=1}^M \int_0^T \langle z_m, \mathrm{d}w\rangle\,, \tag{2.46}$$

where we again write $w := u - y = \mathcal{P}(u, y_0)$ for the play. Note that, for each of the integrals in the sum on the right-hand side of (2.46), we have

$$\int_0^T \langle z_m, \mathrm{d}w\rangle = \int_{\max(0,t_m-\varepsilon_{t_m})}^{\min(T,t_m+\varepsilon_{t_m})} \langle z_m, \mathrm{d}w\rangle = \sum_{i\in\mathcal{A}(y(t_m))} \int_{\max(0,t_m-\varepsilon_{t_m})}^{\min(T,t_m+\varepsilon_{t_m})} \langle z_m, \nu_i\rangle \,\mathrm{d}\lambda_i^m \tag{2.47}$$

by Lemma 2.26, where $\lambda_i^m \in CBV([t_m - \varepsilon_{t_m}, t_m + \varepsilon_{t_m}] \cap [0,T])$, $i \in \mathcal{A}(y(t_m))$, denote the functions from Lemma 2.38 associated with t_m. As the cone property of the linearization cone and the sign of ψ_m imply that $z_m(r) \in \mathcal{Z}(y(r))$ holds for all $r \in [t_m - \varepsilon_{t_m}, t_m + \varepsilon_{t_m}] \cap [0,T]$, we obtain from Lemma 2.39 that all of the integrals in the sum on the right-hand side of (2.47) are nonpositive. In combination with (2.46), this establishes the assertion of the lemma and completes the proof. □

Note that, while trivial in the $W^{1,1}((0,T);\mathbb{R}^d)$-case due to (2.32), the implication (2.44) is not obvious for $CBV([0,T];\mathbb{R}^d)$-solutions y since for this type of regularity one cannot argue in a pointwise manner and since for functions $z \in G([0,T];\mathbb{R}^d)$ satisfying $z(r) \in \mathcal{Z}(y(r))$ for all $r \in [0,T]$ there does not necessarily exist a number $\tau > 0$ satisfying $y(r) + \tau z(r) \in Z$ for all $r \in [0,T]$. In the proof of Lemma 2.40, we have overcome this problem by means of the decomposition in Lemma 2.38 and Lemma 2.28.

3. Proof of (I): directional differentiability in the non-obtuse case

With the necessary preliminaries in place, we can turn our attention to the main results of this paper, namely, points (I) and (II) of Theorem 1.1. The present section is devoted to the proof of (I), *i.e.*, the pointwise directional differentiability of the vectorial stop operator $\mathcal{S}\colon CBV([0,T];\mathbb{R}^d) \times Z \to CBV([0,T];\mathbb{R}^d)$ in the case of a full-dimensional non-obtuse convex polyhedral admissible set Z. The structure of this section is as follows: Section 3.1 introduces basic concepts and results that are needed for our analysis. This subsection also contains our standing assumptions and a rough overview of our method of proof. Section 3.2 is concerned with the notion of *criticality* which expresses that directional derivatives of $\mathcal{S}$ have to satisfy the inequality on the right-hand side of (2.44) with equality. In Sections 3.3 and 3.4, we identify how the directional derivatives jump at their points of discontinuity. At this point of our analysis, the non-obtuseness of the admissible set Z enters in a crucial manner; see the proofs of Lemmas 3.15, 3.19 and 3.21. Section 3.5 introduces the concept of *temporal polyhedricity* – an approximation property that is needed to derive the system (1.5) that uniquely characterizes the directional derivatives of $\mathcal{S}$. In Section 3.6, we finally put all of the pieces of our proof together to arrive at

(I) as desired; see Theorem 3.32. Section 3.7 concludes our discussion of the non-obtuse case by demonstrating that point (I) of Theorem 1.1 indeed contains the directional differentiability and characterization results proved for the case $d = 1$ and $Z = [-r, r]$, $r > 0$, in [15, 17] as a special case.

3.1. Standing assumptions and basic idea of the proof

In all of Section 3, we consider the following situation, unless explicitly stated otherwise.

Assumption 3.1 (standing assumptions for Sect. 3)**.**

i) $d \in \mathbb{N}$ and $T > 0$ are given and fixed;
ii) $Z \subset \mathbb{R}^d$ is a full-dimensional non-obtuse convex polyhedron;
iii) $\mathcal{H} = \{(\nu_i, \alpha_i)\}_{i \in I}$ is a standard description of Z;
iv) $y_0 \in Z$ is a given initial value and $u \in CBV([0,T];\mathbb{R}^d)$ is a given input function;
v) $y := \mathcal{S}(u, y_0)$ is the stop associated with (u, y_0) and $w := \mathcal{P}(u, y_0)$ is the play associated with (u, y_0);
vi) $h_0 \in \mathcal{Z}(y_0)$ is a given perturbation of y_0 and $h \in CBV([0,T];\mathbb{R}^d)$ is a given perturbation of u;
vii) $\tau_0 > 0$ is a number satisfying $y_0 + \tau_0 h_0 \in Z$.

Recall that, by Definition 2.12 and Propositions 2.6 and 2.13, the standard description $\mathcal{H}$ in Assumption 3.1 is unique up to permutations, regular, satisfies $\langle \nu_i, \nu_j \rangle \le 0$ for all $i, j \in I$ with $i \ne j$, and possesses the property (2.11). We introduce:

Definition 3.2 (difference quotients)**.** We denote by $\{\delta_\tau\} \subset CBV([0,T];\mathbb{R}^d)$ the family of difference quotients

$$\delta_\tau := \frac{\mathcal{S}(u + \tau h, y_0 + \tau h_0) - \mathcal{S}(u, y_0)}{\tau}, \qquad 0 < \tau < \tau_0. \tag{3.1}$$

As a direct consequence of Theorem 2.37, Lemma 2.34, and the EVI (V), we obtain:

Lemma 3.3 (EVI for the difference quotients)**.** *The family $\{\delta_\tau\}$ is bounded in $CBV([0,T];\mathbb{R}^d)$. Further, for every $\tau \in (0, \tau_0)$, δ_τ satisfies*

$$\begin{gathered} \delta_\tau(t) \in \frac{1}{\tau}\,(Z - y(t))\,\forall t \in [0,T], \qquad \delta_\tau(0) = h_0, \\ \int_{s_1}^{s_2} \langle z - \delta_\tau, \mathrm{d}(\delta_\tau - h)\rangle - \frac{1}{\tau}\int_{s_1}^{s_2} \langle z - \delta_\tau, \mathrm{d}w\rangle \ge 0 \\ \forall z \in G([s_1, s_2];\mathbb{R}^d) \colon z(t) \in \frac{1}{\tau}\,(Z - y(t))\ \ \forall t \in [s_1, s_2] \qquad \forall\, 0 \le s_1 < s_2 \le T, \end{gathered} \tag{3.2}$$

and

$$\int_{s_1}^{s_2} \langle \delta_\tau, \mathrm{d}w\rangle \le 0 \qquad \forall\, 0 \le s_1 < s_2 \le T. \tag{3.3}$$

Proof. Since Z is full-dimensional, Theorem 2.37 is applicable and the boundedness of $\{\delta_\tau\}$ in $CBV([0,T];\mathbb{R}^d)$ follows straightforwardly from (3.1). The same is true for the first line in (3.2). To obtain the variational inequality in (3.2), we note that (3.1) implies $\mathcal{S}(u + \tau h, y_0 + \tau h_0) = y + \tau\delta_\tau$ for all $0 < \tau < \tau_0$. By plugging this identity into the EVI (2.33) satisfied by $\mathcal{S}(u + \tau h, y_0 + \tau h_0)$, by choosing test functions of the form $v = y + \tau z$

with z as in (3.2), and by using that $w = \mathcal{P}(u, y_0) = u - \mathcal{S}(u, y_0) = u - y$, we obtain that

$$\begin{aligned} 0 &\leq \int_{s_1}^{s_2} \langle v - \mathcal{S}(u + \tau h, y_0 + \tau h_0), \mathrm{d}(\mathcal{S}(u + \tau h, y_0 + \tau h_0) - u - \tau h)\rangle \\ &= \tau^2 \int_{s_1}^{s_2} \langle z - \delta_\tau, \mathrm{d}(\delta_\tau - h)\rangle - \tau \int_{s_1}^{s_2} \langle z - \delta_\tau, \mathrm{d}w\rangle \end{aligned}$$

holds for all z, s_1, and s_2 as in (3.2). Dividing by τ^2 now yields the desired variational inequality. It remains to prove (3.3). To this end, we choose $v = \mathcal{S}(u + \tau h, y_0 + \tau h_0)$ in (2.33) for y. In combination with the definition of w, this gives

$$0 \leq \int_{s_1}^{s_2} \langle \mathcal{S}(u + \tau h, y_0 + \tau h_0) - y, \mathrm{d}(y - u)\rangle = -\tau \int_{s_1}^{s_2} \langle \delta_\tau, \mathrm{d}w\rangle \qquad \forall 0 \leq s_1 < s_2 \leq T.$$

The above establishes (3.3) and completes the proof. □

A central object of the analysis of this section is the following set:

Definition 3.4 (limits of difference quotients)**.** We define

$$\mathcal{D} := \left\{ \delta \in BV([0, T]; \mathbb{R}^d) \;\middle|\; \exists \{\tau_k\} \subset (0, \tau_0) \text{ s.t. } \tau_k \to 0 \text{ and } \delta_{\tau_k}(t) \to \delta(t)\ \forall t \in [0, T] \text{ for } k \to \infty \right\}. \tag{3.4}$$

Note that, from the boundedness of $\{\delta_\tau\}$ in $CBV([0, T]; \mathbb{R}^d)$ and Theorem 2.29, we obtain:

Lemma 3.5 (nontriviality of $\mathcal{D}$)**.** *The set $\mathcal{D}$ is nonempty.*

The main goal of the next five subsections is to prove that the set $\mathcal{D}$ is a singleton. If this is established, then we obtain from a trivial contradiction argument that $\mathcal{S}$ is directionally differentiable at (u, y_0) in direction (h, h_0) in a pointwise manner and point (I) of Theorem 1.1 follows; see Theorem 3.32. To show that $|\mathcal{D}| = 1$ holds, we will successively establish more and more properties of the elements of $\mathcal{D}$ until only one element remains. More precisely, the plan for Sections 3.2 to 3.6 is as follows:

- In Section 3.2, we use (3.2) and (3.3) to show that all $\delta \in \mathcal{D}$ satisfy the inequality on the right-hand side of (2.44) with equality. By means of Lemmas 2.38 and 2.39 and results from Section 2.4, we then transform the resulting notion of *criticality* to obtain information about the pointwise behavior of the left-/right-limit functions of the elements of $\mathcal{D}$; see Theorem 3.11.
- In Sections 3.3 and 3.4, we identify precisely how the elements of $\mathcal{D}$ jump at their points of discontinuity. We remark that this step of the proof is the most intricate one as the identification of the jump conditions has to be achieved solely based on the study of the difference quotients δ_τ which are elements of the space $CBV([0, T]; \mathbb{R}^d)$ and, thus, do not possess any jumps themselves. We overcome this difficulty by exploiting several auxiliary EVIs and vectorial Stampacchia truncation arguments that rely crucially on the non-obtuseness of Z and the properties in Lemma 2.14; see Theorems 3.16 and 3.23.
- In Section 3.5, we combine the information about the pointwise behavior of the elements of $\mathcal{D}$ obtained in Sections 3.2 to 3.4 to define a *critical cone* that contains all right-limit functions δ_+ of elements δ of the set $\mathcal{D}$. The main feature of this critical cone is that it possesses a *temporal polyhedricity* property. Roughly speaking, this property makes it possible to approximate critical directions with critical radial directions which are admissible in (3.2) and for which the passage to the limit $\tau \to 0$ can be performed in the EVI satisfied by δ_τ. For the main result on this topic, see Theorem 3.28.
- In Section 3.6, we carry out the limit transition $\tau \to 0$ in (3.2) to show that all elements δ of $\mathcal{D}$ are solutions of the system (1.5). Using the results from Section 2.4, we then show that this system can have at most one solution, thus proving that $\mathcal{D}$ is indeed a singleton. By putting all of the results of Sections 3.2 to 3.6 together, statement (I) of Theorem 1.1 then follows as desired.

We remark that the above argumentation roughly follows well-known approaches to the sensitivity analysis of elliptic systems; see [30, 31]. Due to the loss of regularity that occurs when passing to the limit $\tau \to 0$ with the difference quotients δ_τ (namely, the transition from $CBV([0,T];\mathbb{R}^d)$ to $BV([0,T];\mathbb{R}^d)$), however, our analysis is completely beyond the scope of standard results. In particular the techniques that we develop for the analysis of the pointwise properties of the elements of the set $\mathcal{D}$ seem to be entirely new.

3.2. Criticality

The first thing that we establish about the elements of the set $\mathcal{D}$ is that they are *critical*.

Definition 3.6 (criticality)**.** A function $z \in G([0,T];\mathbb{R}^d)$ is called a *critical direction* or simply *critical* if

$$z(t) \in \mathcal{Z}(y(t)) \quad \forall t \in [0,T] \qquad \text{and} \qquad \int_0^T \langle z, \mathrm{d}w \rangle = 0. \tag{3.5}$$

Lemma 3.7 (criticality of limits of difference quotients)**.** *If $\delta \in \mathcal{D}$ holds, then δ is a critical direction.*

Proof. Let $\delta \in \mathcal{D}$ be given and let $\{\tau_k\} \subset (0, \tau_0)$ be a sequence for δ as in the definition of $\mathcal{D}$. Then it follows from (3.2) that $\delta_{\tau_k}(t) \in K_{\mathrm{rad}}(y(t); Z)$ for all k and all $t \in [0,T]$. Due to Lemma 2.10iii), this yields $\delta_{\tau_k}(t) \in \mathcal{Z}(y(t))$ and, since $\mathcal{Z}(y(t))$ is closed, $\delta_{\tau_k}(t) \to \delta(t) \in \mathcal{Z}(y(t))$ for all $t \in [0,T]$. From (3.3), (3.2) with $z \equiv 0$, Lemma 2.31, and Lemma 2.22, we further obtain that

$$0 \geq \int_0^T \langle \delta_{\tau_k}, \mathrm{d}w \rangle \geq \tau_k \int_0^T \langle \delta_{\tau_k}, \mathrm{d}(\delta_{\tau_k} - h) \rangle \geq \tau_k \left(\frac{1}{2} |\delta_{\tau_k}(T)|^2 - \frac{1}{2} |\delta_{\tau_k}(0)|^2 - d \, \|\delta_{\tau_k}\|_{C([0,T];\mathbb{R}^d)} \operatorname{var}(h;[0,T]) \right).$$

Due to the boundedness of $\{\delta_\tau\}$ in $CBV([0,T];\mathbb{R}^d)$, the right-hand side of the last inequality goes to zero for $k \to \infty$. In combination with Theorem 2.24 and again the boundedness of $\{\delta_\tau\} \subset CBV([0,T];\mathbb{R}^d)$, this yields

$$0 \geq \lim_{k\to\infty} \int_0^T \langle \delta_{\tau_k}, \mathrm{d}w \rangle = \int_0^T \langle \delta, \mathrm{d}w \rangle \geq 0.$$

This completes the proof. □

We remark that the concept of criticality is well known in the field of sensitivity analysis of elliptic variational inequalities of the first and the second kind. Analogously to the conditions in (3.5), in the elliptic setting, it characterizes tangential directions that are contained in the kernel of the multiplier that is associated with the solution of the variational inequality under consideration. For details on this topic and additional relations to the field of second-order necessary and sufficient optimality conditions for optimization problems, we refer the reader to [5, 20, 21, 30, 45, 58] and the references therein. A feature that distinguishes (3.5) from its elliptic counterpart is that the evolutionary nature of (V) makes it possible to localize the property of being an element of the kernel of the multiplier in time. Indeed, from Lemma 2.40, we readily obtain:

Lemma 3.8 (criticality on subintervals)**.** *A function $z \in G([0,T];\mathbb{R}^d)$ is critical if and only if*

$$z(t) \in \mathcal{Z}(y(t)) \quad \forall t \in [0,T] \qquad \textit{and} \qquad \int_{s_1}^{s_2} \langle z, \mathrm{d}w \rangle = 0 \quad \forall\, 0 \leq s_1 < s_2 \leq T. \tag{3.6}$$

Proof. The implication "⇐" is trivial. The implication "⇒" follows from Lemmas 2.19 and 2.40. □

If we combine the localization property in Lemma 3.8 with the decomposition in Lemma 2.38 and Lemma 2.39, then we get:

Lemma 3.9 (criticality and limits)**.** *Let $t \in [0,T]$ be given. Let $\varepsilon > 0$ and $\lambda_i \in CBV([t-\varepsilon, t+\varepsilon] \cap [0,T])$, $i \in \mathcal{A}(y(t))$, be as in Lemma 2.38. Then, for all critical directions $z \in G([0,T];\mathbb{R}^d)$, the following is true.*

i) If $t < T$ holds, then, for every $i \in \mathcal{A}(y(t))$, there exists $0 < \tilde{\varepsilon} \leq \min(\varepsilon, T-t)$ such that

$$\int_t^{t+\tilde{\varepsilon}} \langle z(t+), \nu_i \rangle \mathrm{d}\lambda_i(s) = 0. \tag{3.7}$$

ii) If $t > 0$ holds, then, for every $i \in \mathcal{A}(y(t))$, there exists $0 < \tilde{\varepsilon} \leq \min(\varepsilon, t)$ such that

$$\int_{t-\tilde{\varepsilon}}^t \langle z(t-), \nu_i \rangle \mathrm{d}\lambda_i(s) = 0. \tag{3.8}$$

Proof. We only prove i). The proof of ii) is analogous. Let $t \in [0,T)$, $\varepsilon > 0$, and $\lambda_i \in CBV([t-\varepsilon, t+\varepsilon] \cap [0,T])$, $i \in \mathcal{A}(y(t))$, be as in the assertion of the lemma and let $z \in G([0,T];\mathbb{R}^d)$ be a critical direction. From (3.6), (2.40), (2.22), and Lemma 2.19, we obtain that

$$0 = \int_{s_1}^{s_2} \langle z, \mathrm{d}w \rangle = \int_{s_1}^{s_2} \left\langle z(s), \mathrm{d}\left(\sum_{i \in \mathcal{A}(y(t))} \lambda_i(s)\nu_i \right) \right\rangle = \sum_{i \in \mathcal{A}(y(t))} \int_{s_1}^{s_2} \langle z(s), \nu_i \rangle \, \mathrm{d}\lambda_i(s) \tag{3.9}$$

holds for all $\max(0, t-\varepsilon) \leq s_1 < s_2 \leq \min(T, t+\varepsilon)$. Since $z(s)$ is an element of $\mathcal{Z}(y(s))$ for all $s \in [0,T]$, Lemma 2.39 implies that the integrals in the sum on the right-hand side of (3.9) are nonpositive each. Thus,

$$\int_{s_1}^{s_2} \langle z(s), \nu_i \rangle \, \mathrm{d}\lambda_i(s) = 0 \qquad \forall \max(0, t-\varepsilon) \leq s_1 < s_2 \leq \min(T, t+\varepsilon) \qquad \forall i \in \mathcal{A}(y(t)). \tag{3.10}$$

Let us now suppose that a number $\tilde{\varepsilon} > 0$ with the properties in i) does not exist for some $i \in \mathcal{A}(y(t))$. Then it follows from Lemmas 2.20, 2.26 and 2.22, (3.10), and the monotonicity and continuity of λ_i that

$$\begin{aligned} 0 < |\langle z(t+), \nu_i \rangle| \, (\lambda_i(t+\tilde{\varepsilon}) - \lambda_i(t)) &= \left| \int_t^{t+\tilde{\varepsilon}} \langle z(t+), \nu_i \rangle \mathrm{d}\lambda_i(s) \right| = \left| \int_t^{t+\tilde{\varepsilon}} \langle z(t+) - z(s), \nu_i \rangle \mathrm{d}\lambda_i(s) \right| \\ &\leq \sup_{t < s \leq t+\tilde{\varepsilon}} |z(t+) - z(s)| \operatorname{var}(\lambda_i; [t, t+\tilde{\varepsilon}]) = \sup_{t < s \leq t+\tilde{\varepsilon}} |z(t+) - z(s)| \, (\lambda_i(t+\tilde{\varepsilon}) - \lambda_i(t)) \quad \forall\, 0 < \tilde{\varepsilon} \leq \min(\varepsilon, T-t). \end{aligned}$$

Since we necessarily have $\lambda_i(t+\tilde{\varepsilon}) - \lambda_i(t) > 0$ in the above, it follows that

$$0 < |\langle z(t+), \nu_i \rangle| \leq \sup_{t < s \leq t+\tilde{\varepsilon}} |z(t+) - z(s)| \qquad \forall\, 0 < \tilde{\varepsilon} \leq \min(\varepsilon, T-t).$$

This produces a contradiction with the definition of $z(t+)$. Thus, a number $\tilde{\varepsilon}$ with the properties in i) has to exist for all $i \in \mathcal{A}(y(t))$ and the proof is complete. □

By exploiting Lemma 2.20 and the monotonicity of the multiplier maps λ_i, we can transform the identities in (3.7) and (3.8) into conditions on the left and the right limits of z. To be able to formulate the resulting conditions concisely, we introduce:

Definition 3.10 (directionally active, strictly active, and biactive indices)**.** Let $t \in [0,T]$ be given. Let $\varepsilon > 0$ and $\lambda_i \in CBV([t-\varepsilon, t+\varepsilon] \cap [0,T])$, $i \in \mathcal{A}(y(t))$, be as in Lemma 2.38. We define

$$
\begin{aligned}
\mathcal{A}^+(t) &:= \begin{cases} \{i \in \mathcal{A}(y(t)) \mid \exists \{s_k\} \subset (t,T] \text{ s.t. } s_k \to t \text{ and } i \in \mathcal{A}(y(s_k)) \ \forall k \in \mathbb{N}\} & \text{if } t \in [0,T), \\ \emptyset & \text{if } t = T, \end{cases} \\
\mathcal{A}^-(t) &:= \begin{cases} \{i \in \mathcal{A}(y(t)) \mid \exists \{s_k\} \subset [0,t) \text{ s.t. } s_k \to t \text{ and } i \in \mathcal{A}(y(s_k)) \ \forall k \in \mathbb{N}\} & \text{if } t \in (0,T], \\ \emptyset & \text{if } t = 0, \end{cases} \\
\mathcal{A}^+_{\text{str}}(t) &:= \begin{cases} \{i \in \mathcal{A}(y(t)) \mid \lambda_i|_{[t,t+s]} \neq \text{const } \forall s \in (0, \min(T-t, \varepsilon))\} & \text{if } t \in [0,T), \\ \emptyset & \text{if } t = T, \end{cases} \\
\mathcal{A}^-_{\text{str}}(t) &:= \begin{cases} \{i \in \mathcal{A}(y(t)) \mid \lambda_i|_{[t-s,t]} \neq \text{const } \forall s \in (0, \min(t, \varepsilon))\} & \text{if } t \in (0,T], \\ \emptyset & \text{if } t = 0, \end{cases} \\
\mathcal{B}^+(t) &:= \mathcal{A}(y(t)) \setminus \mathcal{A}^+_{\text{str}}(t), \\
\mathcal{B}^-(t) &:= \mathcal{A}(y(t)) \setminus \mathcal{A}^-_{\text{str}}(t), \\
V^+(t) &:= \operatorname{span}(\{\nu_i \mid i \in \mathcal{A}^+_{\text{str}}(t)\})^\perp, \\
V^-(t) &:= \operatorname{span}(\{\nu_i \mid i \in \mathcal{A}^-_{\text{str}}(t)\})^\perp.
\end{aligned}
$$

Note that the index sets in Definition 3.10 do not depend on the choice of ε; see the comments after Lemma 2.38. Using the quantities introduced in Definition 3.10, we can reformulate the integral notion of criticality in (3.5) in an equivalent, purely pointwise manner.

Theorem 3.11 (pointwise reformulation of criticality)**.** *A function $z \in G([0,T];\mathbb{R}^d)$ is critical if and only if*

$$
z(t) \in \mathcal{Z}(y(t)), \quad z(t+) \in V^+(t), \quad \textit{and} \quad z(t-) \in V^-(t) \qquad \forall t \in [0,T]. \tag{3.11}
$$

Proof. We begin with the implication "⇒". Suppose that a critical direction $z \in G([0,T];\mathbb{R}^d)$ is given and let $t \in [0,T]$ be fixed. Then $z(t) \in \mathcal{Z}(y(t))$ holds by (3.5). In the case $t = T$, we further have $\mathcal{A}^+_{\text{str}}(t) = \emptyset$ and $V^+(t) = \mathbb{R}^d$ by Definition 3.10, and the inclusion $z(t+) \in V^+(t)$ is trivial. If $t \in [0,T)$ holds, then we obtain from Lemmas 2.20 and 3.9 that, for every $i \in \mathcal{A}^+_{\text{str}}(t)$, there exists a number $0 < \tilde{\varepsilon} \leq \min(\varepsilon, T-t)$ satisfying

$$
0 = \int_t^{t+\tilde{\varepsilon}} \langle z(t+), \nu_i \rangle \mathrm{d}\lambda_i(s) = \langle z(t+), \nu_i \rangle \left(\lambda_i(t+\tilde{\varepsilon}) - \lambda_i(t) \right),
$$

where $\varepsilon > 0$ and $\lambda_i \in CBV([t-\varepsilon, t+\varepsilon] \cap [0,T])$ are again as in Lemma 2.38. By the definition of $\mathcal{A}^+_{\text{str}}(t)$ and the monotonicity of λ_i, we have $\lambda_i(t+\tilde{\varepsilon}) - \lambda_i(t) > 0$. Thus, $\langle z(t+), \nu_i \rangle = 0$ for all $i \in \mathcal{A}^+_{\text{str}}(t)$ and it follows that $z(t+) \in V^+(t)$ holds as desired. To check the inclusion $z(t-) \in V^-(t)$, we can proceed along the exact same lines. This shows that every critical direction z satisfies (3.11) and completes the first half of the proof.

It remains to prove the implication "⇐". To this end, let us assume that a function $z \in G([0,T];\mathbb{R}^d)$ satisfying (3.11) is given. That $z(t) \in \mathcal{Z}(y(t))$ holds for all $t \in [0,T]$ in this case as required in (3.5) is obvious. To establish the integral identity in (3.5), let us assume that an arbitrarily small $\gamma > 0$ is given. As z possesses left and right limits at all points $t \in [0,T]$ and due to Lemma 2.38 and Definition 3.10, for every $t \in [0,T]$, we can find a number $\varepsilon_t > 0$ such that the step function

$$
v_t \colon [0,T] \to \mathbb{R}^d, \qquad v_t(s) := \begin{cases} z(t-) & \text{if } s \in [0,t), \\ z(t) & \text{if } s = t, \\ z(t+) & \text{if } s \in (t,T], \end{cases}
$$

satisfies $\|z - v_t\|_{\infty,[t-\varepsilon_t,t+\varepsilon_t]\cap[0,T]} \le \gamma$, such that the assertions of Lemma 2.38 hold for t on $[t-\varepsilon_t, t+\varepsilon_t] \cap [0,T]$ with functions λ_i^t, $i \in \mathcal{A}(y(t))$, and such that $\lambda_i^t = \text{const}$ holds on $[t-\varepsilon_t, t] \cap [0,T]$ for all $i \in \mathcal{B}^-(t)$ and $\lambda_i^t = \text{const}$ holds on $[t, t+\varepsilon_t] \cap [0,T]$ for all $i \in \mathcal{B}^+(t)$. As $[0,T]$ is compact, we can cover it by finitely many of the intervals $(t-\varepsilon_t, t+\varepsilon_t)$. We denote the times t of this cover and their ε_t, v_t, and λ_i^t by $0 \le t_1 \le ... \le t_M \le T$ and ε_m, v_m, λ_i^m, $i \in \mathcal{A}(y(t_m))$, $m = 1, ..., M$, $M \in \mathbb{N}$, respectively. Let $\{\psi_m\}_{m=1}^M$ be a smooth partition of unity on $[0,T]$ subordinate to the cover $\{(t_m - \varepsilon_m, t_m + \varepsilon_m)\}_{m=1}^M$, *i.e.*, a collection of functions satisfying (2.45). We define

$$v(s) := \sum_{m=1}^M \psi_m(s) v_m(s), \qquad s \in [0,T].$$

Then we clearly have $v \in G([0,T];\mathbb{R}^d)$ and, by construction and Lemmas 2.19, 2.20 and 2.26,

$$\begin{aligned}
\int_0^T \langle v, \mathrm{d}w \rangle &= \sum_{m=1}^M \int_{\max(0,t_m-\varepsilon_m)}^{\min(T,t_m+\varepsilon_m)} \langle \psi_m(s) v_m(s), \mathrm{d}w(s) \rangle \\
&= \sum_{m=1}^M \int_{\max(0,t_m-\varepsilon_m)}^{\min(T,t_m+\varepsilon_m)} \left\langle \psi_m(s) v_m(s), \mathrm{d}\left(\sum_{i \in \mathcal{A}(y(t_m))} \lambda_i^m(s) \nu_i \right) \right\rangle \\
&= \sum_{m:\, t_m \neq 0} \sum_{i \in \mathcal{A}(y(t_m))} \int_{\max(0,t_m-\varepsilon_m)}^{t_m} \langle \psi_m(s) z(t_m-), \nu_i \rangle \mathrm{d}\lambda_i^m(s) \\
&\qquad + \sum_{m:\, t_m \neq T} \sum_{i \in \mathcal{A}(y(t_m))} \int_{t_m}^{\min(T,t_m+\varepsilon_m)} \langle \psi_m(s) z(t_m+), \nu_i \rangle \mathrm{d}\lambda_i^m(s) \\
&= \sum_{m:\, t_m \neq 0} \sum_{i \in \mathcal{A}^-_{\text{str}}(t_m)} \int_{\max(0,t_m-\varepsilon_m)}^{t_m} \langle \psi_m(s) z(t_m-), \nu_i \rangle \mathrm{d}\lambda_i^m(s) \\
&\qquad + \sum_{m:\, t_m \neq T} \sum_{i \in \mathcal{A}^+_{\text{str}}(t_m)} \int_{t_m}^{\min(T,t_m+\varepsilon_m)} \langle \psi_m(s) z(t_m+), \nu_i \rangle \mathrm{d}\lambda_i^m(s) \\
&= 0.
\end{aligned}$$

Here, we have used the inclusions $z(t_m+) \in V^+(t_m)$ and $z(t_m-) \in V^-(t_m)$. On the other hand, we also have

$$\sup_{s \in [0,T]} |v(s) - z(s)| \le \sup_{s \in [0,T]} \sum_{m=1}^M \psi_m(s) |v_m(s) - z(s)| \le \sup_{s \in [0,T]} \sum_{m=1}^M \psi_m(s) \gamma = \gamma$$

and, thus,

$$\left| \int_0^T \langle z, \mathrm{d}w \rangle \right| = \left| \int_0^T \langle z - v, \mathrm{d}w \rangle \right| \le d\gamma \operatorname{var}(w; [0,T])$$

by Lemma 2.22. Since $\gamma > 0$ was arbitrary, the above shows that z satisfies the integral identity in (3.5) and, as a consequence, that z is critical. This completes the proof. □

The pointwise reformulation of the notion of criticality in Theorem 3.11 will be an essential ingredient for the analysis of the concept of *temporal polyhedricity* in Section 3.5 and, ultimately, the derivation of the system (1.5) that uniquely characterizes the directional derivatives of $\mathcal{S}$ in (I). Before we can approach this topic, however, we require more information about the pointwise behavior of the elements of the set $\mathcal{D}$. More precisely, we have to study in detail how the elements of $\mathcal{D}$ jump at their points of discontinuity. As we will see in Section 3.5, analyzing the discontinuities of the elements of the set $\mathcal{D}$ will allow us to show that not only all elements δ of $\mathcal{D}$ are critical directions, but also all right-limit functions δ_+ for $\delta \in \mathcal{D}$. Note that this is not obvious at all since it is a priori completely unclear whether $\delta(t+) \in \mathcal{Z}(y(t))$ holds for all $t \in [0, T]$. To prepare the proof of the inclusion $\delta(t+) \in \mathcal{Z}(y(t))$ and our study of the jump directions of the elements of $\mathcal{D}$, we prove:

Lemma 3.12 (tangentiality of left and right limits)**.** *Let $z \in G([0,T];\mathbb{R}^d)$ be critical and let $t \in [0,T]$ be given.*

i) If $t < T$ holds, then there exists $0 < \varepsilon < T - t$ such that

$$z(t+) \in \mathcal{Z}(y(s)) \qquad \forall s \in (t, t+\varepsilon].$$

ii) If $t > 0$ holds, then there exists $0 < \varepsilon < t$ such that

$$z(t-) \in \mathcal{Z}(y(s)) \qquad \forall s \in [t-\varepsilon, t).$$

Proof. We prove i). The proof of ii) is analogous. Let z be critical and let $t \in [0, T)$ be given. A simple contradiction argument shows that we can find $0 < \varepsilon < T - t$ such that $\mathcal{A}(y(s)) \subset \mathcal{A}^+(t)$ holds for all $s \in (t, t+\varepsilon]$. If $i \in \mathcal{A}^+(t)$ holds and $\{s_k\}$ is a sequence for this i as in the definition of $\mathcal{A}^+(t)$, then we further obtain from (3.5), the definition of $z(t+)$, and $(t, T] \ni s_k \to t$ that $0 \geq \langle \nu_i, z(s_k)\rangle \to \langle \nu_i, z(t+)\rangle$ for $k \to \infty$. Thus, $\langle \nu_i, z(t+)\rangle \leq 0$ for all $i \in \mathcal{A}^+(t)$ and, consequently, $\langle \nu_i, z(t+)\rangle \leq 0$ for all $i \in \mathcal{A}(y(s))$ and $s \in (t, t+\varepsilon]$. □

3.3. Identification of the jumps from the left

We now turn our attention to the jumps $\delta(t) - \delta(t-)$ of the elements δ of the set $\mathcal{D}$. To analyze these jumps from the left, we consider the following situation:

Assumption 3.13 (standing assumptions for Section 3.3)**.**

i) $\delta \in \mathcal{D}$ is arbitrary but fixed;
ii) $t \in (0, T]$ is arbitrary but fixed;
iii) $\gamma > 0$ is an arbitrary but fixed given small number;
iv) $\varepsilon > 0$ is chosen such that all of the following conditions hold:
   (a) $\varepsilon < t$ (trivially achievable);
   (b) $\mathcal{I}(y(t)) \subset \mathcal{I}(y(s))$ for all $s \in [t-\varepsilon, t]$ (achievable by continuity of y);
   (c) $\delta(t-) \in \mathcal{Z}(y(s))$ for all $s \in [t-\varepsilon, t)$ (achievable by Lemmas 3.7 and 3.12);
   (d) the assertions in Lemma 2.38 hold with functions $\lambda_i \in CBV([t-\varepsilon, t+\varepsilon] \cap [0,T])$, $i \in \mathcal{A}(y(t))$;
   (e) $\lambda_i = \text{const}$ holds on $[t-\varepsilon, t]$ for all $i \in \mathcal{B}^-(t)$ (achievable by Definition 3.10);
   (f) $\|h - h(t)\|_{CBV([t-\varepsilon,t];\mathbb{R}^d)} + |\delta(t-\varepsilon) - \delta(t-)| \leq \gamma/L$, where $L > 0$ is as in (2.38) (trivially achievable);
v) $\{\tau_k\}$ is a sequence as in the definition of $\mathcal{D}$ for δ with associated difference quotients $\delta_k := \delta_{\tau_k}$ and $y(t-\varepsilon) + \tau_k \delta(t-) \in Z$ holds for all k (achievable due to $\delta(t-) \in \mathcal{Z}(y(t-\varepsilon))$ and Lemma 2.10iii)).

Note that, in Assumption 3.13iv)f, the CBV-regularity of h enters crucially. To study the jump $\delta(t) - \delta(t-)$, we consider auxiliary variational inequalities of the following type:

$$\begin{gathered}\zeta_k \in CBV([t-\varepsilon,t];\mathbb{R}^d), \qquad \zeta_k(s) \in \frac{1}{\tau_k}\left(Z - y(s)\right) \forall s \in [t-\varepsilon,t], \qquad \zeta_k(t-\varepsilon) = \zeta_0,\\ \int_{t-\varepsilon}^t \langle z - \zeta_k, \mathrm{d}\zeta_k\rangle - \frac{1}{\tau_k}\int_{t-\varepsilon}^t \langle z - \zeta_k, \mathrm{d}w\rangle \geq 0 \qquad \forall z \in G([t-\varepsilon,t];\mathbb{R}^d)\colon z(s) \in \frac{1}{\tau_k}\left(Z - y(s)\right)\ \forall s \in [t-\varepsilon,t].\end{gathered} \tag{AUX$_1$}$$

Lemma 3.14 (unique solvability of (AUX$_1$))**.** *The EVI (AUX$_1$) has a unique solution $\zeta_k \in CBV([t-\varepsilon, t]; \mathbb{R}^d)$ for all ζ_0 satisfying $y(t-\varepsilon) + \tau_k \zeta_0 \in Z$ and all k. Further, for all such ζ_0 and k, it holds*

$$\|\delta_k - \zeta_k\|_{CBV([t-\varepsilon,t];\mathbb{R}^d)} \le L\left(\|h - h(t)\|_{CBV([t-\varepsilon,t];\mathbb{R}^d)} + |\delta_k(t-\varepsilon) - \zeta_0|\right), \tag{3.12}$$

where L is the Lipschitz constant in (2.38) associated with Z.

Proof. To see that (AUX$_1$) admits a unique solution $\zeta_k \in CBV([t-\varepsilon, t]; \mathbb{R}^d)$, we note that, by performing the steps leading to (3.2) in reverse and by exploiting Lemma 2.20, (AUX$_1$) can be recast as

$$\begin{aligned}
&y + \tau_k \zeta_k \in CBV([t-\varepsilon, t]; \mathbb{R}^d), \qquad y(s) + \tau_k \zeta_k(s) \in Z \ \forall s \in [t-\varepsilon, t],\\
&y(t-\varepsilon) + \tau_k \zeta_k(t-\varepsilon) = y(t-\varepsilon) + \tau_k \zeta_0,\\
&\int_{t-\varepsilon}^{t} \left\langle v(s) - y(s) - \tau_k \zeta_k(s), \mathrm{d}\big(y(s) + \tau_k \zeta_k(s) - u(s) - \tau_k h(t)\big)\right\rangle \ge 0 \qquad \forall v \in G([t-\varepsilon, t]; Z).
\end{aligned}$$

The above problem has a unique solution $y + \tau_k \zeta_k \in CBV([t-\varepsilon, t]; \mathbb{R}^d)$ for all ζ_0 satisfying $y(t-\varepsilon) + \tau_k \zeta_0 \in Z$ by the existence result for the stop on $[t-\varepsilon, t]$ with forcing term $u + \tau_k h(t)$ and initial value $y(t-\varepsilon) + \tau_k \zeta_0 \in Z$; *cf.* Proposition 2.33 and Lemma 2.34. This proves the unique solvability of (AUX$_1$). By considering the EVI satisfied by $y + \tau_k \delta_k$ on $[t-\varepsilon, t]$ and by invoking (2.38), we further obtain that

$$\begin{aligned}
\|(y + \tau_k \delta_k) - (y + \tau_k \zeta_k)\|_{CBV([t-\varepsilon,t];\mathbb{R}^d)} \le L\big(&\|(u + \tau_k h) - (u + \tau_k h(t))\|_{CBV([t-\varepsilon,t];\mathbb{R}^d)}\\
&+ |(y(t-\varepsilon) + \tau_k \delta_k(t-\varepsilon)) - (y(t-\varepsilon) + \tau_k \zeta_0)|\big).
\end{aligned} \tag{3.13}$$

The above implies (3.12) as desired and completes the proof. □

The main advantage of the auxiliary EVI (AUX$_1$) is that, in contrast to the variational inequality (3.2) satisfied by the difference quotient δ_k, it is independent of the perturbation h of the forcing term. This property makes it possible to derive formulas for the value $\zeta_k(t)$, as the following lemma shows.

Lemma 3.15 (terminal condition for (AUX$_1$))**.** *For all large enough k, the solution ζ_k of the EVI (AUX$_1$) with initial value $\zeta_0 := \delta(t-)$ satisfies $\zeta_k(t) = \pi_{\mathcal{Z}(y(t))}(\delta(t-))$.*

Proof. That $y(t-\varepsilon) + \tau_k \delta(t-) \in Z$ holds follows from Assumption 3.13. Choosing $\delta(t-)$ as ζ_0 in (AUX$_1$) is thus allowed by Lemma 3.14. Note that $y(t-\varepsilon) + \tau_k \delta(t-) \in Z$ also implies $y(t-\varepsilon) + \tau_k \pi_{\mathcal{Z}(y(t))^\circ}(\delta(t-)) \in Z$ by Lemma 2.14i) and (2.6). Thus, we may also consider (AUX$_1$) with ζ_0 chosen as $\pi_{\mathcal{Z}(y(t))^\circ}(\delta(t-))$.

We begin with the latter case, *i.e.*, (AUX$_1$) with $\zeta_0 := \pi_{\mathcal{Z}(y(t))^\circ}(\delta(t-))$, and denote the associated solution of (AUX$_1$) by ξ_k. Define $z := \pi_{\mathcal{Z}(y(t))^\circ}(\xi_k) \in G([t-\varepsilon, t]; \mathbb{R}^d)$. Since $y(s) + \tau_k \xi_k(s) \in Z$ holds for all $s \in [t-\varepsilon, t]$, we again obtain from Lemma 2.14i) and (2.6) that $y(s) + \tau_k z(s) \in Z$ for all $s \in [t-\varepsilon, t]$. Thus, z is an admissible test function in (AUX$_1$), and we may deduce that

$$\begin{aligned}
0 &\le \int_{t-\varepsilon}^{t} \left\langle \pi_{\mathcal{Z}(y(t))^\circ}(\xi_k) - \xi_k, \mathrm{d}\xi_k \right\rangle - \frac{1}{\tau_k} \int_{t-\varepsilon}^{t} \left\langle \pi_{\mathcal{Z}(y(t))^\circ}(\xi_k) - \xi_k, \mathrm{d}w \right\rangle\\
&= -\int_{t-\varepsilon}^{t} \left\langle \pi_{\mathcal{Z}(y(t))}(\xi_k), \mathrm{d}\xi_k \right\rangle + \frac{1}{\tau_k} \int_{t-\varepsilon}^{t} \left\langle \pi_{\mathcal{Z}(y(t))}(\xi_k), \mathrm{d}w \right\rangle\\
&\le -\int_{t-\varepsilon}^{t} \left\langle \pi_{\mathcal{Z}(y(t))}(\xi_k), \mathrm{d}\xi_k \right\rangle\\
&= -\frac{1}{2} \operatorname{dist}_{\mathcal{Z}(y(t))^\circ}(\xi_k(t))^2 + \frac{1}{2} \operatorname{dist}_{\mathcal{Z}(y(t))^\circ}(\xi_k(t-\varepsilon))^2\\
&= -\frac{1}{2} \operatorname{dist}_{\mathcal{Z}(y(t))^\circ}(\xi_k(t))^2.
\end{aligned}$$

Here, we have used Proposition 2.2, $w = u - y$, the inclusion $\mathcal{Z}(y(t)) \subset \mathcal{Z}(y(s))$ for all $s \in [t-\varepsilon, t]$ obtained from (2.40), Lemma 2.40, Lemma 2.31, and $\xi_k(t-\varepsilon) = \pi_{\mathcal{Z}(y(t))^\circ}(\delta(t-))$. The above shows that $\xi_k(t) \in \mathcal{Z}(y(t))^\circ$ which, in combination with (AUX_1), yields $\xi_k(t) \in \mathcal{Z}(y(t)) \cap \mathcal{Z}(y(t))^\circ$ and, thus, $\xi_k(t) = 0$.

Let us now consider the functions $\eta_k(s) := \xi_k(s) + \pi_{\mathcal{Z}(y(t))}(\delta(t-))$, $s \in [t-\varepsilon, t]$. For these functions, it clearly holds $\eta_k \in CBV([t-\varepsilon, t]; \mathbb{R}^d)$ and $\eta_k(t-\varepsilon) = \pi_{\mathcal{Z}(y(t))^\circ}(\delta(t-)) + \pi_{\mathcal{Z}(y(t))}(\delta(t-)) = \delta(t-)$ by (2.6) and the initial value condition for ξ_k. For all $i \in \mathcal{A}(y(t))$, we further have

$$\langle \nu_i, y(s) + \tau_k \eta_k(s) \rangle = \langle \nu_i, y(s) + \tau_k \xi_k(s) + \tau_k \pi_{\mathcal{Z}(y(t))}(\delta(t-)) \rangle \leq \langle \nu_i, y(s) + \tau_k \xi_k(s) \rangle \leq \alpha_i \qquad \forall s \in [t-\varepsilon, t],$$

and, from our assumption $\mathcal{I}(y(t)) \subset \mathcal{I}(y(s))$ for all $s \in [t-\varepsilon, t]$, the fact that $\{\eta_k\}$ and $\{\xi_k\}$ are bounded in $CBV([t-\varepsilon, t]; \mathbb{R}^d)$ by (3.12) and Lemma 3.3, and $\tau_k \to 0$, we obtain that $\langle \nu_i, y(s) + \tau_k \eta_k(s) \rangle \leq \alpha_i$ holds for all $i \in \mathcal{I}(y(t))$, all $s \in [t-\varepsilon, t]$, and all sufficiently large k. In total, this shows that $y(s) + \tau_k \eta_k(s) \in Z$ for all $s \in [t-\varepsilon, t]$ and all large enough k. Suppose now that a function $z \in G([t-\varepsilon, t]; \mathbb{R}^d)$ satisfying $z(s) \in (1/\tau_k)(Z - y(s))$ for all $s \in [t-\varepsilon, t]$ is given. Then we may compute that

$$\begin{aligned}
&\int_{t-\varepsilon}^{t} \langle z - \eta_k, \mathrm{d}\eta_k \rangle - \frac{1}{\tau_k} \int_{t-\varepsilon}^{t} \langle z - \eta_k, \mathrm{d}w \rangle \\
&= \int_{t-\varepsilon}^{t} \left\langle z - \xi_k - \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \mathrm{d}(\xi_k + \pi_{\mathcal{Z}(y(t))}(\delta(t-))) \right\rangle - \frac{1}{\tau_k} \int_{t-\varepsilon}^{t} \left\langle z - \xi_k - \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \mathrm{d}w \right\rangle \\
&= \int_{t-\varepsilon}^{t} \langle z - \xi_k, \mathrm{d}\xi_k \rangle - \frac{1}{\tau_k} \int_{t-\varepsilon}^{t} \langle z - \xi_k, \mathrm{d}w \rangle - \int_{t-\varepsilon}^{t} \left\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \mathrm{d}\xi_k \right\rangle + \frac{1}{\tau_k} \int_{t-\varepsilon}^{t} \left\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \mathrm{d}w \right\rangle \\
&\geq - \left\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \xi_k(t) \right\rangle + \left\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \xi_k(t-\varepsilon) \right\rangle + \frac{1}{\tau_k} \int_{t-\varepsilon}^{t} \left\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \mathrm{d}\left(\sum_{i \in \mathcal{A}(y(t))} \lambda_i \nu_i \right) \right\rangle \\
&= \left\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \pi_{\mathcal{Z}(y(t))^\circ}(\delta(t-)) \right\rangle + \frac{1}{\tau_k} \sum_{i \in \mathcal{A}(y(t))} (\lambda_i(t) - \lambda_i(t-\varepsilon)) \left\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \nu_i \right\rangle \\
&= \frac{1}{\tau_k} \sum_{i \in \mathcal{A}_{\mathrm{str}}^-(t)} (\lambda_i(t) - \lambda_i(t-\varepsilon)) \left\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \nu_i \right\rangle \\
&= 0.
\end{aligned}$$

Here, we have used Lemma 2.20, the EVI satisfied by ξ_k, $\xi_k(t) = 0$, $\xi_k(t-\varepsilon) = \pi_{\mathcal{Z}(y(t))^\circ}(\delta(t-))$, Assumption 3.13, the fact that $\left\langle \pi_{\mathcal{Z}(y(t))}(x), \pi_{\mathcal{Z}(y(t))^\circ}(x) \right\rangle = 0$ holds for all $x \in \mathbb{R}^d$ by (2.6) and Proposition 2.2, and the fact that $\langle \delta(t-), \nu_i \rangle = 0$ holds for all $i \in \mathcal{A}_{\mathrm{str}}^-(t)$ by Lemma 3.7 and Theorem 3.11, which implies $\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)), \nu_i \rangle = 0$ for all $i \in \mathcal{A}_{\mathrm{str}}^-(t)$ by Lemma 2.14ii). If we combine all of the above, then it follows that, for all sufficiently large k, the solution ζ_k of (AUX_1) with initial value $\zeta_0 = \delta(t-)$ is given by $\zeta_k(s) = \eta_k(s) = \xi_k(s) + \pi_{\mathcal{Z}(y(t))}(\delta(t-))$. In particular, we have $\zeta_k(t) = \xi_k(t) + \pi_{\mathcal{Z}(y(t))}(\delta(t-)) = \pi_{\mathcal{Z}(y(t))}(\delta(t-))$ and the assertion of the lemma follows. □

Note that the proof of Lemma 3.15 makes extensive use of the non-obtuseness of Z. With Lemma 3.15 and the stability estimate (3.12) at hand, we can now determine the jump $\delta(t) - \delta(t-)$. For convenience, we formulate the resulting theorem such that it is independent of Assumption 3.13.

Theorem 3.16 (jumps from the left)**.** *For all $t \in [0, T]$ and all $\delta \in \mathcal{D}$, it holds $\delta(t) = \pi_{\mathcal{Z}(y(t))}(\delta(t-))$.*

Proof. In the situation of Assumption 3.13, we obtain from Lemmas 3.14 and 3.15 that, for all sufficiently large k, we have

$$|\delta_k(t) - \pi_{\mathcal{Z}(y(t))}(\delta(t-))| \leq L \left(\|h - h(t)\|_{CBV([t-\varepsilon, t]; \mathbb{R}^d)} + |\delta_k(t-\varepsilon) - \delta(t-)| \right).$$

After passing to the limit $k \to \infty$, this yields

$$|\delta(t) - \pi_{\mathcal{Z}(y(t))}(\delta(t-))| \le L \left(\|h - h(t)\|_{CBV([t-\varepsilon,t];\mathbb{R}^d)} + |\delta(t-\varepsilon) - \delta(t-)| \right) \le \gamma,$$

where $\gamma > 0$ is the arbitrarily small number from Assumption 3.13. Letting γ go to zero proves the claim for all $t \in (0, T]$. For $t = 0$, the assertion follows from the convention $\delta(0) = \delta(0-)$ and Lemma 3.7. □

3.4. Identification of the jumps from the right

Next, we study the jumps $\delta(t+) - \delta(t)$ of the elements δ of $\mathcal{D}$. To this end, we consider the following situation (similarly to the last subsection).

Assumption 3.17 (standing assumptions for Sect. 3.4)**.**

- i) $\delta \in \mathcal{D}$ is arbitrary but fixed;
- ii) $t \in [0, T)$ is arbitrary but fixed;
- iii) $\gamma > 0$ is an arbitrary but fixed given small number;
- iv) $\varepsilon > 0$ is chosen such that all of the following conditions hold:
  - (a) $\varepsilon < T - t$ (trivially achievable);
  - (b) $\mathcal{I}(y(t)) \subset \mathcal{I}(y(s))$ for all $s \in [t, t + \varepsilon]$ (achievable by continuity of y);
  - (c) the assertions in Lemma 2.38 hold with functions $\lambda_i \in CBV([t - \varepsilon, t + \varepsilon] \cap [0, T])$, $i \in \mathcal{A}(y(t))$;
  - (d) $\lambda_i = \text{const}$ holds on $[t, t + \varepsilon]$ for all $i \in \mathcal{B}^+(t)$ (achievable by Def. 3.10);
  - (e) $\|h - h(t)\|_{CBV([t,t+\varepsilon];\mathbb{R}^d)} \le \gamma/L$ holds, where $L > 0$ is as in (2.38) (trivially achievable);
- v) $\{\tau_k\}$ is a sequence as in the definition of $\mathcal{D}$ for δ with associated difference quotients $\delta_k := \delta_{\tau_k}$.

Note that Assumption 3.17iv)e again relies crucially on the CBV-regularity of h. To identify the jump $\delta(t+) - \delta(t)$ in the situation of Assumption 3.17, we proceed analogously to Section 3.3 and consider suitably defined auxiliary EVIs. The first one reads as follows:

$$\begin{gathered}
\zeta_k \in CBV([t, t+\varepsilon];\mathbb{R}^d), \qquad \zeta_k(s) \in \frac{1}{\tau_k}\left(Z_{\mathcal{A}} - y(s)\right) \forall s \in [t, t+\varepsilon], \qquad \zeta_k(t) = \delta_k(t),\\
\int_t^{t+\varepsilon} \langle z - \zeta_k, \mathrm{d}\zeta_k \rangle - \frac{1}{\tau_k} \int_t^{t+\varepsilon} \langle z - \zeta_k, \mathrm{d}w \rangle \ge 0 \qquad \forall z \in G([t, t+\varepsilon];\mathbb{R}^d)\colon z(s) \in \frac{1}{\tau_k}\left(Z_{\mathcal{A}} - y(s)\right)\ \forall s \in [t, t+\varepsilon].
\end{gathered} \tag{AUX$_2$}$$

Here and in what follows, $Z_{\mathcal{A}}$ is defined by

$$Z_{\mathcal{A}} := \{x \in \mathbb{R}^d \mid \langle \nu_i, x \rangle \le \alpha_i \ \forall i \in \mathcal{A}(y(t))\}.$$

Lemma 3.18 (unique solvability of (AUX$_2$))**.** *For all large enough k, (AUX$_2$) is uniquely solvable and its solution ζ_k satisfies*

$$\|\delta_k - \zeta_k\|_{CBV([t,t+\varepsilon];\mathbb{R}^d)} \le \gamma. \tag{3.14}$$

Proof. From our assumption $\mathcal{I}(y(t)) \subset \mathcal{I}(y(s))$ for all $s \in [t, t + \varepsilon]$, the convergence $\tau_k \to 0$, the Lipschitz estimate (2.38), the continuity of y, and Lemma 2.40, we obtain that, for all large enough k, the set Z in the EVIs satisfied by $y = \mathcal{S}(u, y_0)$ and $\mathcal{S}(u + \tau_k h, y_0 + \tau_k h_0)$ on the interval $[t, t + \varepsilon]$ can be replaced by $Z_{\mathcal{A}}$ without changing the solutions of these problems. This implies that the difference quotients δ_k are, for all large enough

k, on $[t, t+\varepsilon]$ also the (necessarily unique) solutions of the following EVIs (*cf.* the derivation of (3.2)):

$$\begin{aligned}
&\delta_k \in CBV([t,t+\varepsilon];\mathbb{R}^d), \qquad \delta_k(s) \in \frac{1}{\tau_k}\left(Z_{\mathcal{A}} - y(s)\right) \forall s \in [t,t+\varepsilon], \qquad \delta_k(t) = \delta_k(t),\\
&\int_t^{t+\varepsilon} \langle z - \delta_k, \mathrm{d}(\delta_k - h)\rangle - \frac{1}{\tau_k}\int_t^{t+\varepsilon} \langle z - \delta_k, \mathrm{d}w\rangle \geq 0\\
&\qquad\qquad \forall z \in G([t,t+\varepsilon];\mathbb{R}^d)\colon z(s) \in \frac{1}{\tau_k}\left(Z_{\mathcal{A}} - y(s)\right)\ \forall s \in [t,t+\varepsilon].
\end{aligned}$$

Note that, in the case $h|_{[t,t+\varepsilon]} = \text{const}$, the above problem is identical to (AUX_2) by Lemma 2.20. To establish the unique solvability of (AUX_2), we can thus argue along the exact same lines as in the proof of Lemma 3.14. The same is true for the proof of (3.14); *cf.* the arguments in (3.13). ☐

Lemma 3.19 (reformulation of (AUX_2))**.** *For all large enough k, the solution ζ_k of* (AUX_2) *is also the (necessarily unique) solution of the following EVI:*

$$\begin{aligned}
&\zeta_k \in CBV([t,t+\varepsilon];\mathbb{R}^d), \qquad \zeta_k(s) \in \mathcal{Z}(y(t))\ \forall s \in [t,t+\varepsilon], \qquad \zeta_k(t) = \delta_k(t),\\
&\int_t^{t+\varepsilon} \langle z - \zeta_k, \mathrm{d}\zeta_k\rangle - \frac{1}{\tau_k}\int_t^{t+\varepsilon} \langle z - \zeta_k, \mathrm{d}w\rangle \geq 0 \qquad \forall z \in G([t,t+\varepsilon];\mathbb{R}^d)\colon z(s) \in \mathcal{Z}(y(t))\ \forall s \in [t,t+\varepsilon].
\end{aligned} \tag{AUX_3}$$

Proof. For every $z \in G([t,t+\varepsilon];\mathbb{R}^d)$ satisfying $z(s) \in \mathcal{Z}(y(t))$ for all $s \in [t,t+\varepsilon]$, it holds

$$\langle \nu_i, y(s) + \tau_k z(s)\rangle \leq \alpha_i + 0 \qquad \forall s \in [t,t+\varepsilon] \qquad \forall i \in \mathcal{A}(y(t)). \tag{3.15}$$

That ζ_k satisfies the variational inequality in (AUX_3) thus follows immediately from (AUX_2). The same is true for the regularity statement and the initial condition in (AUX_3). It remains to prove that $\zeta_k(s) \in \mathcal{Z}(y(t))$ for all $s \in [t,t+\varepsilon]$. To this end, we note that, for all $s \in [t,t+\varepsilon]$, we have $y(s) + \tau_k \pi_{\mathcal{Z}(y(t))}(\zeta_k(s)) \in Z_{\mathcal{A}}$ (by the same estimate as in (3.15)) and $y(s) + \tau_k \pi_{\mathcal{Z}(y(t))^\circ}(\zeta_k(s)) \in Z_{\mathcal{A}}$ by Lemma 2.14i) and (2.6). This allows us to obtain the following from (AUX_2) for all large enough k:

$$\begin{aligned}
0 &\leq \int_t^s \left\langle \pi_{\mathcal{Z}(y(t))}(\zeta_k) - \zeta_k, \mathrm{d}\zeta_k\right\rangle - \frac{1}{\tau_k}\int_t^s \left\langle \pi_{\mathcal{Z}(y(t))}(\zeta_k) - \zeta_k, \mathrm{d}w\right\rangle\\
&= \int_t^s \left\langle -\pi_{\mathcal{Z}(y(t))^\circ}(\zeta_k), \mathrm{d}\zeta_k\right\rangle + \frac{1}{\tau_k^2}\int_t^s \left\langle \tau_k \pi_{\mathcal{Z}(y(t))^\circ}(\zeta_k), \mathrm{d}w\right\rangle\\
&= -\frac{1}{2}\operatorname{dist}_{\mathcal{Z}(y(t))}(\zeta_k(s))^2 + \frac{1}{2}\operatorname{dist}_{\mathcal{Z}(y(t))}(\zeta_k(t))^2 - \frac{1}{\tau_k^2}\int_t^s \left\langle (y + \tau_k \pi_{\mathcal{Z}(y(t))^\circ}(\zeta_k)) - y, \mathrm{d}(y - u)\right\rangle\\
&\leq -\frac{1}{2}\operatorname{dist}_{\mathcal{Z}(y(t))}(\zeta_k(s))^2 \qquad \forall s \in (t,t+\varepsilon].
\end{aligned}$$

Here, we have used (2.6), (2.7), Lemma 2.31, $\zeta_k(t) = \delta_k(t) \in \mathcal{Z}(y(t))$, and the fact that y satisfies (2.33) on $[t,t+\varepsilon]$ with Z replaced by $Z_{\mathcal{A}}$; *cf.* the proof of Lemma 3.18. The above shows that $\zeta_k(s) \in \mathcal{Z}(y(t))$ holds for all $s \in (t,t+\varepsilon]$. Since this inclusion is trivially true for $s = t$, this shows that ζ_k indeed solves (AUX_3). That (AUX_3) can have at most one solution follows from the standard existence and uniqueness result for the stop operator; *cf.* Proposition 2.33. (Note that the admissible set of (AUX_3) does not vary with time.) ☐

Next, we consider the auxiliary problem

$$\begin{aligned}&\xi_k \in CBV([t,t+\varepsilon];\mathbb{R}^d), \qquad \xi_k(s) \in \tilde{\mathcal{Z}}\ \forall s \in [t,t+\varepsilon], \qquad \xi_k(t) = \xi_0,\\ &\int_t^{t+\varepsilon} \langle z - \xi_k, \mathrm{d}\xi_k\rangle - \frac{1}{\tau_k}\int_t^{t+\varepsilon} \langle z - \xi_k, \mathrm{d}w\rangle \geq 0 \qquad \forall z \in G([t,t+\varepsilon];\mathbb{R}^d)\colon z(s) \in \tilde{\mathcal{Z}}\ \forall s \in [t,t+\varepsilon]\end{aligned} \tag{AUX$_4$}$$

with

$$\tilde{\mathcal{Z}} := \{z \in \mathbb{R}^d \mid \langle \nu_i, z\rangle \leq 0\ \forall i \in \mathcal{A}^+_{\mathrm{str}}(t)\}.$$

Lemma 3.20 (unique solvability of (AUX$_4$))**.** *Problem* (AUX$_4$) *has a unique solution* ξ_k *for all* $\xi_0 \in \tilde{\mathcal{Z}}$ *and* k.

Proof. The assertion follows from the standard existence result for the stop operator; *cf.* Proposition 2.33. (Note that the admissible set is time-independent in (AUX$_4$).) □

Lemma 3.21 (order property of (AUX$_4$))**.** *Let* $\xi_0 \in \tilde{\mathcal{Z}}$ *be given and let* ξ_k *be the corresponding solution of* (AUX$_4$)*. Then it holds*

$$\xi_k(s) \in \tilde{\mathcal{Z}} \cap (\xi_0 + \tilde{\mathcal{Z}}^\circ) \quad \textit{and} \quad \langle \nu_i, \xi_k(s)\rangle \leq \langle \nu_i, \xi_0\rangle \quad \forall s \in [t,t+\varepsilon] \quad \forall i \in I \setminus \mathcal{A}^+_{\mathrm{str}}(t).$$

Proof. Consider the function $z := \pi_{\xi_0+\tilde{\mathcal{Z}}^\circ}(\xi_k) \in G([t,t+\varepsilon];\mathbb{R}^d)$. Then it follows from Lemma 2.14iii) that we have $z(s) \in \tilde{\mathcal{Z}}$ and $\xi_k(s) - z(s) \in \tilde{\mathcal{Z}}$ for all $s \in [t,t+\varepsilon]$. In particular, z is an admissible test function in (AUX$_4$), and we may deduce that

$$\begin{aligned}0 &\leq \int_t^s \left\langle \pi_{\xi_0+\tilde{\mathcal{Z}}^\circ}(\xi_k) - \xi_k, \mathrm{d}\xi_k \right\rangle - \frac{1}{\tau_k}\int_t^s \langle z - \xi_k, \mathrm{d}w\rangle\\ &= -\frac{1}{2}\operatorname{dist}_{\xi_0+\tilde{\mathcal{Z}}^\circ}(\xi_k(s))^2 + \frac{1}{2}\operatorname{dist}_{\xi_0+\tilde{\mathcal{Z}}^\circ}(\xi_0)^2 - \frac{1}{\tau_k}\int_t^s \left\langle z - \xi_k, \mathrm{d}\left(\sum_{i\in\mathcal{A}(y(t))} \lambda_i \nu_i\right)\right\rangle\\ &= -\frac{1}{2}\operatorname{dist}_{\xi_0+\tilde{\mathcal{Z}}^\circ}(\xi_k(s))^2 + \frac{1}{\tau_k}\sum_{i\in\mathcal{A}^+_{\mathrm{str}}(t)} \int_t^s \langle \xi_k - z, \nu_i\rangle \mathrm{d}\lambda_i\\ &\leq -\frac{1}{2}\operatorname{dist}_{\xi_0+\tilde{\mathcal{Z}}^\circ}(\xi_k(s))^2 \qquad \forall s \in (t,t+\varepsilon].\end{aligned}$$

Here, we have used (2.5), Lemma 2.31, (2.40), $\xi_0 \in \xi_0 + \tilde{\mathcal{Z}}^\circ$, our assumption that $\lambda_i = \mathrm{const}$ holds on $[t,t+\varepsilon]$ for all $i \in \mathcal{B}^+(t)$, Lemma 2.20, the fact that λ_i is nondecreasing for all $i \in \mathcal{A}^+_{\mathrm{str}}(t)$, $\xi_k(s) - z(s) \in \tilde{\mathcal{Z}}$ for all $s \in [t,t+\varepsilon]$, the definition of $\tilde{\mathcal{Z}}$, and Lemma 2.28. In combination with the admissibility of ξ_k in (AUX$_4$) and the trivial inclusion $\xi_0 \in \tilde{\mathcal{Z}} \cap (\xi_0 + \tilde{\mathcal{Z}}^\circ)$, the above yields $\xi_k(s) \in \tilde{\mathcal{Z}} \cap (\xi_0 + \tilde{\mathcal{Z}}^\circ)$ for all $s \in [t,t+\varepsilon]$ as desired. Note that, due to the definition of $\tilde{\mathcal{Z}}$ and Lemma 2.11, the latter inclusion implies that, for every $s \in [t,t+\varepsilon]$, there exist numbers $\beta_j \geq 0$, $j \in \mathcal{A}^+_{\mathrm{str}}(t)$, satisfying

$$\xi_k(s) = \xi_0 + \sum_{j\in\mathcal{A}^+_{\mathrm{str}}(t)} \beta_j \nu_j.$$

As Z is non-obtuse, the above yields

$$\langle \nu_i, \xi_k(s)\rangle = \left\langle \nu_i, \xi_0 + \sum_{j\in\mathcal{A}^+_{\mathrm{str}}(t)} \beta_j \nu_j \right\rangle = \langle \nu_i, \xi_0\rangle + \sum_{j\in\mathcal{A}^+_{\mathrm{str}}(t)} \beta_j \langle \nu_i, \nu_j\rangle \leq \langle \nu_i, \xi_0\rangle \qquad \forall i \in I \setminus \mathcal{A}^+_{\mathrm{str}}(t).$$

This completes the proof of the lemma. ☐

Note that the proofs of Lemmas 3.19 and 3.21 again rely crucially on the non-obtuseness of Z. Next, we establish a connection between (AUX$_2$) and (AUX$_4$). Recall that $V^+(t) := \operatorname{span}(\{\nu_i \mid i \in \mathcal{A}^+_{\text{str}}(t)\})^\perp$ and, consequently, $V^+(t)^\perp = \operatorname{span}(\{\nu_i \mid i \in \mathcal{A}^+_{\text{str}}(t)\})$; see Definition 3.10.

Lemma 3.22 (relationship between the auxiliary problems (AUX$_2$) and (AUX$_4$))**.** *Let ξ_k be the solution of (AUX$_4$) with initial value $\xi_0 := \pi_{V^+(t)^\perp}(\delta_k(t))$. Then, for all sufficiently large k, it holds $\zeta_k = \xi_k + \pi_{V^+(t)}(\delta_k(t))$, where ζ_k is the solution of the EVI (AUX$_2$).*

Proof. Note that $\delta_k(t) \in \mathcal{Z}(y(t))$ and the definition of $V^+(t)$ yield $\langle \nu_i, \xi_0\rangle = \langle \nu_i, \pi_{V^+(t)^\perp}(\delta_k(t))\rangle = \langle \nu_i, \delta_k(t)\rangle \le 0$ for all $i \in \mathcal{A}^+_{\text{str}}(t)$. Thus, $\xi_0 \in \tilde{\mathcal{Z}}$ and ξ_0 is an admissible initial value in (AUX$_4$). Define $\hat{\xi}_k := \xi_k + \pi_{V^+(t)}(\delta_k(t))$. We claim that this function solves (AUX$_4$) with initial value $\hat{\xi}_0 := \delta_k(t) \in \mathcal{Z}(y(t)) \subset \tilde{\mathcal{Z}}$. To see that this is true, we note that $\hat{\xi}_k$ trivially satisfies $\hat{\xi}_k \in CBV([t, t+\varepsilon]; \mathbb{R}^d)$ and

$$\hat{\xi}_k(t) = \xi_k(t) + \pi_{V^+(t)}(\delta_k(t)) = \pi_{V^+(t)^\perp}(\delta_k(t)) + \pi_{V^+(t)}(\delta_k(t)) = \delta_k(t) = \hat{\xi}_0.$$

As $\langle \nu_i, \pi_{V^+(t)}(\delta_k(t))\rangle = 0$ holds for all $i \in \mathcal{A}^+_{\text{str}}(t)$ and due to $\xi_k \in G([t, t+\varepsilon]; \tilde{\mathcal{Z}})$, we further have

$$\langle \nu_i, \hat{\xi}_k(s)\rangle = \langle \nu_i, \xi_k(s) + \pi_{V^+(t)}(\delta_k(t))\rangle = \langle \nu_i, \xi_k(s)\rangle \le 0 \qquad \forall i \in \mathcal{A}^+_{\text{str}}(t) \qquad \forall s \in [t, t+\varepsilon]$$

and, thus, $\hat{\xi}_k(s) \in \tilde{\mathcal{Z}}$ for all $s \in [t, t+\varepsilon]$. To prove that $\hat{\xi}_k$ indeed solves (AUX$_4$) with $\hat{\xi}_0 := \delta_k(t)$, it remains to check the variational inequality in (AUX$_4$). So let $z \in G([t, t+\varepsilon]; \mathbb{R}^d)$ satisfying $z(s) \in \tilde{\mathcal{Z}}$ for all $s \in [t, t+\varepsilon]$ be given. From Lemma 3.21 and Lemma 2.11, we obtain that $\xi_k(s) \in (\xi_0 + \tilde{\mathcal{Z}}^\circ) \subset \operatorname{span}(\{\nu_i \mid i \in \mathcal{A}^+_{\text{str}}(t)\}) = V^+(t)^\perp$ holds for all $s \in [t, t+\varepsilon]$, and from (2.40), Lemma 2.20, and our assumption that $\lambda_i = \text{const}$ holds on $[t, t+\varepsilon]$ for all $i \in \mathcal{B}^+(t)$, it follows that

$$\begin{aligned}
\int_t^{t+\varepsilon} \left\langle \pi_{V^+(t)}(\delta_k(t)), \mathrm{d}w\right\rangle &= \sum_{i \in \mathcal{A}(y(t))} \int_t^{t+\varepsilon} \left\langle \pi_{V^+(t)}(\delta_k(t)), \nu_i\right\rangle \mathrm{d}\lambda_i \\
&= \sum_{i \in \mathcal{A}(y(t))} \langle \nu_i, \pi_{V^+(t)}(\delta_k(t))\rangle(\lambda_i(t+\varepsilon) - \lambda_i(t)) \\
&= \sum_{i \in \mathcal{A}^+_{\text{str}}(t)} \langle \nu_i, \pi_{V^+(t)}(\delta_k(t))\rangle(\lambda_i(t+\varepsilon) - \lambda_i(t)) = 0.
\end{aligned}$$

In combination with the EVI for ξ_k and again Lemma 2.20, this yields

$$\begin{aligned}
&\int_t^{t+\varepsilon} \left\langle z - \hat{\xi}_k, \mathrm{d}\hat{\xi}_k\right\rangle - \frac{1}{\tau_k}\int_t^{t+\varepsilon} \left\langle z - \hat{\xi}_k, \mathrm{d}w\right\rangle \\
&= \int_t^{t+\varepsilon} \left\langle z - \xi_k - \pi_{V^+(t)}(\delta_k(t)), \mathrm{d}(\xi_k + \pi_{V^+(t)}(\delta_k(t)))\right\rangle - \frac{1}{\tau_k}\int_t^{t+\varepsilon} \left\langle z - \xi_k - \pi_{V^+(t)}(\delta_k(t)), \mathrm{d}w\right\rangle \\
&= -\int_t^{t+\varepsilon} \left\langle \pi_{V^+(t)}(\delta_k(t)), \mathrm{d}\xi_k\right\rangle + \int_t^{t+\varepsilon} \langle z - \xi_k, \mathrm{d}\xi_k\rangle - \frac{1}{\tau_k}\int_t^{t+\varepsilon} \langle z - \xi_k, \mathrm{d}w\rangle \\
&\ge -\int_t^{t+\varepsilon} \left\langle \pi_{V^+(t)}(\delta_k(t)), \mathrm{d}\xi_k\right\rangle \\
&= \left\langle \pi_{V^+(t)}(\delta_k(t)), \xi_k(t) - \xi_k(t+\varepsilon)\right\rangle = 0,
\end{aligned}$$

where, in the last line, we have again used that $\xi_k(s) \in V^+(t)^\perp$ holds for all $s \in [t, t+\varepsilon]$. Thus, $\hat{\xi}_k := \xi_k + \pi_{V^+(t)}(\delta_k(t))$ indeed solves (AUX$_4$) with initial value $\hat{\xi}_0 := \delta_k(t)$. Next, we prove that $\hat{\xi}_k = \zeta_k$ holds for all large enough k. Note that, since ζ_k solves (AUX$_3$) for all large enough k, since $\mathcal{Z}(y(t)) \subset \tilde{\mathcal{Z}}$ holds, since ζ_k and $\hat{\xi}_k$ have the same initial value, and since (AUX$_3$) is uniquely solvable, to show that $\hat{\xi}_k$ is identical ζ_k, it suffices to prove that $\hat{\xi}_k(s) \in \mathcal{Z}(y(t))$ for all $s \in [t, t+\varepsilon]$. To see this, we note that $\xi_k \in G([t, t+\varepsilon]; \tilde{\mathcal{Z}})$ yields

$$\langle \nu_i, \hat{\xi}_k(s)\rangle = \langle \nu_i, \xi_k(s) + \pi_{V^+(t)}(\delta_k(t))\rangle = \langle \nu_i, \xi_k(s)\rangle \leq 0 \qquad \forall i \in \mathcal{A}^+_{\text{str}}(t) \qquad \forall s \in [t, t+\varepsilon]$$

and that Lemma 3.21 and the inclusion $\delta_k(t) \in \mathcal{Z}(y(t))$ imply

$$\langle \nu_i, \hat{\xi}_k(s)\rangle \leq \langle \nu_i, \hat{\xi}_0\rangle = \langle \nu_i, \delta_k(t)\rangle \leq 0 \quad \forall s \in [t, t+\varepsilon] \quad \forall i \in \mathcal{A}(y(t)) \setminus \mathcal{A}^+_{\text{str}}(t).$$

Thus, we indeed have $\hat{\xi}_k(s) \in \mathcal{Z}(y(t))$ for all $s \in [t, t+\varepsilon]$ and the assertion of the lemma follows. ☐

By combining Lemmas 3.18 to 3.22, we can now prove the main result of this subsection. Analogously to Theorem 3.16, we formulate it independently of Assumption 3.17.

Theorem 3.23 (jumps from the right)**.** *For all $t \in [0, T]$ and all $\delta \in \mathcal{D}$, it holds $\delta(t+) = \pi_{V^+(t)}(\delta(t))$.*

Proof. For $t = T$, we have $V^+(t) = \mathbb{R}^d$ and $\delta(t+) = \delta(t)$ by definition/convention and the assertion is trivial. We may thus assume in the following that $t \in [0, T)$. Suppose that $\gamma > 0$ is given and consider the situation in Assumption 3.17. Let k be large enough and let ξ_k and ζ_k be as in Lemma 3.22. Then we know from Lemma 3.18 that $\|\delta_k - \zeta_k\|_{CBV([t,t+\varepsilon];\mathbb{R}^d)} \leq \gamma$ holds, from Lemma 3.22 that $\zeta_k = \xi_k + \pi_{V^+(t)}(\delta_k(t))$ holds on $[t, t+\varepsilon]$, and from Lemma 3.21, $\xi_0 = \pi_{V^+(t)^\perp}(\delta_k(t))$, and Lemma 2.11 that $\xi_k(s) \in \operatorname{span}(\{\nu_i \mid i \in \mathcal{A}^+_{\text{str}}(t)\}) = V^+(t)^\perp$. Combining all of this with (2.4) yields

$$\begin{aligned}
\|\pi_{V^+(t)}(\delta_k) - \pi_{V^+(t)}(\delta_k(t))\|_{C([t,t+\varepsilon];\mathbb{R}^d)} &= \|\pi_{V^+(t)}(\delta_k) - \pi_{V^+(t)}(\delta_k(t)) - \pi_{V^+(t)}(\xi_k)\|_{C([t,t+\varepsilon];\mathbb{R}^d)} \\
&= \|\pi_{V^+(t)}(\delta_k) - \pi_{V^+(t)}(\zeta_k)\|_{C([t,t+\varepsilon];\mathbb{R}^d)} \\
&\leq \|\delta_k - \zeta_k\|_{C([t,t+\varepsilon];\mathbb{R}^d)} \\
&\leq \gamma
\end{aligned}$$

for all large enough k and, after passing to the limit $k \to \infty$,

$$\left|\pi_{V^+(t)}(\delta(s)) - \pi_{V^+(t)}(\delta(t))\right| \leq \gamma \qquad \forall s \in [t, t+\varepsilon].$$

Due to the definition of $\delta(t+)$, this shows that $\left|\pi_{V^+(t)}(\delta(t+)) - \pi_{V^+(t)}(\delta(t))\right| \leq \gamma$ and, since $\delta(t+) \in V^+(t)$ holds by Lemma 3.7 and Theorem 3.11, that $\left|\delta(t+) - \pi_{V^+(t)}(\delta(t))\right| \leq \gamma$. As $\gamma > 0$ was arbitrary, the assertion of the theorem now follows. ☐

3.5. Temporal polyhedricity

With Theorems 3.16 and 3.23 at hand, we can revisit the notion of criticality from Section 3.2 to arrive at the following important corollary.

Corollary 3.24 (criticality of right limits)**.** *If $\delta \in \mathcal{D}$ holds, then δ_+ is a critical direction.*

Proof. Let $\delta \in \mathcal{D}$ be given. We first check that $\delta_+(t) \in \mathcal{Z}(y(t))$ holds for all $t \in [0, T]$. For $t = T$, this inclusion is trivial since our conventions for the right limit and Lemma 3.7 imply $\delta_+(T) = \delta(T+) = \delta(T) \in \mathcal{Z}(y(T))$. For $t \in [0, T)$, we obtain from Lemma 3.7 and Theorem 3.23 that $\delta(t+) = \pi_{V^+(t)}(\delta(t))$ with $\delta(t) \in \mathcal{Z}(y(t))$ and $V^+(t) := \operatorname{span}(\{\nu_i \mid i \in \mathcal{A}^+_{\text{str}}(t)\})^\perp$. In view of Lemma 2.14iv), this implies $\delta(t+) \in \mathcal{Z}(y(t))$ for all $t \in [0, T)$.

We thus indeed have $\delta_+(t) \in \mathcal{Z}(y(t))$ for all $t \in [0,T]$. From (2.1), Lemma 3.7, and Theorem 3.11, we further obtain that $\delta_+(t-) = \delta(t-) \in V^-(t)$ for all $t \in (0,T]$ and $\delta_+(t+) = \delta(t+) \in V^+(t)$ for all $t \in [0,T)$. Since $V^-(0) = V^+(T) = \mathbb{R}^d$ holds, this shows that δ_+ satisfies the conditions in (3.11). By invoking Theorem 3.11, it now follows that δ_+ is critical and the proof is complete. □

Corollary 3.24 motivates the following definition.

Definition 3.25 (critical cone in $G_r([0,T];\mathbb{R}^d)$)**.** We define

$$\mathcal{K}^{\text{crit}}_{G_r}(y,u) := \left\{ z \in G_r([0,T];\mathbb{R}^d) \;\middle|\; z \text{ is a critical direction} \right\}.$$

Note that, by Definition 3.6, Lemma 3.8, and Theorem 3.11, we have

$$\begin{aligned}\mathcal{K}^{\text{crit}}_{G_r}(y,u) &= \left\{ z \in G_r([0,T];\mathbb{R}^d) \;\middle|\; z(t) \in \mathcal{Z}(y(t))\ \forall t \in [0,T] \text{ and } \int_0^T \langle z, \mathrm{d}w\rangle = 0 \right\} \\ &= \left\{ z \in G_r([0,T];\mathbb{R}^d) \;\middle|\; z(t) \in \mathcal{Z}(y(t))\ \forall t \in [0,T] \text{ and } \int_{s_1}^{s_2} \langle z, \mathrm{d}w\rangle = 0\ \forall\, 0 \le s_1 < s_2 \le T \right\} \\ &= \left\{ z \in G_r([0,T];\mathbb{R}^d) \;\middle|\; z(t) \in \mathcal{Z}(y(t)),\ z(t+) \in V^+(t),\ z(t-) \in V^-(t)\ \forall t \in [0,T] \right\}.\end{aligned} \tag{3.16}$$

From Corollary 3.24, we further obtain:

Corollary 3.26 (critical cone and right limits)**.** *For every $\delta \in \mathcal{D}$, it holds $\delta_+ \in \mathcal{K}^{\text{crit}}_{G_r}(y,u)$.*

The main goal of this subsection is to demonstrate that elements of the cone $\mathcal{K}^{\text{crit}}_{G_r}(y,u)$ can be approximated (in an appropriate sense) by directions that are (a) more regular, (b) critical, and (c) admissible in the EVI (3.2) satisfied by the difference quotients δ_τ. As we will see in Section 3.6, this approximation property is crucial for the derivation of the system (1.5) that implies that $\mathcal{D}$ is a singleton and, ultimately, characterizes the directional derivatives of $\mathcal{S}$. We remark that, in the context of elliptic variational inequalities, density results of this type are typically referred to as *polyhedricity* conditions; see [20, 30, 45, 59] and the references therein. Accordingly, we call the resulting approximation property *temporal polyhedricity*. The set of critical radial directions that we use for the approximation of the elements of $\mathcal{K}^{\text{crit}}_{G_r}(y,u)$ is defined as follows:

Definition 3.27 (critical radial BV_r-directions with zero left limits at jump points)**.** We define

$$\mathcal{K}^{\text{rad,crit},0}_{BV_r}(y,u) := \Big\{ z \in \mathcal{K}^{\text{crit}}_{G_r}(y,u) \cap BV_r([0,T];\mathbb{R}^d) \;\Big|\; \begin{array}{l}\text{there exists a number } \tau > 0 \text{ independent of } s \text{ such that} \\ y(s) + \tau z(s) \in Z \text{ is true for all } s \in [0,T] \text{ and it holds} \\ z(s-) = 0 \text{ for all } s \in [0,T] \text{ with } z(s-) \neq z(s) \Big\}.\end{array}$$

We can now prove the main result of this subsection.

Theorem 3.28 (temporal polyhedricity)**.** *Let $z \in \mathcal{K}^{\text{crit}}_{G_r}(y,u)$ be given. Then there exist $z_{k,l}, z_l \in G_r([0,T];\mathbb{R}^d)$, $k,l \in \mathbb{N}$, such that the following is true:*

$$\begin{gathered} z_{k,l} \in \mathcal{K}^{\text{rad,crit},0}_{BV_r}(y,u), \qquad \|z_{k,l}\|_\infty \le \|z\|_\infty\ \forall k,l, \\ z_l \in \mathcal{K}^{\text{crit}}_{G_r}(y,u) \cap BV_r([0,T];\mathbb{R}^d), \qquad \|z_l\|_\infty \le \|z\|_\infty\ \forall l, \\ z_{k,l} \to z_l \text{ pointwise in } [0,T] \text{ for } k \to \infty \text{ for all } l, \\ z_l \to z \text{ uniformly on } [0,T] \text{ for } l \to \infty. \end{gathered} \tag{3.17}$$

Proof. Let $z \in \mathcal{K}_{G_r}^{\text{crit}}(y,u)$ be given. To construct the sequence $\{z_l\}$, we proceed along the lines of the proof of Theorem 3.11. Let $l \in \mathbb{N}$ be arbitrary but fixed. As z is right-continuous and regulated, for every $t \in [0,T]$, there exists a number $\varepsilon_t > 0$ such that the step function

$$v_t\colon [0,T] \to \mathbb{R}^d, \qquad v_t(s) := \begin{cases} z(t-) & \text{if } s \in [0,t), \\ z(t+) & \text{if } s \in [t,T], \end{cases}$$

satisfies $\|z - v_t\|_{\infty,[t-\varepsilon_t,t+\varepsilon_t]\cap[0,T]} \leq 1/l$, such that the assertions of Lemma 2.38 hold for t on $[t-\varepsilon_t, t+\varepsilon_t]\cap[0,T]$ with functions λ_i^t, $i \in \mathcal{A}(y(t))$, such that $\lambda_i^t = \text{const}$ holds on $[t-\varepsilon_t, t]\cap[0,T]$ for all $i \in \mathcal{B}^-(t)$, such that $\lambda_i^t = \text{const}$ holds on $[t, t+\varepsilon_t]\cap[0,T]$ for all $i \in \mathcal{B}^+(t)$, such that $I \setminus \mathcal{A}^+(t) \subset \mathcal{I}(y(s))$ holds for all $s \in (t, t+\varepsilon_t]\cap[0,T]$, and such that $I \setminus \mathcal{A}^-(t) \subset \mathcal{I}(y(s))$ holds for all $s \in [t-\varepsilon_t, t)\cap[0,T]$. Due to its compactness, we can cover the interval $[0,T]$ by finitely many of the intervals $(t-\varepsilon_t, t+\varepsilon_t)$. We denote the times t of this cover and the associated ε_t, v_t, and λ_i^t by $0 \leq t_1 \leq ... \leq t_M \leq T$ and ε_m, v_m, λ_i^m, $i \in \mathcal{A}(y(t_m))$, $m = 1, ..., M$, $M \in \mathbb{N}$, respectively, and again consider a smooth partition of unity $\{\psi_m\}_{m=1}^M$ on $[0,T]$ subordinate to the cover $\{(t_m - \varepsilon_m, t_m + \varepsilon_m)\}_{m=1}^M$, as in (2.45). We assume w.l.o.g. that the points t_m are distinct and that $\psi_m(t_m) = 1$ holds for all $m = 1, ..., M$ and $\psi_m(t_j) = 0$ for all $m \neq j$. For this choice of the partition of unity, we define

$$z_l(s) := \sum_{m=1}^M \psi_m(s) v_m(s) \qquad s \in [0,T]. \tag{3.18}$$

The function z_l obtained along these lines is clearly in $BV_r([0,T];\mathbb{R}^d)$ and satisfies

$$\int_0^T \langle z_l, \mathrm{d}w\rangle = 0 \qquad \text{and} \qquad \sup_{s\in[0,T]} |z_l(s) - z(s)| \leq \frac{1}{l}$$

by the exact same arguments as in the proof of Theorem 3.11. From the definition of the left and the right limit, the definitions of $\mathcal{A}^+(t_m)$ and $\mathcal{A}^-(t_m)$, our choice of ε_m, and $z(t) = z(t+) \in \mathcal{Z}(y(t))$ for all $t \in [0,T]$, it follows further that $\langle \nu_i, z(t_m-)\rangle \leq 0$ holds for all $i \in \mathcal{A}^-(t_m)$, that $\langle \nu_i, z(t_m+)\rangle \leq 0$ holds for all $i \in \mathcal{A}(y(t_m))$, that $\mathcal{A}(y(s)) \subset \mathcal{A}^-(t_m)$ holds for all $s \in [t_m - \varepsilon_m, t_m)\cap[0,T]$, and that $\mathcal{A}(y(s)) \subset \mathcal{A}(y(t_m))$ holds for all $s \in [t_m, t_m + \varepsilon_m]\cap[0,T]$. By combining all of this, we obtain that $z(t_m-) \in \mathcal{Z}(y(s))$ holds for all $s \in [t_m - \varepsilon_m, t_m)\cap[0,T]$, that $z(t_m+) \in \mathcal{Z}(y(s))$ holds for all $s \in [t_m, t_m+\varepsilon_m]\cap[0,T]$, and – due to the definition of z_l and the properties of $\{\psi_m\}$ – that $z_l(s) \in \mathcal{Z}(y(s))$ holds for all $s \in [0,T]$. In total and in view of (3.16), this shows that the functions $\{z_l\}$ satisfy $\{z_l\} \subset \mathcal{K}_{G_r}^{\text{crit}}(y,u) \cap BV_r([0,T];\mathbb{R}^d)$ and $\|z_l - z\|_\infty \leq 1/l \to 0$ for $l \to \infty$ as desired. That $\|z_l\|_\infty \leq \|z\|_\infty$ holds for all l follows immediately from (3.18). This completes the construction of the sequence $\{z_l\}$.

It remains to construct the functions $\{z_{k,l}\}$. To this end, we consider a function $\varphi \in C^\infty(\mathbb{R})$ satisfying

$$0 \leq \varphi(s) \leq 1 \;\forall s \in \mathbb{R}, \quad \varphi(s) = 1 \;\forall s \in (-\infty,-1], \quad \varphi(s) = 0 \;\forall s \in [0,\infty),$$

and define

$$z_{k,l}(s) := \sum_{m=1}^M \left[\varphi\left(\frac{s - t_m + 1/k}{1/k}\right) + \mathbb{1}_{[t_m,T]}(s)\right] \psi_m(s) v_m(s) \qquad s \in [0,T] \qquad k,l \in \mathbb{N}. \tag{3.19}$$

Note that these functions trivially satisfy $\|z_{k,l}\|_\infty \leq \|z\|_\infty$ for all l,k and $z_{k,l}(s) \to z_l(s)$ for all $s \in [0,T]$ and $k \to \infty$. Moreover, it holds $\{z_{k,l}\} \subset \mathcal{K}_{G_r}^{\text{crit}}(y,u) \cap BV_r([0,T];\mathbb{R}^d)$ for the exact same reasons as in the case of the sequence $\{z_l\}$ and $z_{k,l}(s-) = 0$ for all $s \in [0,T]$ with $z_{k,l}(s-) \neq z_{k,l}(s)$ (namely, the points t_m with $v_m(t_m) \neq 0$) by construction. To see that $\{z_{k,l}\} \subset \mathcal{K}_{BV_r}^{\text{rad,crit},0}(y,u)$ holds, it remains to show that there exists $\tau > 0$ such that

$y(s) + \tau z_{k,l}(s) \in Z$ holds for all $s \in [0,T]$. To this end, let $m \in \{1, ..., M\}$ be fixed. From our choice of ε_m, we obtain that $\mathcal{A}(y(s)) \subset \mathcal{A}(y(t_m))$ holds for all $s \in [t_m, t_m + \varepsilon_m] \cap [0,T]$ and that $z(t_m) = z(t_m+) \in \mathcal{Z}(y(t_m))$. This implies that

$$\langle \nu_i, y(s) + \tau z(t_m+)\rangle \leq \alpha_i \quad \forall i \in \mathcal{A}(y(t_m)) \quad \forall s \in [t_m, t_m + \varepsilon_m] \cap [0,T] \quad \forall \tau > 0$$

and

$$\langle \nu_i, y(s)\rangle < \alpha_i \quad \forall i \in \mathcal{I}(y(t_m)) \quad \forall s \in [t_m, t_m + \varepsilon_m] \cap [0,T].$$

Due to the continuity of y, the above shows that there exists $\hat{\tau}_0 > 0$ satisfying $y(s) + \tau z(t_m+) \in Z$ for all $0 < \tau < \hat{\tau}_0$ and all $s \in [t_m, t_m + \varepsilon_m] \cap [0,T]$. Let us next consider the interval $[t_m - \varepsilon_m, t_m) \cap [0,T]$. Due to our choice of ε_m and the definitions of $\mathcal{A}^-(t_m)$ and $z(t_m-)$, it holds

$$\langle \nu_i, y(s) + \tau z(t_m-)\rangle \leq \alpha_i \quad \forall i \in \mathcal{A}^-(t_m) \quad \forall s \in [t_m - \varepsilon_m, t_m] \cap [0,T] \quad \forall \tau > 0$$

and

$$\langle \nu_i, y(s)\rangle < \alpha_i \quad \forall i \in I \setminus \mathcal{A}^-(t_m) \quad \forall s \in [t_m - \varepsilon_m, t_m - 1/k] \cap [0,T].$$

Due to the continuity of y, this implies that there exists $\tilde{\tau}_0 \in (0, \hat{\tau}_0)$ satisfying $y(s) + \tau z(t_m-) \in Z$ for all $0 < \tau < \tilde{\tau}_0$ and all $s \in [t_m - \varepsilon_m, t_m - 1/k] \cap [0,T]$. In total, we have now shown that

$$y(s) + \tau \left(\mathbb{1}_{[t_m - \varepsilon_m, t_m - 1/k] \cap [0,T]}(s) z(t_m-) + \mathbb{1}_{[t_m, t_m + \varepsilon_m] \cap [0,T]}(s) z(t_m+)\right) \in Z \qquad \forall s \in [0,T] \qquad \forall \tau \in (0, \tilde{\tau}_0).$$

Due to the properties of ψ_m and φ, the definition of v_m, and the convexity of Z, this yields

$$y(s) + \tau \left[\varphi\left(\frac{s - t_m + 1/k}{1/k}\right) + \mathbb{1}_{[t_m, T]}(s)\right] \psi_m(s) v_m(s) \in Z \qquad \forall s \in [0,T] \qquad \forall \tau \in (0, \tilde{\tau}_0).$$

Due to the convexity of Z and the definition (3.19), it now follows immediately that there exists $\tau > 0$ such that $y(s) + \tau z_{k,l}(s) \in Z$ holds for all $s \in [0,T]$. This completes the proof. □

Remark 3.29.

- In the special case $d = 1$ and $Z = [-r, r]$, $r > 0$, a density result analogous to that in Theorem 3.28 has already been proven in [48], Theorem 6.5 (with roughly the same techniques). Note that, in the latter theorem, the approximation is even accomplished with elements of the set $\mathcal{K}^{\text{rad,crit},0}_{BV_r}(y, u) \cap C^\infty([0,T])$. The price that one pays for this improved smoothness (that is essential in the optimal control context of [48]) is that density is only obtained for a certain subset of the critical cone $\mathcal{K}^{\text{crit}}_{G_r}(y, u)$, which, however, still contains the right-limit functions of the directional derivatives of $\mathcal{S}$ in the one-dimensional setting. Whether Theorem 3.28 can be refined in a similar manner for $d > 1$ is an open problem. For further remarks on this topic, we refer to Section 5.
- As already mentioned, in the elliptic setting, density properties similar to that in Theorem 3.28 are typically referred to as *polyhedricity*-results and well established; see [20, 30, 45, 59]. A major difference between our notion of temporal polyhedricity and the classical concept of polyhedricity is that, in (3.17), the approximation is not performed w.r.t. a mode of convergence that is "natural" for the underlying EVI (V) (this would be convergence in $CBV([0,T]; \mathbb{R}^d)$ or $C([0,T]; \mathbb{R}^d)$ as these are the spaces in which Lipschitz estimates are available for $\mathcal{S}$) but w.r.t. pointwise convergence. Working with this weaker notion of convergence is crucial, given that the directional derivatives of $\mathcal{S}$ are typically neither

left- nor right-continuous and, thus, cannot be approximated with elements of the spaces $BV_r([0,T];\mathbb{R}^d)$, $CBV([0,T];\mathbb{R}^d)$, or $C([0,T];\mathbb{R}^d)$ when convergence w.r.t. the norm $\|\cdot\|_\infty$ is considered.

3.6. Unique characterization of the limit of the difference quotients

We are now in the position to prove that the set $\mathcal{D}$ is a singleton.

Lemma 3.30 (EVI for limits of difference quotients)**.** *For every $\delta \in \mathcal{D}$, it holds*

$$\int_{s_1}^{s_2} \langle z, \mathrm{d}(\delta - h)\rangle + \int_{s_1}^{s_2} \langle \delta, \mathrm{d}h\rangle - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 \geq 0 \qquad \forall\, 0 \leq s_1 < s_2 \leq T \qquad \forall z \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y,u). \tag{3.20}$$

Proof. Let $\delta \in \mathcal{D}$ be given, with a sequence $\{\tau_k\}$ and difference quotients $\delta_k := \delta_{\tau_k}$ as in Definition 3.4. Let $0 \leq s_1 < s_2 \leq T$ and $z \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y,u)$ be fixed. Choose sequences $\{z_{m,l}\} \subset \mathcal{K}^{\mathrm{rad,crit},0}_{BV_r}(y,u)$ and $\{z_l\} \subset \mathcal{K}^{\mathrm{crit}}_{G_r}(y,u)$ for z as in Theorem 3.28. Then it follows from the definition of the set $\mathcal{K}^{\mathrm{rad,crit},0}_{BV_r}(y,u)$ and the convexity of Z that, for all m,l, we have $y(t) + \tau_k z_{m,l}(t) \in Z$ for all $t \in [0,T]$ for all sufficiently large k. This means that $z_{m,l}$ is admissible in the variational inequality (3.2) for δ_k for k large enough and that we have

$$\int_{s_1}^{s_2} \langle z_{m,l} - \delta_k, \mathrm{d}(\delta_k - h)\rangle - \frac{1}{\tau_k}\int_{s_1}^{s_2} \langle z_{m,l} - \delta_k, \mathrm{d}w\rangle \geq 0. \tag{3.21}$$

Due to the inclusion $\mathcal{K}^{\mathrm{rad,crit},0}_{BV_r}(y,u) \subset \mathcal{K}^{\mathrm{crit}}_{G_r}(y,u)$, (3.16), Lemma 2.31, (3.3), the CBV-regularity of δ_k, and the integration by parts formula in Lemma 2.27, (3.21) implies

$$-\int_{s_1}^{s_2} \langle z_{m,l} - \delta_k, \mathrm{d}h\rangle + \langle z_{m,l}(s_2), \delta_k(s_2)\rangle - \langle z_{m,l}(s_1), \delta_k(s_1)\rangle - \int_{s_1}^{s_2} \langle \delta_k, \mathrm{d}z_{m,l}\rangle - \frac{1}{2}|\delta_k(s_2)|^2 + \frac{1}{2}|\delta_k(s_1)|^2 \geq 0$$

for all large enough k. Using Theorem 2.24, Lemma 3.3, and the pointwise convergence $\delta_k(t) \to \delta(t)$ for all $t \in [0,T]$, we can pass to the limit $k \to \infty$ in the above. This yields

$$-\int_{s_1}^{s_2} \langle z_{m,l} - \delta, \mathrm{d}h\rangle + \langle z_{m,l}(s_2), \delta(s_2)\rangle - \langle z_{m,l}(s_1), \delta(s_1)\rangle - \int_{s_1}^{s_2} \langle \delta, \mathrm{d}z_{m,l}\rangle - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 \geq 0.$$

Integrating by parts again (keeping in mind that both the integrand and the integrator may now possess discontinuities), it follows that

$$\begin{aligned} -\int_{s_1}^{s_2} \langle z_{m,l} - \delta, \mathrm{d}h\rangle &+ \int_{s_1}^{s_2} \langle z_{m,l}, \mathrm{d}\delta\rangle - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 \\ &- \sum_{s_1 < t \leq s_2} \langle \delta(t) - \delta(t-), z_{m,l}(t) - z_{m,l}(t-)\rangle + \sum_{s_1 \leq t < s_2} \langle \delta(t) - \delta(t+), z_{m,l}(t) - z_{m,l}(t+)\rangle \geq 0. \end{aligned}$$

Since $z_{m,l}$ is right-continuous with zero left limits at all points of discontinuity, the last inequality simplifies to

$$\begin{aligned} -\int_{s_1}^{s_2} \langle z_{m,l} - \delta, \mathrm{d}h\rangle &+ \int_{s_1}^{s_2} \langle z_{m,l}, \mathrm{d}\delta\rangle - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 \\ &- \sum_{s_1 < t \leq s_2,\ z_{m,l}(t) \neq z_{m,l}(t-)} \langle \delta(t) - \delta(t-), z_{m,l}(t)\rangle \geq 0. \end{aligned}$$

Recall that $\delta(t) = \pi_{\mathcal{Z}(y(t))}(\delta(t-))$ holds for all $t \in [0,T]$ by Theorem 3.16 and that $z_{m,l}(t) \in \mathcal{Z}(y(t))$ by the definition of $\mathcal{K}^{\text{crit}}_{G_r}(y,u)$. In view of (2.10), this yields

$$\sum_{s_1<t\leq s_2,\ z_{m,l}(t)\neq z_{m,l}(t-)} \langle \delta(t)-\delta(t-), z_{m,l}(t)\rangle = \sum_{s_1<t\leq s_2,\ z_{m,l}(t)\neq z_{m,l}(t-)} \left\langle \pi_{\mathcal{Z}(y(t))}(\delta(t-)) - \delta(t-), z_{m,l}(t)\right\rangle \geq 0.$$

We may thus conclude that

$$-\int_{s_1}^{s_2} \langle z_{m,l}-\delta, \mathrm{d}h\rangle + \int_{s_1}^{s_2} \langle z_{m,l}, \mathrm{d}\delta\rangle - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 \geq 0.$$

Letting m go to infinity in the above and then l, using Theorem 2.24 and the properties of $\{z_{m,l}\}$ and $\{z_l\}$, we arrive at the desired variational inequality (3.20). □

Note that, if δ is continuous on $[0,T]$, then Lemmas 2.19 and 2.31 allow to rewrite (3.20) as

$$\int_{s_1}^{s_2} \langle z-\delta, \mathrm{d}(\delta-h)\rangle \geq 0 \qquad \forall\, 0\leq s_1<s_2\leq T \qquad \forall z\in\mathcal{K}^{\text{crit}}_{G_r}(y,u). \tag{3.22}$$

The EVI (3.20) can thus be interpreted as a weak formulation of (3.22); *cf.* [22], Theorem 4.1, System (4.2). By combining (3.20) with the results of the previous subsections, we now arrive at:

Theorem 3.31 (uniqueness of limits of difference quotients)**.** *The set $\mathcal{D}$ is a singleton and its sole element is uniquely characterized by the system*

$$\begin{gathered}\delta\in BV([0,T];\mathbb{R}^d), \qquad \delta_+\in\mathcal{K}^{\text{crit}}_{G_r}(y,u), \qquad \delta(0)=h_0,\\ \delta(t)=\pi_{\mathcal{Z}(y(t))}(\delta(t-))\ \forall t\in[0,T], \qquad \delta(t+)=\pi_{V^+(t)}(\delta(t))\ \forall t\in[0,T],\\ \int_{s_1}^{s_2}\langle z,\mathrm{d}(\delta-h)\rangle + \int_{s_1}^{s_2}\langle\delta,\mathrm{d}h\rangle - \frac{1}{2}|\delta(s_2)|^2+\frac{1}{2}|\delta(s_1)|^2\geq 0 \quad \forall 0\leq s_1<s_2\leq T \quad \forall z\in\mathcal{K}^{\text{crit}}_{G_r}(y,u).\end{gathered} \tag{3.23}$$

Proof. From Lemma 3.5, it follows that $\mathcal{D}$ is nonempty, and from Theorems 3.16 and 3.23, Corollary 3.26, (3.2), and Lemma 3.30, we obtain that all elements of $\mathcal{D}$ satisfy the conditions in (3.23). To establish the assertion of the theorem, it remains to prove that the system (3.23) can have at most one solution. To this end, let us assume that $\delta,\eta\in BV([0,T];\mathbb{R}^d)$ both solve (3.23) and that a number $\gamma>0$ is given. As δ and η are of bounded variation, we can find points $0 =: t_0 < t_1 < \ldots < t_{M-1} < t_M := T$, $M\in\mathbb{N}$, such that

$$4\|\eta\|_\infty \sum_{t\in[0,T]\setminus\{t_0,\ldots,t_M\}} |\delta(t+)-\delta(t-)| \leq \gamma. \tag{3.24}$$

Let $j\in\{1,\ldots,M\}$ be arbitrary and let $s_1,s_2\in[0,T]$ be points such that $t_{j-1}<s_1<s_2<t_j$ holds and such that both δ and η are continuous at s_1 and s_2. (Recall that the set of points, where one of the functions δ and η is discontinuous, is at most countable.) Then, by choosing η_+ as the test function in the variational inequality for δ, by choosing δ_+ as the test function in the variational inequality for η, by adding the resulting inequalities, by exploiting the CBV-regularity of h and Lemma 2.25, by invoking Lemma 2.26, and by using the integration

by parts formula in Lemma 2.27, we obtain that

$$
\begin{aligned}
0 &\leq \int_{s_1}^{s_2} \langle \eta_+, \mathrm{d}\delta \rangle - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 + \int_{s_1}^{s_2} \langle \delta_+, \mathrm{d}\eta \rangle - \frac{1}{2}|\eta(s_2)|^2 + \frac{1}{2}|\eta(s_1)|^2 \\
&= \int_{s_1}^{s_2} \langle \eta_+, \mathrm{d}\delta_+ \rangle + \int_{s_1}^{s_2} \langle \delta_+, \mathrm{d}\eta_+ \rangle - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 - \frac{1}{2}|\eta(s_2)|^2 + \frac{1}{2}|\eta(s_1)|^2 \\
&= \langle \delta(s_2), \eta(s_2) \rangle - \langle \delta(s_1), \eta(s_1) \rangle + \sum_{s_1 < t < s_2} \langle \delta(t+) - \delta(t-), \eta(t+) - \eta(t-) \rangle \\
&\qquad - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 - \frac{1}{2}|\eta(s_2)|^2 + \frac{1}{2}|\eta(s_1)|^2.
\end{aligned}
$$

Using the binomial identities and trivial estimates in the above yields

$$
|\delta(s_2) - \eta(s_2)|^2 \leq |\delta(s_1) - \eta(s_1)|^2 + 4\|\eta\|_\infty \sum_{s_1 < t < s_2} |\delta(t+) - \delta(t-)|
$$

and, after passing to the limits $s_1 \to t_{j-1}$ and $s > s_2 \to s \leq t_j$ for an arbitrary $t_{j-1} < s \leq t_j$ (which is possible due to the density of points of continuity),

$$
|\delta(s-) - \eta(s-)|^2 \leq |\delta(t_{j-1}+) - \eta(t_{j-1}+)|^2 + 4\|\eta\|_\infty \sum_{t_{j-1} < t < t_j} |\delta(t+) - \delta(t-)| \qquad \forall s \in (t_{j-1}, t_j]. \tag{3.25}
$$

By plugging the projection identities in (3.23) into (3.25) and by exploiting (2.4), it now follows that

$$
|\delta(s-) - \eta(s-)|^2 \leq |\delta(t_{j-1}-) - \eta(t_{j-1}-)|^2 + 4\|\eta\|_\infty \sum_{t_{j-1} < t < t_j} |\delta(t+) - \delta(t-)| \quad \forall s \in (t_{j-1}, t_j] \quad \forall j = 1, ..., M,
$$

which, after a trivial induction, yields

$$
|\delta(s-) - \eta(s-)|^2 \leq |h_0 - h_0|^2 + 4\|\eta\|_\infty \sum_{j=1}^{M} \left(\sum_{t_{j-1} < t < t_j} |\delta(t+) - \delta(t-)| \right) \leq \gamma \qquad \forall s \in (0, T].
$$

Here, in the last step, we have used (3.24). As $\gamma > 0$ was arbitrary, the above shows that $\delta_- = \eta_-$ holds in $(0, T]$. Due to the initial condition and the projection identities in (3.23), this implies $\eta = \delta$ as desired. This shows that $\mathcal{D}$ is indeed a singleton and completes the proof. □

As already mentioned in Section 3.1, with Theorem 3.31 at hand, assertion (I) of Theorem 1.1 follows immediately. For the sake of completeness, we state this result in the following theorem in a formulation that is independent of Assumption 3.1.

Theorem 3.32 (point (I) of Thm. 1.1)**.** *Let $T > 0$ be given and let $Z \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a full-dimensional convex non-obtuse polyhedron. Then the solution map $\mathcal{S}\colon CBV([0,T];\mathbb{R}^d) \times Z \to CBV([0,T];\mathbb{R}^d)$, $(u, y_0) \mapsto y$, of (V) is directionally differentiable in the sense that, for all tuples $(u, y_0) \in CBV([0,T];\mathbb{R}^d) \times Z$ and all directions $(h, h_0) \in CBV([0,T];\mathbb{R}^d) \times \mathbb{R}^d$ satisfying $y_0 + \tau_0 h_0 \in Z$ for some $\tau_0 > 0$, there exists a unique $\delta := \mathcal{S}'((u, y_0); (h, h_0)) \in BV([0,T];\mathbb{R}^d)$ satisfying*

$$
\lim_{(0,\tau_0) \ni \tau \to 0} \frac{\mathcal{S}(u + \tau h, y_0 + \tau h_0)(t) - \mathcal{S}(u, y_0)(t)}{\tau} = \delta(t) \qquad \forall t \in [0, T]. \tag{3.26}
$$

Moreover, the directional derivative δ of $\mathcal{S}$ at (u, y_0) in direction (h, h_0) is uniquely characterized by the system (3.23), with the set $\mathcal{K}^{\mathrm{crit}}_{G_r}(y, u)$ defined by (3.16) with $y := \mathcal{S}(u, y_0)$ and $w := \mathcal{P}(u, y_0)$ and the spaces $V^{\pm}(t)$ defined as in Definition 3.10 w.r.t. an arbitrary but fixed standard description $\{(\nu_i, \alpha_i)\}_{i\in I}$ of Z.

Proof. If (u, y_0) and (h, h_0) are tuples as in the assertion of the theorem, then Assumption 3.1 holds and it follows from the definition of the set $\mathcal{D}$ in (3.4), a trivial contradiction argument based on Theorems 2.29 and 2.37, and Theorem 3.31 that $\mathcal{S}$ is pointwisely directionally differentiable at (u, y_0) in direction (h, h_0) in the sense of (3.26). That the directional derivative $\mathcal{S}'((u, y_0); (h, h_0))$ is uniquely characterized by the system (3.23) is also a straightforward consequence of Theorem 3.31. This completes the proof. □

3.7. Relation to the known one-dimensional results

We conclude Section 3 by discussing how Theorem 3.32 relates to the results for the one-dimensional play and stop proved in [15–17, 48]. In all what follows, we still assume that d, T, Z, $\{(\nu_i, \alpha_i)\}_{i\in I}$, y_0, u, y, w, h_0, and h are as in Assumption 3.1. Note that, in the case $d = 1$, every full-dimensional convex polyhedron Z is of the form $[a, b]$, $(-\infty, b]$, $[a, \infty)$, or $\mathbb{R}$ with $-\infty < a < b < \infty$. In particular, every such polyhedron is non-obtuse with a standard description involving (at most) the outer unit normals 1 and -1; see Definition 2.12. We begin with an observation on the jumps of the directional derivatives in the one-dimensional case.

Corollary 3.33 (jumps in 1d)**.** *If $d = 1$ holds, then the directional derivative $\delta := \mathcal{S}'((u, y_0); (h, h_0))$ satisfies the following for all $t \in [0, T]$.*

i) $\delta(t) \in \{\delta(t+), \delta(t-)\}$;
ii) $\delta(t+) = 0$ *if* $\delta(t+) \neq \delta(t-)$;
iii) $\delta(t+)(\delta(t+) - \delta(t-)) = \delta(t+)(\delta(t+) - \delta(t)) = \delta(t)(\delta(t) - \delta(t-)) = 0$;
iv) $z(\delta(t) - \delta(t-)) \geq 0$ *for all* $z \in \mathcal{Z}(y(t))$;
v) $z(\delta(t+) - \delta(t)) = 0$ *for all* $z \in V^+(t)$;
vi) $z(\delta(t+) - \delta(t-)) \geq 0$ *for all* $z \in \mathcal{Z}(y(t)) \cap V^+(t)$.

Proof. For $d = 1$, we have $\mathcal{Z}(y(t)) \in \{(-\infty, 0], [0, \infty), \mathbb{R}\}$ and $V^+(t) \in \{\{0\}, \mathbb{R}\}$ for all $t \in [0, T]$. This implies that, for all $t \in [0, T]$, all $x \in \mathbb{R}$, and all $F \in \{\pi_{\mathcal{Z}(y(t))}, \pi_{V^+(t)}\}$, we have $F(x) \in \{0, x\}$ and $F(0) = 0$. If we combine these observations with (2.10) and the jump conditions $\delta(t) = \pi_{\mathcal{Z}(y(t))}(\delta(t-))$ and $\delta(t+) = \pi_{V^+(t)}(\delta(t))$ from Theorems 3.16 and 3.23, then the assertions in i) to vi) follow immediately. □

Corollary 3.33i) demonstrates that our results are consistent with [17], Theorem 2.1, which establishes that $\mathcal{S}'((u, y_0); (h, h_0)) \in BV_{rl}([0, T])$ holds for the one-dimensional stop $\mathcal{S}$. Points ii), iii), and vi) of Corollary 3.33 further reproduce the observations made in [48], Corollary 4.12; *cf.* also the proof of [17], Theorem 2.1. We remark that BV_{rl}-regularity cannot be expected for $\mathcal{S}'((u, y_0); (h, h_0))$ if d is greater than one; see Example 3.36 below. Next, we study how the system (3.23) is related to the EVIs that have been derived for the directional derivatives of the one-dimensional stop $\mathcal{S}$ and their right limits in [17], Theorem 2.1 and [48], Corollary 4.13.

Corollary 3.34 (EVI for the right limits of the directional derivatives in 1d)**.** *If $d = 1$ holds, then the right-limit function $\eta := \mathcal{S}'((u, y_0); (h, h_0))_+ \in BV_r([0, T])$ is the unique solution of the variational inequality*

$$\begin{gathered} \eta \in BV_r([0, T]), \qquad \eta \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y, u), \qquad \eta(0) = \pi_{V^+(0)}(h_0), \\ \int_0^T (z - \eta)\mathrm{d}(\eta - h) \geq 0 \quad \forall z \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y, u). \end{gathered} \tag{3.27}$$

Further, for all $s \in (0, T]$, it holds

$$\int_0^s (z - \eta)\mathrm{d}(\eta - h) \geq 0 \quad \forall z \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y, u). \tag{3.28}$$

Proof. Define $\delta := \mathcal{S}'((u, y_0); (h, h_0))$ and $\eta := \delta_+$. We first prove that η solves (3.27) and (3.28). The first line of (3.27) is a straightforward consequence of (3.23). To establish the second line and (3.28), we proceed similarly to the proof of Theorem 3.31. Let $s \in (0, T]$ and $z \in \mathcal{K}^{\text{crit}}_{G_r}(y, u)$ be given and let $\gamma > 0$ be arbitrarily small. As η has bounded variation, we can find points $0 =: t_0 < t_1 < ... < t_{M-1} < t_M := s$, $M \in \mathbb{N}$, such that

$$\|\eta\|_\infty \sum_{t \in [0,s] \setminus \{t_0,...,t_M\}} |\eta(t) - \eta(t-)| \leq \gamma. \tag{3.29}$$

Let $j \in \{1, ..., M\}$ be arbitrary and let $s_1, s_2 \in [0, s]$ be points such that $t_{j-1} < s_1 < s_2 < t_j$ holds and such that both η and δ are continuous at s_1 and s_2. Then it follows from (3.23), Lemmas 2.25 and 2.26, the identities $\delta(s_1) = \delta(s_1+) = \eta(s_1)$ and $\delta(s_2) = \delta(s_2+) = \eta(s_2)$, and Lemma 2.27 (with $f = g = \eta$) that

$$\begin{aligned}
0 &\leq \int_{s_1}^{s_2} \langle z, \mathrm{d}(\delta - h)\rangle + \int_{s_1}^{s_2} \langle \delta, \mathrm{d}h\rangle - \frac{1}{2}|\delta(s_2)|^2 + \frac{1}{2}|\delta(s_1)|^2 \\
&= \int_{s_1}^{s_2} \langle z, \mathrm{d}(\eta - h)\rangle + \int_{s_1}^{s_2} \langle \eta, \mathrm{d}h\rangle - \frac{1}{2}|\eta(s_2)|^2 + \frac{1}{2}|\eta(s_1)|^2 \\
&= \int_{s_1}^{s_2} \langle z - \eta, \mathrm{d}(\eta - h)\rangle + \frac{1}{2} \sum_{s_1 < t < s_2} |\eta(t) - \eta(t-)|^2 \\
&= \int_{t_{j-1}}^{t_j} \mathbb{1}_{[s_1,s_2]}(z - \eta)\mathrm{d}(\eta - h) + \frac{1}{2} \sum_{s_1 < t < s_2} |\eta(t) - \eta(t-)|^2.
\end{aligned} \tag{3.30}$$

As δ and η possess at most countably many points of discontinuity, we may let s_1 and s_2 go to t_{j-1} and t_j, respectively, and invoke the bounded convergence theorem to obtain the following from (3.30):

$$0 \leq \int_{t_{j-1}}^{t_j} \mathbb{1}_{(t_{j-1},t_j)}(z - \eta)\mathrm{d}(\eta - h) + \|\eta\|_\infty \sum_{t_{j-1} < t < t_j} |\eta(t) - \eta(t-)|. \tag{3.31}$$

Note that Lemma 2.21, the right-continuity of η, (2.1), points iii) and vi) of Corollary 3.33, and the inclusion $z(t_j) = z(t_j+) \in \mathcal{Z}(y(t_j)) \cap V^+(t_j)$ obtained from (3.16) and $z \in \mathcal{K}^{\text{crit}}_{G_r}(y, u)$ imply

$$\begin{aligned}
\int_{t_{j-1}}^{t_j} (\mathbb{1}_{\{t_j\}} + \mathbb{1}_{\{t_{j-1}\}})(z - \eta)\mathrm{d}(\eta - h) &= z(t_j)(\eta(t_j) - \eta(t_j-)) - \eta(t_j)(\eta(t_j) - \eta(t_j-)) \\
&= z(t_j)(\delta(t_j+) - \delta(t_j-)) - \delta(t_j+)(\delta(t_j+) - \delta(t_j-)) \\
&\geq 0.
\end{aligned} \tag{3.32}$$

In combination with (3.31) and Lemma 2.19, this yields

$$0 \leq \int_{t_{j-1}}^{t_j} (z - \eta)\mathrm{d}(\eta - h) + \|\eta\|_\infty \sum_{t_{j-1} < t < t_j} |\eta(t) - \eta(t-)|$$

and, after summing over j and using (3.29) and Lemma 2.19,

$$0 \leq \sum_{j=1}^{M} \int_{t_{j-1}}^{t_j} (z - \eta)\mathrm{d}(\eta - h) + \|\eta\|_\infty \sum_{t_{j-1} < t < t_j} |\eta(t) - \eta(t-)| \leq \int_0^s (z - \eta)\mathrm{d}(\eta - h) + \gamma. \tag{3.33}$$

Since γ was arbitrary, this proves that η satisfies (3.28) for all $s \in (0, T]$ and the second line of (3.27).

It remains to prove that (3.27) has at most one solution. To this end, let us assume that η_1 and η_2 solve (3.27). By choosing η_1 as the test function in the EVI for η_2 and vice versa, by adding the resulting inequalities, by again invoking Lemma 2.27 (with $f = g = \eta_1 - \eta_2$), and by exploiting that $\eta_1(0) = \eta_2(0)$, we obtain that

$$0 \leq -\int_0^T (\eta_1 - \eta_2)\mathrm{d}(\eta_1 - \eta_2) = -\frac{1}{2}\left(|(\eta_1 - \eta_2)(T)|^2 + \sum_{0<t\leq T} |(\eta_1 - \eta_2)(t) - (\eta_1 - \eta_2)(t-)|^2\right).$$

The above implies $\eta_1(T) = \eta_2(T)$ and, since $\eta_1 - \eta_2$ is in $BV_r([0,T])$, $\eta_1 - \eta_2 \in CBV([0,T])$. Next, we choose for an arbitrary but fixed $s \in (0,T)$ the function $z := \mathbb{1}_{[0,s)}\eta_2 + \mathbb{1}_{[s,T]}\eta_1$ as the test function in the EVI for η_1 and $z := \mathbb{1}_{[0,s)}\eta_1 + \mathbb{1}_{[s,T]}\eta_2$ as the test function in the EVI for η_2 and add the resulting inequalities. (Note that these functions are indeed elements of $\mathcal{K}^{\mathrm{crit}}_{G_r}(y,u)$, see the last equality in (3.16).) In combination with the continuity of $\eta_1 - \eta_2$, again Lemmas 2.19, 2.27 and 2.26, and the initial conditions, this yields

$$0 \leq \int_0^T \mathbb{1}_{[0,s)}(\eta_2 - \eta_1)\mathrm{d}(\eta_1 - \eta_2) = \int_0^s (\eta_2 - \eta_1)\mathrm{d}(\eta_1 - \eta_2) = -\frac{1}{2}|\eta_1(s) - \eta_2(s)|^2.$$

In combination with what we have proven before and the initial condition in (3.27), this shows that $\eta_1 = \eta_2$ holds on $[0,T]$. Thus, (3.27) has precisely one solution, namely $\mathcal{S}'((u,y_0);(h,h_0))_+$, and the proof is complete. □

Note that the formulas for $\mathcal{K}^{\mathrm{crit}}_{G_r}(y,u)$ in (3.16) and [48], Proposition 4.9 imply that, if $d = 1$ holds and $K^{\mathrm{ptw}}_{\mathrm{crit}}(y,u)\colon [0,T] \rightrightarrows \mathbb{R}$ is defined as in [48], Definition 4.7 (*i.e.*, $K^{\mathrm{ptw}}_{\mathrm{crit}}(y,u)(t) := \mathcal{Z}(y(t)) \cap V^+(t)$), then we have

$$z \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y,u) \quad \Longleftrightarrow \quad z \in G\left([0,T]; K^{\mathrm{ptw}}_{\mathrm{crit}}(y,u)\right) \cap G_r([0,T]). \tag{3.34}$$

This shows that Corollary 3.34 reproduces the EVI derived in [48], Corollary 4.13, albeit with the modification that only right-continuous test functions are considered and that h_0 is also allowed to be nonzero. (We expect that, by exploiting (3.34) and approximation arguments analogous to those in the proofs of Theorems 3.11 and 3.28, one can also get rid of the right-continuity of the test functions in Corollary 3.34. We omit discussing this topic in detail in this paper.)

To obtain a system of EVIs that characterizes the directional derivative $\delta := \mathcal{S}'((u,y_0);(h,h_0))$ itself, we can exploit the evolutionary nature of (V). This nature implies that, by restricting the function $\mathcal{S}(u,y_0)$ to the interval $[0,s]$, $0 < s < T$, we obtain precisely the solution of the variational inequality that arises when the terminal time T in (V) is replaced by s. For the directional derivative δ, this means that δ_+ not only satisfies the EVI (3.27) in the case $d = 1$, but also analogue EVIs formulated on the intervals $[0,s]$, $0 < s < T$. As (3.27) uniquely characterizes $\delta_+(T) = \delta(T)$ (due to the conventions for the right limit at the interval endpoint), the corresponding EVIs on the intervals $[0,s]$ uniquely characterize the values $\delta(s)$ for all $0 < s < T$ and, thus, δ in its entirety. Rigorously, we can formulate this result as follows:

Corollary 3.35 (EVI-system for the directional derivatives in 1d)**.** *If $d = 1$ holds, then the directional derivative $\delta := \mathcal{S}'((u,y_0);(h,h_0))$ is the unique solution of the EVI-system*

$$\begin{aligned} &\delta \in BV([0,T]), && \delta(0) = h_0, \\ &\delta_{+,[0,s]} \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y,u)[0,s], && \int_0^s (z - \delta_{+,[0,s]})\mathrm{d}(\delta - h) \geq 0 \quad \forall z \in \mathcal{K}^{\mathrm{crit}}_{G_r}(y,u)[0,s] \quad \forall s \in (0,T]. \end{aligned} \tag{3.35}$$

Here, $\delta_{+,[0,s]}$ and $\mathcal{K}_{G_r}^{\mathrm{crit}}(y,u)[0,s]$ denote the right-limit function and the critical cone from (3.16) associated with the interval $[0,s]$, respectively, with the corresponding conventions for the interval endpoint s, i.e.,

$$\begin{aligned}\delta_{+,[0,s]}(t) &:= \begin{cases}\delta_+(t) & \text{if } 0 \le t < s,\\ \delta(s) & \text{if } t = s,\end{cases}\\ \mathcal{K}_{G_r}^{\mathrm{crit}}(y,u)[0,s] &:= \left\{ z \in G_r([0,s]) \;\middle|\; z(t) \in \mathcal{Z}(y(t))\ \forall t \in [0,s] \text{ and } \int_0^s z\,\mathrm{d}w = 0 \right\}.\end{aligned} \tag{3.36}$$

Proof. From Corollary 3.34, the arguments outlined above, and (3.23), we obtain that δ satisfies $\delta \in BV([0,T])$, $\delta(0) = h_0$, and

$$\delta_{+,[0,s]} \in \mathcal{K}_{G_r}^{\mathrm{crit}}(y,u)[0,s], \quad \int_0^s (z - \delta_{+,[0,s]})\mathrm{d}(\delta_{+,[0,s]} - h) \ge 0 \quad \forall z \in \mathcal{K}_{G_r}^{\mathrm{crit}}(y,u)[0,s] \quad \forall s \in (0,T].$$

Using Lemmas 2.19, 2.21, and 2.26, (2.1), (3.36), (2.3), (3.27), and (3.16), one further easily checks that

$$\begin{aligned}\int_0^s (z - \delta_{+,[0,s]})\mathrm{d}(\delta_{+,[0,s]} - \delta) &= \int_0^s (z - \delta_{+,[0,s]})(\mathbb{1}_{\{0\}} + \mathbb{1}_{\{s\}})\mathrm{d}(\delta_{+,[0,s]} - \delta)\\ &= \big(z(0) - \delta_{+,[0,s]}(0)\big)\big(\delta_{+,[0,s]}(0+) - \delta(0+) - \delta_{+,[0,s]}(0) + \delta(0)\big)\\ &\qquad + \big(z(s) - \delta_{+,[0,s]}(s)\big)\big(\delta_{+,[0,s]}(s) - \delta(s) - \delta_{+,[0,s]}(s-) + \delta(s-)\big)\\ &= (z(0) - \delta(0+))(\delta(0) - \delta(0+))\\ &= \big(z(0) - \pi_{V^+(0)}(h_0)\big)\big(h_0 - \pi_{V^+(0)}(h_0)\big)\\ &= 0 \qquad \forall z \in \mathcal{K}_{G_r}^{\mathrm{crit}}(y,u)[0,s].\end{aligned}$$

By combining all of the above and by exploiting Lemma 2.19, we obtain that δ is indeed a solution of (3.35). To see that the system (3.35) possesses at most one solution, we can proceed along the lines of the proof of Corollary 3.34. If δ and $\hat{\delta}$ both solve (3.35), then we trivially have $\delta(0) = \hat{\delta}(0)$ and, for every $s \in (0,T]$, we obtain from the EVIs satisfied by δ and $\hat{\delta}$, the initial condition in (3.35), and [17], Proposition 6.2 (which is an easy consequence of Lemmas 2.21, 2.26, and 2.27) that

$$0 \le \int_0^s (\hat{\delta}_{+,[0,s]} - \delta_{+,[0,s]})\mathrm{d}(\delta - \hat{\delta}) \le -\frac{1}{2}|\hat{\delta}_{+,[0,s]}(s) - \delta_{+,[0,s]}(s)|^2 = -\frac{1}{2}|\hat{\delta}(s) - \delta(s)|^2. \tag{3.37}$$

The above implies $\hat{\delta}(s) = \delta(s)$ and, consequently, $\hat{\delta} = \delta$ on $[0,T]$. This completes the proof of the corollary. ☐

The system (3.35) corresponds to the one in [48], Theorem 4.11 and [17], Theorem 2.1, Equation (2.6), as one may easily check by means of the equivalence in (3.34). We would like to emphasize that one has to state the system (3.35) very carefully to get a unique characterization for δ. In particular, it is not enough to merely demand $\delta_+ \in \mathcal{K}_{G_r}^{\mathrm{crit}}(y,u)$ in (3.35), analogously to the EVI (3.27), as this condition misses the information $\delta(s) \in \mathcal{Z}(y(s))$ for all $s \in (0,T]$ that is encoded in the inclusions $\delta_{+,[0,s]} \in \mathcal{K}_{G_r}^{\mathrm{crit}}(y,u)[0,s]$ for all $s \in (0,T]$ *via* the endpoint convention in (3.36) and that is essential for the estimate (3.37).

As a final remark, we would like to point out that the proof of Corollary 3.34 cannot be extended (easily) to the case $d > 1$. The problem is that, for $d > 1$, it is not necessarily true that the directional derivative $\delta := \mathcal{S}'((u,y_0);(h,h_0))$ satisfies $\langle \delta(t+), \delta(t+) - \delta(t-)\rangle \le 0$ for all $t \in [0,T]$, as used in the final step of (3.32); *cf.* Corollary 3.33iii). This makes it (seemingly) impossible to correct the function values of the integrand in (3.31) at the right endpoints t_j of the decomposition intervals $[t_{j-1}, t_j]$ used in (3.29) and, thus, to combine the integrals over the subintervals $[t_{j-1}, t_j]$ to a meaningful global EVI, as done in (3.33). Note that this problem is also intrinsically related to the fact that the directional derivatives of $\mathcal{S}$ are typically not in $BV_{rl}([0,T];\mathbb{R}^d)$ if $d>1$ holds. In view of these effects, it seems likely that the characterization of δ by means of the system (3.23) is the best that one can expect for dimensions $d>1$ and that it is not possible in general to characterize δ *via* a classical EVI without separate jump conditions. We conclude Section 3 with an example which demonstrates that situations with $\langle \delta(t+), \delta(t+)-\delta(t-)\rangle > 0$ and $\delta \notin BV_{rl}([0,T];\mathbb{R}^d)$ are indeed possible for $d>1$.

Example 3.36 (jumps in a 2d situation)**.** Consider the case

$$d=2, \qquad T=2, \qquad Z=\{x\in\mathbb{R}^2 \mid 0\le x_2\le x_1\}, \qquad h\equiv 0,$$
$$y_0=\begin{pmatrix}2\\1\end{pmatrix}, \qquad h_0=\begin{pmatrix}0\\2\end{pmatrix}, \qquad u\in W^{1,\infty}((0,2);\mathbb{R}^2), \quad u(t)=\mathbb{1}_{[0,1]}(t)\begin{pmatrix}2\\1\end{pmatrix}(1-t)+\mathbb{1}_{(1,2]}(t)\begin{pmatrix}1\\-1\end{pmatrix}(t-1).$$

For the above data, one easily verifies that the conditions in Assumption 3.1 hold with $\tau_0=1/2$ and that the solution $y:=\mathcal{S}(u,y_0)$ of (V) and the associated play $w:=u-y$ are given by

$$y,w\in W^{1,\infty}((0,2);\mathbb{R}^2), \qquad y(t)=\mathbb{1}_{[0,1]}(t)\begin{pmatrix}2\\1\end{pmatrix}(1-t)+\mathbb{1}_{(1,2]}(t)\begin{pmatrix}1\\0\end{pmatrix}(t-1), \qquad w(t)=\mathbb{1}_{(1,2]}(t)\begin{pmatrix}0\\-1\end{pmatrix}(t-1).$$

Using the calculus rules from Section 2.4, one also easily checks that the system (3.23) is solved by the function

$$\delta(t):=\mathbb{1}_{[0,1)}(t)\begin{pmatrix}0\\2\end{pmatrix}+\mathbb{1}_{\{1\}}(t)\begin{pmatrix}1\\1\end{pmatrix}+\mathbb{1}_{(1,2]}(t)\begin{pmatrix}1\\0\end{pmatrix} \tag{3.38}$$

in this situation. As (3.23) is uniquely solvable by Theorem 3.31, this implies $\delta=\mathcal{S}'((u,y_0);(h,h_0))$. As δ is clearly neither left- nor right-continuous at $t=1$, this demonstrates that there are indeed situations with $\mathcal{S}'((u,y_0);(h,h_0))\notin BV_{rl}([0,T];\mathbb{R}^d)$. From (3.38), we further obtain that $\langle\delta(1+),\delta(1+)-\delta(1-)\rangle=1>0$. The inequality $\langle\delta(t+),\delta(t+)-\delta(t-)\rangle\le 0$ that is valid for $d=1$ by Corollary 3.33iii) and that entered the proof of Corollary 3.34 crucially in (3.32) is thus indeed in general wrong for $d>1$, as mentioned after Corollary 3.35.

4. Proof of (II): non-differentiability in the obtuse case

It remains to prove part (II) of Theorem 1.1. To establish this second main result of our analysis, we use Proposition 2.16 to reduce the situation to two dimensions and then explicitly construct a counterexample.

Theorem 4.1 (point (II) of Thm. 1.1)**.** *Let $T>0$ be given and let $Z\subset\mathbb{R}^d$, $d\in\mathbb{N}$, be a full-dimensional convex polyhedron that is not non-obtuse. Let $\vartheta\in(0,T)$ be an arbitrary number. Then there exist $y_0\in Z$, $u\in W^{1,\infty}((0,T);\mathbb{R}^d)$, $h_0\in\mathcal{Z}(y_0)$, $\tau_0>0$, and $h\in C^\infty([0,T];\mathbb{R}^d)$ such that $y_0+\tau h_0\in Z$ holds for all $\tau\in[0,\tau_0]$, such that the solution operator $\mathcal{S}\colon CBV([0,T];\mathbb{R}^d)\times Z\to CBV([0,T];\mathbb{R}^d)$ of (V) satisfies*

$$\left.\frac{\mathcal{S}(u,y_0+\tau h_0)-\mathcal{S}(u,y_0)}{\tau}\right|_{[\vartheta,T]}=\left.\frac{\mathcal{S}(u+\tau h,y_0)-\mathcal{S}(u,y_0)}{\tau}\right|_{[\vartheta,T]}=\text{const}=c_\tau \qquad \forall\, 0<\tau\le\tau_0$$

for some $c_\tau\in\mathbb{R}^d$, and such that the limit $\lim_{(0,\tau_0)\ni\tau\to 0}c_\tau$ does not exist.

Proof. Let T and Z be as in the theorem, let $\{(\nu_i,\alpha_i)\}_{i\in I}$ be a standard description of Z, and let $\vartheta\in(0,T)$ be given. From Proposition 2.16, we obtain that there exists $w\in Z$ such that $\mathcal{A}(w)=\{i_1,i_2\}$ holds for some $i_1,i_2\in I$ with $\langle\nu_{i_1},\nu_{i_2}\rangle>0$. We assume w.l.o.g. that $w=0$, $\alpha_{i_1}=\alpha_{i_2}=0$, and $i_1=1$, $i_2=2$. (This can always be achieved by translating Z and reordering $\{(\nu_i,\alpha_i)\}_{i\in I}$.) As $\mathcal{A}(0)=\{1,2\}$, we can find $\varepsilon\in(0,\vartheta)$ such that $\langle\nu_i,x\rangle<\alpha_i$ holds for all $i\in I\setminus\{1,2\}$ and all $x\in B_\varepsilon(0)$. Since $\langle\nu_1,\nu_2\rangle>0$ and $\nu_1\ne\nu_2$, there further exist unique $e_1,e_2\in\mathbb{R}^d$ satisfying $|e_1|=|e_2|=1$, $e_1,e_2\in\operatorname{span}(\{\nu_1,\nu_2\})$, $\langle\nu_1,e_1\rangle=\langle\nu_2,e_2\rangle=0$, $\langle\nu_1,e_2\rangle<0$, $\langle\nu_2,e_1\rangle<0$, and $\langle e_1, e_2\rangle < 0$ (*cf.* Lem. 2.10iv)). Consider now an arbitrary but fixed number $\gamma \in [0,1]$ and choose functions $\varphi \in C_c^\infty(\mathbb{R})$ and $\beta_1 \in L^\infty(0,T)$ such that

$$\operatorname{supp}(\varphi) \subset \left(0, \frac{\varepsilon}{2}\right), \qquad \varphi \geq 0, \qquad \int_0^T \varphi \,\mathrm{d}s = 1, \qquad \beta_1 = 0 \text{ a.e. in } (\varepsilon, T), \qquad 0 \leq \beta_1 \leq \frac{\langle e_1, e_2\rangle}{\langle \nu_1, e_2\rangle} \text{ a.e. in } (0,\varepsilon),$$

and such that the limit

$$\lim_{(0,\varepsilon)\ni\tau\to 0} \frac{1}{\tau}\int_{\varepsilon-\tau}^{\varepsilon} \beta_1(s)\mathrm{d}s \tag{4.1}$$

does not exist. (Functions φ and β_1 with these properties can be constructed easily.) We define

$$y_0 := \varepsilon e_1, \qquad \tau_0 := \frac{\varepsilon}{2}, \qquad h_0 := -\gamma e_1,$$
$$\lambda_1(t) := \int_0^t \beta_1(s)\mathrm{d}s, \qquad u(t) := \max(0, \varepsilon - t)e_1 + \lambda_1(t)\nu_1, \qquad h(t) := -(1-\gamma)\int_0^t \varphi(s)\mathrm{d}s\, e_1.$$

For this data, one easily checks that $y_0 + \tau h_0 \in Z$ for all $\tau \in [0,\tau_0]$, $u \in W^{1,\infty}((0,T);\mathbb{R}^d)$, and $h \in C^\infty([0,T];\mathbb{R}^d)$. We claim that $\mathcal{S}(u+\tau h, y_0 + \tau h_0)$ is equal to the function

$$y^\tau(t) := \begin{cases} (\varepsilon - \gamma\tau - t)e_1 + \tau h(t) & \text{if } t \in [0, \varepsilon-\tau], \\ \displaystyle\int_{\varepsilon-\tau}^t \langle \dot u(s), e_2\rangle \,\mathrm{d}s\, e_2 & \text{if } t \in (\varepsilon-\tau, T], \end{cases} \tag{4.2}$$

for all $\tau \in [0,\tau_0]$. To see this, let $\tau \in [0,\tau_0]$ be fixed. Due to (4.2) and $0 \leq \tau \leq \tau_0 < \varepsilon$, we trivially have $y^\tau(0) = (\varepsilon - \gamma\tau)e_1 = y_0 + \tau h_0$. From the properties of φ and the definition $\tau_0 := \varepsilon/2$, we further obtain that

$$\begin{aligned}(\varepsilon - \gamma\tau - (\varepsilon-\tau))e_1 + \tau h(\varepsilon - \tau) &= (1-\gamma)\tau e_1 - (1-\gamma)\tau\int_0^{\varepsilon-\tau}\varphi(s)\mathrm{d}s\, e_1 \\ &= (1-\gamma)\tau e_1 - (1-\gamma)\tau \int_0^{\varepsilon/2}\varphi(s)\mathrm{d}s\, e_1 = 0 = \int_{\varepsilon-\tau}^{\varepsilon-\tau}\langle \dot u(s), e_2\rangle\,\mathrm{d}s\, e_2.\end{aligned}$$

In combination with (4.2), this yields $y^\tau \in W^{1,\infty}((0,T);\mathbb{R}^d)$. Similarly, one also checks that

$$0 \leq \varepsilon - \gamma\tau - t - (1-\gamma)\tau\int_0^t \varphi(s)\mathrm{d}s \leq \varepsilon - \gamma\tau \leq \varepsilon \qquad \forall t \in [0, \varepsilon-\tau]$$

and, thus,

$$y^\tau(t) = (\varepsilon - \gamma\tau - t)\, e_1 + \tau h(t) = \left(\varepsilon - \gamma\tau - t - (1-\gamma)\tau\int_0^t\varphi(s)\mathrm{d}s\right) e_1 \in \operatorname{conv}(\{0, \varepsilon e_1\}) \subset Z \qquad \forall t \in [0, \varepsilon-\tau].$$

For the derivative $\dot y^\tau$, we obtain

$$\dot y^\tau(t) - (\dot u + \tau\dot h)(t) = -e_1 + \tau \dot h(t) - \left(-e_1 + \dot\lambda_1(t)\nu_1 + \tau\dot h(t)\right) = -\beta_1(t)\nu_1 \in -\mathcal{N}_Z(y^\tau(t)) \text{ for a.a. } t \in (0,\varepsilon-\tau).$$

Thus, $y^\tau(0) = y_0 + \tau h_0$, $y^\tau(t) \in Z$ for all $t \in [0,\varepsilon-\tau]$, and $\dot y^\tau(t) - (\dot u + \tau\dot h)(t) \in -\mathcal{N}_Z(y^\tau(t))$ for a.a. $t \in (0,\varepsilon-\tau)$. To prove that $y^\tau = \mathcal{S}(u+\tau h, y_0+\tau h_0)$ holds, it remains to check that the last inclusions also hold on the interval

$(\varepsilon - \tau, T]$; see Proposition 2.33. To this end, we note that our assumptions on β_1 yield

$$\langle \dot{u}(t), e_2 \rangle = \mathbb{1}_{[0,\varepsilon]}(t) \langle -e_1, e_2 \rangle + \beta_1(t) \langle \nu_1, e_2 \rangle = \left(\mathbb{1}_{[0,\varepsilon]}(t) - \frac{\langle \nu_1, e_2 \rangle}{\langle e_1, e_2 \rangle} \beta_1(t) \right) \langle -e_1, e_2 \rangle \in [0, 1] \text{ for a.a. } t \in (\varepsilon - \tau, T). \tag{4.3}$$

This implies

$$y^\tau(t) = \int_{\varepsilon-\tau}^{t} \langle \dot{u}(s), e_2 \rangle \, \mathrm{d}s \, e_2 = \int_{\varepsilon-\tau}^{\min(\varepsilon,t)} \langle \dot{u}(s), e_2 \rangle \, \mathrm{d}s \, e_2 \in \operatorname{conv}(\{0, \tau e_2\}) \subset \operatorname{conv}(\{0, \varepsilon e_2\}) \subset Z \quad \forall t \in (\varepsilon - \tau, T]. \tag{4.4}$$

Similarly, we may also compute that

$$\langle \dot{u}(t), \nu_2 \rangle = \mathbb{1}_{[0,\varepsilon]}(t) \langle -e_1, \nu_2 \rangle + \beta_1(t) \langle \nu_1, \nu_2 \rangle \geq 0 \text{ for a.a. } t \in (\varepsilon - \tau, T).$$

In combination with the properties of φ, the estimate $\varepsilon - \tau \geq \varepsilon/2$, the inclusion $y^\tau(t) \in \operatorname{conv}(\{0, \varepsilon e_2\})$ for all $t \in (\varepsilon - \tau, T]$ in (4.4), and the fact that e_2 and ν_2 form an orthonormal basis of $\operatorname{span}(\{e_1, e_2\}) = \operatorname{span}(\{\nu_1, \nu_2\})$, this implies

$$\dot{y}^\tau(t) - (\dot{u} + \tau \dot{h})(t) = \dot{y}^\tau(t) - \dot{u}(t) = \langle \dot{u}(t), e_2 \rangle \, e_2 - \dot{u}(t) = - \langle \dot{u}(t), \nu_2 \rangle \, \nu_2 \in -\mathcal{N}_Z(y^\tau(t)) \text{ for a.a. } t \in (\varepsilon - \tau, T).$$

We thus indeed have $y^\tau(0) = y_0 + \tau h_0$, $y^\tau(t) \in Z$ for all $t \in [0, T]$, and $\dot{y}^\tau(t) - (\dot{u} + \tau \dot{h})(t) \in -\mathcal{N}_Z(y^\tau(t))$ for a.a. $t \in (0, T)$ and, as a consequence, $y^\tau = \mathcal{S}(u + \tau h, y_0 + \tau h_0)$ as claimed. Note that, from (4.2), (4.3), and $\beta_1 = 0$ a.e. in (ε, T), we obtain that

$$\left. \frac{y^\tau - y^0}{\tau} \right|_{[\varepsilon,T]} = \left. \frac{\mathcal{S}(u + \tau h, y_0 + \tau h_0) - \mathcal{S}(u, y_0)}{\tau} \right|_{[\varepsilon,T]} = \text{const} \qquad \forall\, 0 < \tau \leq \tau_0 \tag{4.5}$$

and

$$\frac{y^\tau(\varepsilon) - y^0(\varepsilon)}{\tau} = \frac{1}{\tau} \int_{\varepsilon-\tau}^{\varepsilon} \langle \dot{u}(s), e_2 \rangle \, \mathrm{d}s \, e_2 = \langle -e_1, e_2 \rangle \, e_2 + \frac{1}{\tau} \int_{\varepsilon-\tau}^{\varepsilon} \beta_1(s) \mathrm{d}s \, \langle \nu_1, e_2 \rangle \, e_2 \qquad \forall\, 0 < \tau \leq \tau_0. \tag{4.6}$$

As β_1 was chosen such that the limit in (4.1) does not exist, since $\langle \nu_1, e_2 \rangle \neq 0$ and $0 < \varepsilon < \vartheta$, and since, for $\gamma \in \{0, 1\}$, we obtain situations in which only the initial value y_0 and the forcing term u are perturbed, the assertion of the theorem now follows immediately from (4.2), (4.5), and (4.6). □

Remark 4.2.

- The construction in the proof of Theorem 4.1 suggests that it might be possible to obtain directional differentiability results for the operator $\mathcal{S}$ in the presence of obtuse interior angles if one restricts the attention to situations in which the time-derivative of the function $\mathcal{P}(u, y_0)\colon [0, T] \to \mathbb{R}^d$ possesses left and right limits everywhere in $[0, T]$. We leave this topic for future research.
- As (4.6) shows, in the situation considered in the proof of Theorem 4.1, limits (of subsequences) of difference quotients of $\mathcal{S}$ do not only depend on where the multiplier maps λ_i in (2.40) are nonconstant, but also on how fast the maps λ_i grow. This is a major difference to the non-obtuse case; see (3.23), (3.16), and Definition 3.10. In the elliptic setting, dependencies of this type typically occur when the admissible set of the considered variational inequality (interpreted as a subset of the underlying function space) is not (extended) polyhedric but possesses proper nonzero curvature at its boundary; *cf.* [58], Lemmas 5.3, 5.8 and the references therein. The observations in Section 3 and the proof of Theorem 4.1 thus suggest that the admissible set $\{v \in CBV([0, T]; \mathbb{R}^d) \mid v(0) = y_0,\ v(t) \in Z \text{ for all } t \in [0, T]\}$ of the EVI (V), considered

as a subset of the space $BV([0,T];\mathbb{R}^d)$, possesses proper nonzero curvature in the sense of the second-order epi-derivatives studied in [58] when Z has an obtuse ridge and is extended polyhedric (in a suitably defined sense) when Z is non-obtuse. Note that such structural differences between non-obtuse and obtuse polyhedra are completely unheard of in the context of elliptic variational inequalities; see [60]. We again leave this topic for future research.

5. Remarks on applications in optimal control

We conclude this paper by briefly commenting on applications of Theorem 1.1 in the field of optimal control. For illustration purposes, we restrict our attention to the following model problem governed by (V):

$$
\begin{aligned}
\text{Minimize} \quad & \mathcal{J}(y, y(T), u) \\
\text{w.r.t.} \quad & y \in CBV([0,T];\mathbb{R}^d), \quad u \in U, \\
\text{s.t.} \quad & \begin{cases} y(t) \in Z \quad \forall t \in [0,T], \qquad y(0) = y_0, \\ \displaystyle\int_0^T \langle v - y, \mathrm{d}(y-u)\rangle \geq 0 \quad \forall v \in C([0,T];Z). \end{cases}
\end{aligned}
\tag{P}
$$

Our standing assumptions on the quantities in (P) are as follows (*cf.* [48], Asm. 3.1, Cor. 5.1):

Assumption 5.1 (standing assumptions for Sect. 5)**.**

i) $T > 0$ and $d \in \mathbb{N}$ are given and fixed.
ii) $(U, \|\cdot\|_U)$ is a reflexive real Banach space such that $U \subset CBV([0,T];\mathbb{R}^d)$ holds and such that U is continuously, densely, and compactly embedded into $C([0,T];\mathbb{R}^d)$.
iii) $\mathcal{J} \colon L^\infty((0,T);\mathbb{R}^d) \times \mathbb{R}^d \times U \to \mathbb{R}$ is a function with the following properties:
  (a) $\mathcal{J}$ is Fréchet differentiable and $\partial_1 \mathcal{J}(y, y(T), u) \in L^1((0,T);\mathbb{R}^d)$ for all $(y,u) \in CBV([0,T];\mathbb{R}^d) \times U$. Here, $L^1((0,T);\mathbb{R}^d)$ is identified with a subset of $L^\infty((0,T);\mathbb{R}^d)^*$ in the canonical way.
  (b) $\mathcal{J}$ is lower semicontinuous in the sense that, for all $\{(y_k, z_k, u_k)\} \subset C([0,T];\mathbb{R}^d) \times \mathbb{R}^d \times U$ satisfying $y_k \to y$ in $C([0,T];\mathbb{R}^d)$, $z_k \to z$ in $\mathbb{R}^d$, and $u_k \rightharpoonup u$ in U, we have
  $$\liminf_{k\to\infty} \mathcal{J}(y_k, z_k, u_k) \geq \mathcal{J}(y, z, u).$$
  (c) $\mathcal{J}$ is radially unbounded in the sense that there exists a function $\rho\colon [0,\infty) \to \mathbb{R}$ satisfying
  $$\rho(s) \to \infty \text{ for } s \to \infty \qquad \text{and} \qquad \mathcal{J}(y,z,u) \geq \rho\left(\|u\|_U\right) \quad \forall (y,z,u) \in C([0,T];\mathbb{R}^d) \times \mathbb{R}^d \times U.$$
iv) $Z \subset \mathbb{R}^d$ is a full-dimensional convex non-obtuse polyhedron.
v) $y_0 \in Z$ is a given and fixed starting value.

A prototypical example covered by the setting in Assumption 5.1 is

$$U := H^1((0,T);\mathbb{R}^d), \qquad \mathcal{J}(y,z,u) := \frac{\omega_1}{2}\|y - y_D\|^2_{L^2((0,T);\mathbb{R}^d)} + \frac{\omega_2}{2}|z - y_T|^2 + \frac{\omega_3}{2}\|u\|^2_{H^1((0,T);\mathbb{R}^d)},$$

with given $\omega_1, \omega_2, \omega_3 > 0$, $y_D \in L^2((0,T);\mathbb{R}^d)$, and $y_T \in \mathbb{R}^d$. From standard arguments, we obtain:

Theorem 5.2 (solvability of (P))**.** *Problem* (P) *possesses a solution* $(\bar{u}, \bar{y}) \in U \times CBV([0,T];\mathbb{R}^d)$.

Proof. The solvability of (P) follows straightforwardly from the properties of U and $\mathcal{J}$ in Assumption 5.1, the Lipschitz estimate (2.39), and the direct method of the calculus of variations. □

The main insight that Theorem 1.1 provides in the context of the minimization problem (P) is that it is possible to formulate a classical Bouligand stationarity condition for local minimizers, *i.e.*, a standard first-order necessary optimality condition in terms of the directional derivatives of the reduced objective function $U \ni u \mapsto \mathcal{J}(\mathcal{S}(u, y_0), \mathcal{S}(u, y_0)(T), u) \in \mathbb{R}$.

Corollary 5.3 (Bouligand stationarity condition)**.** *Suppose that $\bar{u} \in U$ is a locally optimal control of* (P) *with associated state $\bar{y} := \mathcal{S}(\bar{u}, y_0)$. Then the map $\mathcal{S}$ is directionally differentiable at $(\bar{u}, y_0)$ in the sense of part (I) of Theorem 1.1 and it holds*

$$\begin{aligned}\int_0^T \langle \partial_1 \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}), \mathcal{S}'((\bar{u}, y_0); (h, 0))\rangle \mathrm{d}s + \langle \partial_2 \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}), \mathcal{S}'((\bar{u}, y_0); (h, 0))(T)\rangle \\ + \langle \partial_3 \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}), h\rangle_{U^*,U} \geq 0 \quad \forall h \in U.\end{aligned} \tag{5.1}$$

Proof. If $\bar{u}$ is a locally optimal control of the problem (P) with state $\bar{y} := \mathcal{S}(\bar{u}, y_0)$ and $h \in U$ is fixed, then our assumptions, Theorem 1.1, the Lipschitz estimate (2.39), and the local optimality of $\bar{u}$ imply that $\mathcal{S}$ is directionally differentiable at $(\bar{u}, y_0)$ in the sense of (1.4) and that

$$\begin{aligned}0 &\leq \mathcal{J}(\mathcal{S}(\bar{u} + \tau h, y_0), \mathcal{S}(\bar{u} + \tau h, y_0)(T), \bar{u} + \tau h) - \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}) \\ &= \int_0^T \langle \partial_1 \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}), \mathcal{S}(\bar{u} + \tau h, y_0) - \bar{y}\rangle \mathrm{d}s + \langle \partial_2 \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}), \mathcal{S}(\bar{u} + \tau h, y_0)(T) - \bar{y}(T)\rangle \\ &\qquad + \langle \partial_3 \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}), \tau h\rangle_{U^*,U} + o(\tau)\end{aligned}$$

holds for all sufficiently small $\tau > 0$, where the Landau notation refers to the limit $\tau \to 0$. If we divide by τ in the above and pass to the limit $\tau \to 0$, using (1.3), the pointwise convergence in (1.4), and the dominated convergence theorem, then the assertion of the corollary follows. ☐

Note that it is typically hard to use the Bouligand stationarity condition (5.1) directly as a point of departure for the design of numerical solution algorithms. The main appeal of (5.1) is that it constitutes the most natural first-order necessary optimality condition for the problem (P) and, thus, can be used as a point of reference when comparing stationarity systems that have been derived *via* regularization or discretization techniques; see [13, 32–43] and the comparisons in [45]. A very interesting question in this context is whether the characterization of the directional derivatives of $\mathcal{S}$ by means of (1.5) can be used to reformulate the Bouligand stationarity condition (5.1) as an equivalent primal-dual optimality system in qualified form. In the case $d = 1$ and $Z = [-r, r]$, $r > 0$, it has been demonstrated in [48] that this is indeed possible and that a control $\bar{u} \in U$ with state $\bar{y} := \mathcal{S}(\bar{u}, y_0)$ satisfies (5.1) if and only if there exist an adjoint state $\bar{p} \in BV([0, T])$ and a multiplier $\bar{\mu} \in G_r([0, T])^*$ such that the following *strong stationarity system* is satisfied:

$$\begin{gathered}\bar{p}(0) = \bar{p}(T) = 0, \qquad \bar{p}(t) = \bar{p}(t-)\ \forall t \in [0, T), \qquad \bar{p}(t-) \in K_{\text{crit}}^{\text{ptw}}(\bar{y}, \bar{u})(t)\ \forall t \in [0, T], \\ \langle \bar{\mu}, z\rangle_{G_r([0,T])^*, G_r([0,T])} \geq 0 \quad \forall z \in \mathcal{K}_{G_r}^{\text{red,crit}}(\bar{y}, \bar{u}), \qquad \int_0^T h\, \mathrm{d}\bar{p} = \langle \partial_3 \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}), h\rangle_{U^*,U}\ \ \forall h \in U, \\ -\int_0^T z\, \mathrm{d}\bar{p} = \int_0^T \partial_1 \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}) z\, \mathrm{d}s + \partial_2 \mathcal{J}(\bar{y}, \bar{y}(T), \bar{u}) z(T) - \langle \bar{\mu}, z\rangle_{G_r([0,T])^*, G_r([0,T])} \qquad \forall z \in G_r([0, T]).\end{gathered} \tag{5.2}$$

Here, $K_{\text{crit}}^{\text{ptw}}(\bar{y}, \bar{u})\colon [0, T] \rightrightarrows \mathbb{R}$ is the same map as in (3.34) and $\mathcal{K}_{G_r}^{\text{red,crit}}(\bar{y}, \bar{u})$ denotes a suitably defined subset of the critical cone $\mathcal{K}_{G_r}^{\text{crit}}(\bar{y}, \bar{u})$; see [48], Definitions 4.7, 6.1. The main instrument for establishing (5.2) in [48] is a version of Theorem 3.28 which shows that the space $BV_r([0, T])$ in the first line of (3.17) can be replaced by $C^\infty([0, T])$ when functions $z \in \mathcal{K}_{G_r}^{\text{red,crit}}(\bar{y}, \bar{u})$ are considered and $d = 1$ holds; see [48], Theorem 6.5. Whether it is possible to refine the temporal polyhedricity result in Theorem 3.28 in a similar way for $d > 1$ and to obtain a strong stationarity system for (P) analogous to (5.2) for arbitrary dimensions is an open problem.

Data availability statement

No new data/codes were created or analyzed in this study.

References

[1] W.W. Desch and J. Turi, The stop operator related to a convex polyhedron. *J. Diff. Equ.* **157** (1999) 329–347.

[2] P. Dupuis and H. Ishii, On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications. *Stoch. Stoch. Rep.* **35** (1991) 31–62.

[3] M.A. Krasnosel'skii and A.V. Pokrovskii, *Systems with Hysteresis*, 1st edn. Springer, Berlin, Heidelberg (1989).

[4] P. Krejčí, Evolution variational inequalities and multidimensional hysteresis operators, in *Nonlinear Differential Equations*, vol. 404 of *Research Notes in Mathematics*, edited by P. Drábek, P. Krejčí and P. Takáč. Chapman & Hall/CRC, London (1999) 47–110.

[5] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*. Springer Series in Operations Research. Springer, New York (2000).

[6] J.-J. Moreau, Problème d'évolution associé à un convexe mobile d'un espace hilbertien. *C. R. Acad. Sci. Paris Sér. A-B* **276** (1973) A791–A794.

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

[8] J.-J. Moreau, On unilateral constraints, friction and plasticity, in *New Variational Techniques in Mathematical Physics*, edited by G. Capriz and G. Stampacchia. Springer, Berlin, Heidelberg (2011) 171–322.

[9] P. Krejčí, *Hysteresis, Convexity and Dissipation in Hyperbolic Equations*, Vol. 8 of *Mathematical Sciences and Applications*. Gakkōtosho, Tokyo (1996).

[10] V. Acary, O. Bonnefon and B. Brogliato, *Nonsmooth Modeling and Simulation for Switched Circuits*. Lecture Notes in Electrical Engineering. Springer (2010).

[11] T.H. Cao, B.S. Mordukhovich, D. Nguyen and T. Nguyen, Applications of controlled sweeping processes to nonlinear crowd motion models with obstacles. *IEEE Control Syst. Lett.* **6** (2022) 740–745.

[12] A. Mielke and T. Roubíček, *Rate-Independent Systems*, vol. 193 of *Applied Mathematical Sciences*. Springer, New York (2015).

[13] L. Adam and J. Outrata, On optimal control of a sweeping process coupled with an ordinary differential equation. *Discrete Contin. Dyn. Syst. Ser. B* **19** (2014) 2709–2738.

[14] A.H. Siddiqi, P. Manchanda and M. Brokate, On some recent developments concerning Moreau's sweeping process, in *Trends in Industrial and Applied Mathematics: Proceedings of the 1st International Conference on Industrial and Applied Mathematics of the Indian Subcontinent*, Edited by A. H. Siddiqi and M. Kočvara. Springer, Boston, MA (2002) 339–354.

[15] M. Brokate and P. Krejčí, Weak differentiability of scalar hysteresis operators. *Discrete Contin. Dyn. Syst.* **35** (2015) 2405–2421.

[16] M. Brokate, Newton and Bouligand derivatives of the scalar play and stop operator. *Math. Model. Nat. Phenom.* **15** (2020) Art. 51.

[17] M. Brokate and P. Krejčí, A variational inequality for the derivative of the scalar play operator. *J. Appl. Numer. Optim.* **3** (2021) 263–283.

[18] L. Betz, Strong stationarity for optimal control of a nonsmooth coupled system: application to a viscous evolutionary variational inequality coupled with an elliptic PDE. *SIAM J. Optim.* **29** (2019) 3069–3099.

[19] J. Jarušek, M. Krbec, M. Rao and J. Sokołowski, Conical differentiability for evolution variational inequalities. *J. Diff. Equ.* **193** (2003) 131–146.

[20] A. Haraux, How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities. *J. Math. Soc. Japan* **29** (1977) 615–631.

[21] F. Mignot, Contrôle dans les inéquations variationelles elliptiques. *J. Funct. Anal.* **22** (1976) 130–185.

[22] C. Christof, Sensitivity analysis and optimal control of obstacle-type evolution variational inequalities. *SIAM J. Control Optim.* **57** (2019) 192–218.

[23] E.H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory I and II. *Contrib. Nonlinear Funct. Anal.* (1971) 237–424.

[24] S. Adly and L. Bourdin, Sensitivity analysis of variational inequalities via twice epi-differentiability and proto-differentiability of the proximity operator. *SIAM J. Optim.* **28** (2018) 1699–1725.

[25] J.F. Bonnans, R. Cominetti and A. Shapiro, Sensitivity analysis of optimization problems under second order regular constraints. *Math. Oper. Res.* **23** (1998) 806–831.

[26] S. Fitzpatrick and R.R. Phelps, Differentiability of the metric projection in Hilbert space. *Trans. Am. Math. Soc.* **270** (1982) 483–501.

[27] D. Noll, Directional differentiability of the metric projection in Hilbert space. *Pacific J. Math.* **170** (1995) 567–592.

[28] A. Shapiro, Existence and differentiability of metric projections in Hilbert spaces. *SIAM J. Optim.* **4** (1994) 130–141.

[29] C.N. Do, Generalized second-order derivatives of convex functions in reflexive Banach spaces. *Trans. Am. Math. Soc.* **334** (1992) 281–301.

[30] C. Christof, *Sensitivity Analysis of Elliptic Variational Inequalities of the First and the Second Kind.* PhD thesis, Technische Universität Dortmund (2018).

[31] C. Christof and G. Wachsmuth, Differential sensitivity analysis of variational inequalities with locally Lipschitz continuous solution operators. *Appl. Math. Optim.* **81** (2020) 23–62.

[32] T.H. Cao, N.T. Khalil, B.S. Mordukhovich, D. Nguyen, T. Nguyen and F.L. Pereira, Optimization of controlled free-time sweeping processes with applications to marine surface vehicle modeling. *IEEE Control Syst. Lett.* **6** (2022) 782–787.

[33] T.H. Cao and B.S. Mordukhovich, Optimal control of a perturbed sweeping process via discrete approximations. *Discrete Contin. Dyn. Syst. Ser. B* **21** (2016) 3331–3358.

[34] T.H. Cao and B.S. Mordukhovich, Optimal control of a nonconvex perturbed sweeping process. *J. Diff. Equ.* **266** (2019) 1003–1050.

[35] G. Colombo, R. Henrion, D.H. Nguyen and B.S. Mordukhovich, Optimal control of the sweeping process over polyhedral controlled sets. *J. Diff. Equ.* **260** (2016) 3397–3447.

[36] G. Colombo, B.S. Mordukhovich and D. Nguyen, Optimization of a perturbed sweeping process by constrained discontinuous controls. *SIAM J. Control Optim.* **58** (2020) 2678–2709.

[37] R. Henrion, A. Jourani and B.S. Mordukhovich, Controlled polyhedral sweeping processes: Existence, stability, and optimality conditions. *J. Diff. Equ.* **366** (2023) 408–443.

[38] C.E. Arroud and G. Colombo, A maximum principle for the controlled sweeping process. *Set-Valued Var. Anal.* **26** (2018) 607–629.

[39] M. Brokate and P. Krejčí, Optimal control of ODE systems involving a rate independent variational inequality. *Discrete Contin. Dyn. Syst. Ser. B* **18** (2013) 331–348.

[40] M.d.R. de Pinho, M.M.A. Ferreira and G.V. Smirnov, Optimal control involving sweeping processes. *Set-Valued Var. Anal.* **27** (2019) 523–548.

[41] R. Herzog, D. Knees, C. Meyer, M. Sievers, A. Stötzner and S. Thomas. Rate-independent systems and their viscous regularizations: analysis, simulation, and optimal control, in *Non-Smooth and Complementarity-Based Distributed Parameter Systems: Simulation and Hierarchical Optimization*, edited by M. Hintermüller, R. Herzog, C. Kanzow, M. Ulbrich and S. Ulbrich. Springer, Cham (2022) 121–144.

[42] C. Nour and V. Zeidan, Optimal control of nonconvex sweeping processes with separable endpoints: nonsmooth maximum principle for local minimizers. *J. Diff. Equ.* **318** (2022) 113–168.

[43] U. Stefanelli, D. Wachsmuth and G. Wachsmuth, Optimal control of a rate-independent evolution equation via viscous regularization. *Discrete Contin. Dyn. Syst. Ser. S* **10** (2017) 1467–1485.

[44] C. Christof, C. Meyer, B. Schweizer and S. Turek, Strong stationarity for optimal control of variational inequalities of the second kind, in *Non-Smooth and Complementarity-Based Distributed Parameter Systems: Simulation and Hierarchical Optimization*, edited by M. Hintermüller, R. Herzog, C. Kanzow, M. Ulbrich and S. Ulbrich. Springer, Cham (2022) 307–327.

[45] F. Harder and G. Wachsmuth, Comparison of optimality systems for the optimal control of the obstacle problem. *GAMM-Mitt.* **40** (2018) 312–338.

[46] M. Hintermüller and I. Kopacka, Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. *SIAM J. Optim.* **20** (2008) 868–902.

[47] F. Mignot and J.P. Puel, Optimal control in some variational inequalities. *SIAM J. Control Optim.* **22** (1984) 466–476.

[48] M. Brokate and C. Christof, Strong stationarity conditions for optimal control problems governed by a rate-independent evolution variational inequality. *SIAM J. Control Optim.* **61** (2023) 2222–2250.

[49] P. Krejčí and V. Recupero, Comparing BV solutions of rate independent processes. *J. Convex Anal.* **21** (2014) 121–146.

[50] H. Attouch, G. Buttazzo and G. Michaille, *Variational Analysis in Sobolev and BV Spaces.* SIAM, Philadelphia (2006).

[51] C. Christof and G. Wachsmuth, Lipschitz stability and Hadamard directional differentiability for elliptic and parabolic obstacle-type quasi-variational inequalities. *SIAM J. Control Optim.* **60** (2022) 3430–3456.

[52] G.A. Monteiro, A. Slavík and M. Tvrdý, *Kurzweil-Stieltjes Integral: Theory and Applications*, vol. 15 of *Series in Real Analysis.* World Scientific, Singapore (2019).

[53] W. Bruns and J. Gubeladze, *Polytopes, Rings, and K-Theory.* Springer Monographs in Mathematics. Springer, New York (2009).

[54] M. Berger, M. Cole and S. Levy, *Geometry II.* Universitext. Springer, Berlin, Heidelberg (2009).

[55] U. Faigle, W. Kern and G. Still, *Algorithmic Principles of Mathematical Programming.* Texts in the Mathematical Sciences. Springer (2002).

[56] O.L. Mangasarian, *Nonlinear Programming*, vol. 10 of *SIAM's Classics in Applied Mathematics*, 2nd edn. SIAM (1994).

[57] L.C. Evans, *Partial Differential Equations*, 2nd edn. AMS, Providence, RI (2010).

[58] C. Christof and G. Wachsmuth, No-gap second-order conditions via a directional curvature functional. *SIAM J. Optim.* **28** (2018) 2097–2130.

[59] G. Wachsmuth, A guided tour of polyhedric sets: basic properties, new results on intersections and applications. *J. Convex Anal.* **26** (2019) 153–188.

[60] G. Wachsmuth, Pointwise constraints in vector-valued Sobolev spaces: with applications in optimal control. *Appl. Math. Optim.* **77** (2018) 463–497.