J. Differential Equations 253 (2012) 203–224

Contents lists available at SciVerse ScienceDirect

Journal of Differential Equations

www.elsevier.com/locate/jde

Subdifferential analysis of differential inclusions via discretization

C.H. Jeffrey Pang

Massachusetts Institute of Technology, Department of Mathematics, 2-334, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, United States

A R T I C L E   I N F O

Article history:
Received 9 December 2011
Revised 19 March 2012
Available online 6 April 2012

MSC:
34A60
49K05
49K15
49K40

Keywords:
Differential inclusions
Optimality conditions
Subdifferential
Reachable map
Coderivatives

A B S T R A C T

The framework of differential inclusions encompasses modern optimal control and the calculus of variations. Necessary optimality conditions in the literature identify potentially optimal paths, but do not show how to perturb paths to optimality. We first look at the corresponding discretized inclusions, estimating the subdifferential dependence of the optimal value in terms of the endpoints of the feasible paths. Our approach is to first estimate the coderivative of the reachable map. The discretized (nonsmooth) Euler–Lagrange and Transversality Conditions follow as a corollary. We obtain corresponding results for differential inclusions by passing discretized inclusions to the limit.

Contents

E-mail address: chj2pang@mit.edu.

0022-0396/$ – see front matter
http://dx.doi.org/10.1016/j.jde.2012.03.019

1. Introduction

The subject of this paper is the analysis of discretized differential inclusions by calculating the coderivatives of the discretized reachable map. We then pass these results to the limit to obtain results on differential inclusions. We say that S is a *set-valued map* or a *multifunction*, denoted by $S : X \rightrightarrows Y$, if $S(x) \subset Y$ for all $x \in X$. For $F : [0, T] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and $C \subset \mathbb{R}^n \times \mathbb{R}^n$, consider the *differential inclusion*:

$$\begin{aligned} \min_{x(\cdot)\in AC([0,T],\mathbb{R}^n)} \quad & \varphi\big(x(0), x(T)\big) \\ \text{s.t.} \quad & x'(t) \in F\big(t, x(t)\big) \quad \text{for } t \in [0, T] \text{ a.e.} \end{aligned} \tag{1.1}$$

Here, $AC([0, T], \mathbb{R}^n)$ is the set of absolutely continuous functions of the form $x : [0, T] \to \mathbb{R}^n$. The constraint

$$\big(x(0), x(T)\big) \in C \subset \mathbb{R}^n \times \mathbb{R}^n$$

is sometimes included in the differential inclusion problem (1.1), but this constraint can be easily incorporated into the objective function φ. More details on differential inclusions can be obtained in the texts [1–3,9,14,15]. As is popularized in these texts, the differential inclusion framework (1.1) encompasses optimal control and the calculus of variations.

In order to optimize (1.1), much attention has focused on necessary optimality conditions for a path $x(\cdot)$. Such research was undertaken in the last few decades by Clarke, Loewen, Rockafellar, Ioffe, Vinter, Mordukhovich, Kaskosz and Lojasiewicz, Milyutin, Smirnov, Zheng, Zhu and others, building on results in the calculus of variations and optimal control. For a history of the development of the necessary optimality conditions, we refer to the previously mentioned texts. The following conditions are currently understood as useful necessary optimality conditions for a feasible path $\bar{x}(\cdot)$ of (1.1):

(TC) (Transversality Condition)

$$\big(-p(0), p(T)\big) \in \partial\varphi\big(\bar{x}(0), \bar{x}(T)\big).$$

(EL) (Euler–Lagrange Condition)

$$p'(t) \in -\overline{\text{co}}\, D_x^* F\big(t, \bar{x}(t) \,\big|\, \bar{x}'(t)\big)\big(p(t)\big) \quad \text{for } t \in [0, T] \text{ a.e.}$$

(WP) (Weierstrass–Pontryagin Maximum Principle)

$$\big\langle -p(t), v - \bar{x}'(t)\big\rangle \leqslant 0 \quad \text{for all } v \in F\big(t, \bar{x}(t)\big),\ t \in [0, T] \text{ a.e.}$$

While such necessary conditions are helpful in finding candidates for a minimizing path, the deficiency in such necessary conditions is that they give no indication on how to perturb a feasible path to optimality. As a first step, we study the discrete inclusions corresponding to the differential inclusion and calculate the dependence of the differential inclusion on its initial point.

Define the *reachable map* (or *attainable map*) $R:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by

$$\begin{aligned} R(x_0) := \big\{y\colon \exists x(\cdot) \in AC\big([0,T],\mathbb{R}^n\big) \text{ s.t.} \\ x'(t) \in F\big(t,x(t)\big) \text{ for } t \in [0,T] \text{ a.e.,} \\ x(0) = x_0 \text{ and } x(T) = y\big\}. \end{aligned} \tag{1.2}$$

In order to study (1.1), we consider

$$\begin{aligned} f(x) := \min_x \quad & \varphi(x,y) \\ \text{s.t.} \quad & y \in R(x). \end{aligned} \tag{1.3}$$

We study (1.3) under the broader framework of marginal functions. For a set-valued map $G : X \rightrightarrows Y$ and a function $\varphi : X \times Y \to \mathbb{R}$, the *marginal function* $f : X \to \mathbb{R}$ is

$$f(x) := \inf\big\{\varphi(x,y)\colon\ y \in G(x)\big\}. \tag{1.4}$$

One can view the value x as a parameter of an optimization problem in terms of y. A well-studied example of a set-valued map G is the map $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ defined by

$$\begin{aligned} G(x) &= \big\{y \bigm| y \in F(x) + \big[\{0\} \times \mathbb{R}^{m_2}_-\big]\big\} \quad \text{where } m_1 + m_2 = m \text{ and } F : \mathbb{R}^n \to \mathbb{R}^m \text{ is smooth} \\ &= \big\{y \bigm| y_i = F_i(x) \text{ for } 1 \leqslant i \leqslant m_1 \text{ and } y_i \leqslant F_i(x) \text{ for } m_1 + 1 \leqslant i \leqslant m\big\}. \end{aligned}$$

The sensitivity analysis of marginal functions can be analyzed with tools of variational analysis and generalized differentiation. We denote the composition $S_2 \circ S_1 : X \rightrightarrows Z$ of set-valued maps $S_1 : X \rightrightarrows Y$ and $S_2 : Y \rightrightarrows Z$ in the usual way by

$$S_2 \circ S_1(x) = \bigcup_{y \in S_1(x)} S_2(y).$$

Denote the epigraphical mapping of φ and f by $E_\varphi : X \times Y \rightrightarrows \mathbb{R}$ and $E_f : X \rightrightarrows \mathbb{R}$ respectively. Then E_φ and E_f satisfy the relation

$$E_f(x) = E_\varphi \circ \bar{G}(x), \tag{1.5}$$

where $\bar{G} : X \rightrightarrows X \times Y$ is defined by $\bar{G}(x) = \{x\} \times G(x)$. The relationship (1.5) and a set-valued chain rule can be used to express differentiability properties of f in terms of the coderivatives of G and φ.

1.1. Contributions of this paper

In this work, we focus on the subdifferential analysis of the discretized differential inclusion problem by finding the subdifferential of f_N, where f_N is the discretized analogue of (1.3). Our approach is to look at the marginal function framework and calculate the coderivatives of the discretized reachable map $R_N(\cdot)$. The coderivative of the reachable map gives new insight on the Euler–Lagrange Condition (EL). We also study the limitations of a discrete analogue of the Weierstrass–Pontryagin Maximum Principle (WP).

For a set-valued map $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ between finite dimensional spaces, [11] recently established that the convexified limiting coderivative characterizes the set of positively homogeneous maps that are generalized derivatives of S as defined in [10]. We will recall on this relation in Section 2, limiting our analysis to the finite dimensional case. By making use of this result, we can obtain the convexified limiting coderivative of the reachable map $R(\cdot)$ by passing a sequence of discrete problems to the limit. The marginal function framework allows us to calculate the subdifferential dependence of f of the differential inclusion in terms of its initial value.

An advantage of our subdifferential analysis over the necessary optimality conditions is that our analysis gives a better indication of how to perturb a feasible path to optimality. Our approach in analyzing differential inclusions can be compared to previous discrete approximation methods. In particular, the notes in [15, Chapter 7] highlighted [7,8,13]. For a more comprehensive commentary and further references, we refer the reader to [9, Section 6.5.12] and [14].

1.2. Outline

In Section 2, we recall standard definitions in variational analysis and some results in [11] that will be used in the later part of the paper. In Section 3, we recall chain rules for coderivatives, and show how these results can be easily extended for the convexified limiting coderivative. In Section 4, we study the discretized differential inclusion problem. Finally, in Section 5, we study the continuous inclusion problem by passing the discretized problems in Section 4 to the limit, and find formulas for the convexified limiting coderivative of the reachable map.

2. Preliminaries and notation

This section recalls some standard definitions in variational analysis and some other results in [11] that will be used in the remainder of this paper. The texts [12,9] contain many standard definitions in variational analysis, like inner and outer semicontinuity (isc and osc) and the Pompieu–Hausdorff distance $\mathbf{d}(\cdot,\cdot)$. We highlight some of definitions used most often in this paper. We denote the set $\{1,2,\ldots,N\}$ by $\overline{1,N}$. For set-valued maps $H_i:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$, $i=1,2$, we let $H_1\subset H_2$ denote $H_1(x)\subset H_2(x)$ for all x, or equivalently $\mathrm{Graph}(H_1)\subset\mathrm{Graph}(H_2)$.

We recall the definition of coderivatives.

Definition 2.1 *(Normal cones).* For a set $C\subset\mathbb{R}^n$, the *regular normal cone* at $\bar{x}$ is defined as

$$\hat{N}_C(\bar{x}):=\bigl\{y \bigm| \langle y,x-\bar{x}\rangle\leqslant o\bigl(\|x-\bar{x}\|\bigr)\text{ for all }x\in C\bigr\}.$$

The *limiting* (or *Mordukhovich*) *normal cone* $N_C(\bar{x})$ is defined as $\limsup_{x\xrightarrow[C]{}\bar{x}}\hat{N}_C(x)$, or as

$$N_C(\bar{x})=\bigl\{y \bigm| \text{there exist } x_i\xrightarrow[C]{}\bar{x},\ y_i\in\hat{N}_C(x_i)\text{ such that }y_i\to y\bigr\}.$$

Definition 2.2 *(Coderivatives).* For a set-valued map $S:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ locally closed at $(\bar{x},\bar{y})\in\mathrm{Graph}(S)$, the *regular coderivative* at $(\bar{x},\bar{y})$, denoted by $\hat{D}^*S(\bar{x}\mid\bar{y}):\mathbb{R}^m\rightrightarrows\mathbb{R}^n$, is defined by

$$\begin{aligned}v\in\hat{D}^*S(\bar{x}\mid\bar{y})(u)\quad&\Leftrightarrow\quad(v,-u)\in\hat{N}_{\mathrm{Graph}(S)}(\bar{x},\bar{y})\\&\Leftrightarrow\quad\bigl\langle(v,-u),(x,y)-(\bar{x},\bar{y})\bigr\rangle\leqslant o\bigl(\bigl\|(x,y)-(\bar{x},\bar{y})\bigr\|\bigr)\quad\text{for all }(x,y)\in\mathrm{Graph}(S).\end{aligned}$$

The *limiting* (or *Mordukhovich*) *coderivative* at $(\bar{x},\bar{y})\in\mathrm{Graph}(S)$ is denoted by $D^*S(\bar{x}\mid\bar{y}):\mathbb{R}^m\rightrightarrows\mathbb{R}^n$ and is defined by

$$v\in D^*S(\bar{x}\mid\bar{y})(u)\quad\Leftrightarrow\quad(v,-u)\in N_{\mathrm{Graph}(S)}(\bar{x},\bar{y}).$$

The *convexified limiting coderivative* $\overline{\mathrm{co}}\,D^*S(\bar{x}\mid\bar{y}):\mathbb{R}^m\rightrightarrows\mathbb{R}^n$ is defined in the natural manner.

We recall the definition of subdifferentials.

Definition 2.3 *(Subdifferentials).* Consider a function $f:\mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$ at a point $\bar{x}$ where $f(\bar{x})$ is finite. Then the *limiting* (or *Mordukhovich*) *subdifferential* $\partial f(\bar{x})$, *horizon* (or *singular*) *subdifferential* $\partial^\infty f(\bar{x})$ and the *Clarke* (or *generalized*) *subdifferential* $\partial_C f(\bar{x})$ are defined respectively by

$$\begin{aligned}
\partial f(\bar{x}) &:= \bigl\{v \bigm| (v,-1) \in N_{\mathrm{epi}(f)}\bigl(\bar{x}, f(\bar{x})\bigr)\bigr\} \\
&= D^* E_f\bigl(\bar{x} \bigm| f(\bar{x})\bigr)(1), \\
\partial^\infty f(\bar{x}) &:= \bigl\{v \bigm| (v,0) \in N_{\mathrm{epi}(f)}\bigl(\bar{x}, f(\bar{x})\bigr)\bigr\} \\
&= D^* E_f\bigl(\bar{x} \bigm| f(\bar{x})\bigr)(0), \quad \text{and} \\
\partial_C f(\bar{x}) &:= \overline{\mathrm{co}}\, \partial f(\bar{x}) \\
&= \overline{\mathrm{co}}\, D^* E_f\bigl(\bar{x} \bigm| f(\bar{x})\bigr)(1).
\end{aligned}$$

The limiting and Clarke subdifferentials coincide with the usual definition of the subdifferential for convex functions. The subdifferential $\partial f(\bar{x})$ gives important information on how f varies with respect to x when close to $\bar{x}$.

We now recall the definition of generalized derivatives of set-valued maps in the sense of [10]. Let $\mathbb{B}$ denote the unit ball in the appropriate space.

Definition 2.4 *(Generalized differentiability).* (See [10].) Let $S:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be such that S is locally closed at $(\bar{x},\bar{y}) \in \mathrm{Graph}(S)$, and let $H:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a positively homogeneous map. The map S is *pseudo strictly H-differentiable at* $(\bar{x},\bar{y})$ if for any $\delta > 0$, there are neighborhoods U_δ of $\bar{x}$ and V_δ of $\bar{y}$ such that

$$S(x) \cap V_\delta \subset S\bigl(x'\bigr) + H\bigl(x - x'\bigr) + \delta\bigl\|x - x'\bigr\|\mathbb{B} \quad \text{for all } x, x' \in U_\delta.$$

We write

$$(H+\delta)(w) := H(w) + \delta\|w\|\mathbb{B}$$

to reduce notation. The map S has the *Aubin property* (or the *Lipschitz-like property*, or the *pseudo-Lipschitz* property) with modulus $\kappa \geqslant 0$ if S is pseudo strictly H-differentiable for some H defined by $H(w) = \kappa\|w\|\mathbb{B}$. The *graphical modulus* is the infimum of all such κ, and is denoted by $\mathrm{lip}\, S(\bar{x} \mid \bar{y})$.

We now recall the definition of prefans and the generalized derivative set $\mathscr{H}(D)$.

Definition 2.5 *(Prefans).* (See [6].) We say that $H:\mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a *prefan* if

(1) $H(p)$ is nonempty, convex and compact for all $p \in \mathbb{R}^n$,
(2) H is positively homogeneous, and
(3) $\|H\|^+ := \sup_{\|w\|\leqslant 1} \sup_{z\in H(w)} \|z\|$ is finite.

Definition 2.6 *(Generalized derivative set).* (See [11].) Let $D:\mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a positively homogeneous, osc set-valued map s.t. $\|D\|^+$ is finite. We define the *generalized derivative set* by

$$\mathscr{H}(D) := \Big\{ H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m\colon\ H \text{ is a prefan}, \\ \text{and for all } p \in \mathbb{R}^n \backslash \{0\} \text{ and } u \in \mathbb{R}^m, \\ \min_{y \in H(p)} \langle u, y \rangle \leqslant \min_{v \in \overline{\text{co}}\, D(u)} \langle v, p \rangle \Big\}.$$

The Aubin criterion characterizes the graphical modulus $\operatorname{lip} S(\bar{x} \mid \bar{y})$ in terms of graphical derivatives (which are in turn defined in terms of tangent cones), while the Mordukhovich criterion characterizes $\operatorname{lip} S(\bar{x} \mid \bar{y})$ in terms of coderivatives. Theorem 2.7 and Lemma 2.8 below characterize the set of possible generalized derivatives at a point $(\bar{x}, \bar{y}) \in \operatorname{Graph}(S)$, and can be seen as a generalization of the Mordukhovich criterion. While the proof in [11] makes heavy use of graphical derivatives and recent work in [5] (who in turn acknowledged Frankowska's contribution), the main results in finite dimensions have an appealing formulation in terms of coderivatives.

Theorem 2.7 *(Characterization of generalized derivatives). (See [11].) Let $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be locally closed at $(\bar{x}, \bar{y}) \in \operatorname{Graph}(S)$ and let $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a prefan. Then S is pseudo strictly H-differentiable at $(\bar{x}, \bar{y})$ if and only if $H \in \mathscr{H}(D^*S(\bar{x} \mid \bar{y}))$. (Note that $\mathscr{H}(D^*S(\bar{x} \mid \bar{y})) = \mathscr{H}(\overline{\text{co}}\, D^*S(\bar{x} \mid \bar{y}))$.)*

Lemma 2.8 *(Convexified coderivatives and generalized derivatives). (See [11].) Suppose $D_i : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ are positively homogeneous, osc, and $\|D_i\|^+$ are finite for $i = 1, 2$. Then the following strict reverse inclusion properties hold*:

(1) $\mathscr{H}(D_1) \supset \mathscr{H}(D_2)$ *iff* $\overline{\text{co}}\, D_1 \subset \overline{\text{co}}\, D_2$.
(2) $\mathscr{H}(D_1) \supsetneq \mathscr{H}(D_2)$ *iff* $\overline{\text{co}}\, D_1 \subsetneq \overline{\text{co}}\, D_2$.
(3) $\mathscr{H}(D_1) = \mathscr{H}(D_2)$ *iff* $\overline{\text{co}}\, D_1 = \overline{\text{co}}\, D_2$.

These results show that the convexified limiting coderivative $\overline{\text{co}}\, D^*S(\cdot \mid \cdot)(\cdot)$ is an effective tool for studying the generalized derivatives of set-valued maps, just like the way the Clarke subdifferential is useful for studying the generalized differentiability of single-valued maps.

We recall the definition of inner semicompactness that will be used in the chain rules for set-valued maps in this paper.

Definition 2.9 *(Inner semicompactness).* We say that $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *inner semicompact* at $\bar{x} \in \operatorname{dom}(S)$ if for every sequence $x_k \to \bar{x}$, there is a sequence $y_k \in S(x_k)$ that contains a convergent subsequence as $k \to \infty$.

In finite dimensions, if there is a neighborhood U of $\bar{x}$ and a bounded neighborhood V such that $S(U) \subset V$, then S is inner semicompact at $\bar{x}$.

Finally, we recall the definition of regularity and a straightforward consequence of graphical regularity.

Definition 2.10 *(Regularity).* We say that $C \subset \mathbb{R}^n$ is *Clarke regular* (or *normally regular*) at $\bar{x} \in C$ if C is locally closed at $\bar{x}$ and $N_C(\bar{x}) = \hat{N}_C(\bar{x})$. We say that $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *graphically regular* at $(\bar{x}, \bar{y}) \in \operatorname{Graph}(S)$ if $\operatorname{Graph}(S)$ is Clarke regular at $(\bar{x}, \bar{y})$.

Fact 2.11 *(Convexified limiting coderivatives under graph regularity). If $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is graphically regular at $(\bar{x}, \bar{y}) \in \operatorname{Graph}(S)$, then $\operatorname{Graph}(D^*S(\bar{x} \mid \bar{y})) = \operatorname{Graph}(\hat{D}^*S(\bar{x} \mid \bar{y}))$. Furthermore, $\operatorname{Graph}(\hat{D}^*S(\bar{x} \mid \bar{y}))$ is a convex cone, and we have $\overline{\text{co}}\, D^*S(\bar{x} \mid \bar{y}) \equiv D^*S(\bar{x} \mid \bar{y})$.*

3. Calculus of convexified limiting coderivatives

In this section, we discuss how the chain rule for the convexified limiting coderivatives can be obtained directly from the coderivative chain rules, removing parts irrelevant in the finite dimensional case. In Lemma 3.3, we deduce that the convexified limiting coderivative, together with the limiting subdifferential, are sufficient in calculating the Clarke subdifferential of marginal functions. This suggests that the convexified limiting coderivative of the reachable map as calculated in Section 5, while not as precise as the coderivative, can be a satisfactory conclusion.

We first write down the chain rule for finite dimensional coderivatives based on [9, Theorem 3.13] and [12, Theorem 10.37]. The formulas (3.2) and (3.3) for convexified limiting coderivatives are straightforward.

Theorem 3.1 *(Coderivative chain rule). Let* $G : \mathbb{R}^l \rightrightarrows \mathbb{R}^m$, $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, $\bar{z} \in (F \circ G)(\bar{x})$, *and*

$$S(x, z) := G(x) \cap F^{-1}(z) = \big\{y \in G(x):\ z \in F(y)\big\}.$$

The following assertions hold:

(1) *Given* $\bar{y} \in S(\bar{x}, \bar{z})$, *assume that* S *is inner semicontinuous at* $(\bar{x}, \bar{z}, \bar{y})$, *that the graphs of* F *and* G *are locally closed around the points* $(\bar{y}, \bar{z})$ *and* $(\bar{x}, \bar{y})$ *respectively, and that the qualification condition*

$$D^*F(\bar{y} \mid \bar{z})(0) \cap -D^*G^{-1}(\bar{y} \mid \bar{x})(0) = \{0\} \tag{3.1}$$

is fulfilled. Then one has

$$D^*(F \circ G)(\bar{x} \mid \bar{z}) \subset D^*G(\bar{x} \mid \bar{y}) \circ D^*F(\bar{y} \mid \bar{z}),$$

which in turn implies

$$\overline{\text{co}}\, D^*(F \circ G)(\bar{x} \mid \bar{z}) \subset \overline{\text{co}}\, D^*G(\bar{x} \mid \bar{y}) \circ D^*F(\bar{y} \mid \bar{z}). \tag{3.2}$$

(2) *Assume that* S *is inner semicompact at* $(\bar{x}, \bar{z})$, *that* G *and* F^{-1} *are closed-graph whenever* x *is near* $\bar{x}$ *and* z *is near* $\bar{z}$, *respectively, and that* (3.1) *holds for every* $\bar{y} \in S(\bar{x}, \bar{z})$. *Then*

$$D^*(F \circ G)(\bar{x} \mid \bar{z}) \subset \bigcup_{\bar{y} \in S(\bar{x}, \bar{z})} D^*G(\bar{x} \mid \bar{y}) \circ D^*F(\bar{y} \mid \bar{z}),$$

which in turn implies

$$\overline{\text{co}}\, D^*(F \circ G)(\bar{x} \mid \bar{z}) \subset \overline{\text{co}} \bigcup_{\bar{y} \in S(\bar{x}, \bar{z})} \overline{\text{co}}\, D^*G(\bar{x} \mid \bar{y}) \circ D^*F(\bar{y} \mid \bar{z}). \tag{3.3}$$

(3) *If* S *is locally bounded at* $(\bar{x}, \bar{z})$, (3.1) *holds for every* $\bar{y} \in S(\bar{x}, \bar{z})$, *and* F *and* G *are both graph convex* (*i.e., have convex graphs*), *then* $F \circ G$ *is also graph convex, and*

$$D^*(F \circ G)(\bar{x} \mid \bar{z}) = D^*G(\bar{x} \mid \bar{y}) \circ D^*F(\bar{y} \mid \bar{z}) \quad \textit{for any } \bar{y} \in S(\bar{x}, \bar{z}).$$

The formula (3.3) is not any stronger if its RHS is replaced by

$$\overline{\text{co}} \bigcup_{\bar{y} \in S(\bar{x}, \bar{z})} D^*G(\bar{x} \mid \bar{y}) \circ D^*F(\bar{y} \mid \bar{z}),$$

since this formula is equal to the RHS of (3.3). Therefore, to find the convexified limiting coderivative $\overline{\text{co}}\, D^*(F \circ G)(\bar{x} \mid \bar{z})$, the convexified limiting coderivative of G, i.e., $\overline{\text{co}}\, D^*G$, is sufficient. We explore the possibilities if we had relaxed the formulas (3.2) and (3.3) by replacing the relevant formulas with $\overline{\text{co}}\, D^*G(\bar{x} \mid \bar{y}) \circ \overline{\text{co}}\, D^*F(\bar{y} \mid \bar{z})$ instead.

Example 3.2 *(Tightness of chain rules).* Consider the set-valued maps $G_i : \mathbb{R} \rightrightarrows \mathbb{R}$, $i = 1, 2, 3$ and $F : \mathbb{R} \rightrightarrows \mathbb{R}$ defined by

$$\begin{aligned} G_1(x) &= \begin{cases} \mathbb{R} & \text{if } x \leqslant 0, \\ (-\infty, -x] \cup [x, \infty) & \text{if } x \geqslant 0, \end{cases} \\ G_2(x) &= \big[\min(0, x), \infty\big), \\ G_3(x) &= \big[\max(x/2, x), \infty\big), \quad \text{and} \\ F(x) &= \{-x\} \cup \{x\}. \end{aligned}$$

As illustrated in Table 1, we have

$$\begin{aligned} \overline{\text{co}}\, D^*(F \circ G_1)(0 \mid 0) = \overline{\text{co}}\, D^*G_1(0 \mid 0) \circ D^*F(0 \mid 0) \subsetneq \overline{\text{co}}\, D^*G_1(0 \mid 0) \circ \overline{\text{co}}\, D^*F(0 \mid 0), \\ \overline{\text{co}}\, D^*(F \circ G_2)(0 \mid 0) \subsetneq \overline{\text{co}}\, D^*G_2(0 \mid 0) \circ D^*F(0 \mid 0) = \overline{\text{co}}\, D^*G_2(0 \mid 0) \circ \overline{\text{co}}\, D^*F(0 \mid 0), \quad \text{and} \\ \overline{\text{co}}\, D^*(F \circ G_3)(0 \mid 0) \subsetneq \overline{\text{co}}\, D^*G_3(0 \mid 0) \circ D^*F(0 \mid 0) \subsetneq \overline{\text{co}}\, D^*G_3(0 \mid 0) \circ \overline{\text{co}}\, D^*F(0 \mid 0). \end{aligned}$$

The following general principle in the optimization of marginal functions will be used later. We take this result from [9, Theorem 3.38].

Lemma 3.3 *(Subdifferential of marginal functions). For the marginal function* (1.4)*, define the argminimum mapping by*

$$M(x) := \big\{ y \in G(x) \;\big|\; \varphi(x, y) = f(x) \big\}.$$

The following hold:

(1) *Given $\bar{y} \in M(\bar{x})$, assume that M is inner semicontinuous at $(\bar{x}, \bar{y})$, that $\varphi(x, y)$ is lsc around $(\bar{x}, \bar{y})$, and that* Graph(G) *is locally closed at $(\bar{x}, \bar{y})$. Suppose also that the qualification condition*

$$\partial^\infty \varphi(\bar{x}, \bar{y}) \cap -N_{\text{Graph}(G)}(\bar{x}, \bar{y}) = \{0\} \tag{3.4}$$

is satisfied. Then one has the inclusion

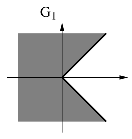
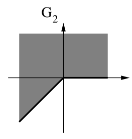
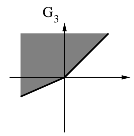
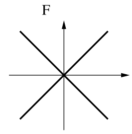
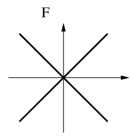
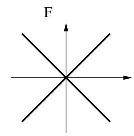
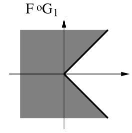
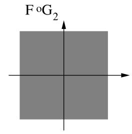
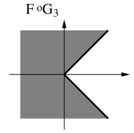
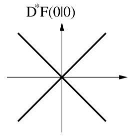
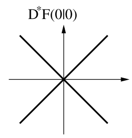
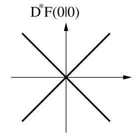
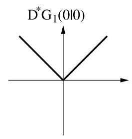
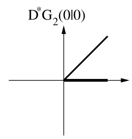
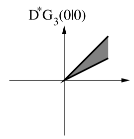
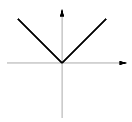
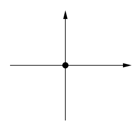
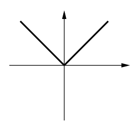
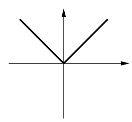
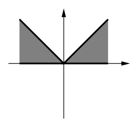
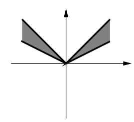
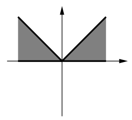
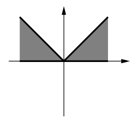
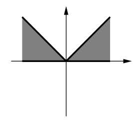
$$\begin{aligned} \partial f(\bar{x}) &\subset \bigcup_{(x^*, y^*) \in \partial\varphi(\bar{x}, \bar{y})} \big[x^* + D^*G(\bar{x} \mid \bar{y})\big(y^*\big)\big], \quad \text{and} \\ \partial_C f(\bar{x}) &\subset \overline{\text{co}} \bigcup_{(x^*, y^*) \in \partial\varphi(\bar{x}, \bar{y})} \big[x^* + \overline{\text{co}}\, D^*G(\bar{x} \mid \bar{y})\big(y^*\big)\big]. \end{aligned} \tag{3.5}$$

(2) *Assume that M is inner semicompact at $\bar{x}$, that G is closed-graph and φ is lsc on* Graph(G) *whenever x is near $\bar{x}$, and that the other assumptions in* (1) *are satisfied for every $\bar{y} \in M(\bar{x})$. Then one has analogs of inclusion* (3.5)*, where the sets on the right-hand sides are replaced by their unions over $\bar{y} \in M(\bar{x})$.*

(3) *Assume that $M(\cdot)$ is locally bounded at $\bar{x}$,* (3.4) *is satisfied for every $\bar{y} \in M(\bar{x})$, G is graph-convex and φ is convex. Then f is convex, and*

$$\partial f(\bar{x}) = \big\{x^* + D^*G(\bar{x} \mid \bar{y})\big(y^*\big) \;\big|\; \big(x^*, y^*\big) \in \partial\varphi(\bar{x}, \bar{y})\big\} \quad \textit{for any } \bar{y} \in M(\bar{x}).$$

Table 1
Possible scenarios in chain rule of set-valued maps from Example 3.2.

	$i=1$	2	3
G_i			
F			
$F \circ G_i$			
$D^*F(0 \mid 0)$			
$D^*G_i(0 \mid 0)$			
$\overline{\text{co}}\, D^*(F \circ G_i)(0 \mid 0)$			
$\overline{\text{co}}\, D^*G_i(0 \mid 0) \circ D^*F(0 \mid 0)$			
$\overline{\text{co}}\, D^*G_i(0 \mid 0) \circ \overline{\text{co}}\, D^*F(0 \mid 0)$			

Proof. Cases (1) and (2) are exactly the statement of [9, Theorem 3.38], and we prove only (3) from Theorem 3.1(3). Consider the map $\bar{G} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m$ defined by $\bar{G}(x) = \{x\} \times G(x)$. The coderivative $D^*\bar{G}(\bar{x} \mid (\bar{x}, \bar{y})) : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is easily evaluated to be

$$D^*\bar{G}\big(\bar{x} \bigm| (\bar{x}, \bar{y})\big)(p, q) = p + D^*G(\bar{x} \mid \bar{y})(q).$$

Noting that $E_f = E_\varphi \circ \bar{G}$, the constraint qualification we need to check in Theorem 3.1(3) is

$$D^* E_\varphi\big((\bar{x}, \bar{y}) \bigm| f(\bar{x})\big)(0) \cap -D^*\bar{G}^{-1}\big((\bar{x}, \bar{y}) \bigm| \bar{x}\big)(0) = \{0\}. \tag{3.6}$$

Note that $\partial^\infty \varphi(\bar{x}, \bar{y}) = D^* E_\varphi((\bar{x}, \bar{y}) \mid f(\bar{x}))(0)$. Now, $(p, q) \in D^*\bar{G}^{-1}((\bar{x}, \bar{y}) \mid \bar{x})(0)$ if and only if $(p, q, 0) \in N_{\mathrm{Graph}(\bar{G}^{-1})}(\bar{x}, \bar{y}, \bar{x})$, which is in turn equivalent to $(0, p, q) \in N_{\mathrm{Graph}(\bar{G})}(\bar{x}, \bar{x}, \bar{y})$. We see that $\mathrm{Graph}(\bar{G})$ is the image of a linear map of $\mathrm{Graph}(G)$ and use a rule of normal cones on linear maps in [12, Theorem 6.43] to obtain

$$N_{\mathrm{Graph}(\bar{G})}(\bar{x}, \bar{x}, \bar{y}) = \big\{(u, w, v) \bigm| (u + w, v) \in N_{\mathrm{Graph}(G)}(\bar{x}, \bar{y})\big\}.$$

Thus $(0, p, q) \in N_{\mathrm{Graph}(\bar{G})}(\bar{x}, \bar{x}, \bar{y})$ iff $(p, q) \in N_{\mathrm{Graph}(G)}(\bar{x}, \bar{y})$. Therefore (3.6) is equivalent to (3.4). We then apply Theorem 3.1(3) to get

$$\begin{aligned} \partial f(\bar{x}) &= D^* E_f\big(\bar{x} \bigm| f(\bar{x})\big)(1) \\ &= D^*(E_\varphi \circ \bar{G})\big(\bar{x} \bigm| f(\bar{x})\big)(1). \end{aligned}$$

Then for any $\bar{y} \in M(\bar{x})$,

$$\begin{aligned} \partial f(\bar{x}) &= D^*\bar{G}\big(\bar{x} \bigm| (\bar{x}, \bar{y})\big) \circ D^* E_\varphi\big((\bar{x}, \bar{y}) \bigm| f(\bar{x})\big)(1) \\ &= D^*\bar{G}\big(\bar{x} \bigm| (\bar{x}, \bar{y})\big)\big(\partial\varphi(\bar{x}, \bar{y})\big) \\ &= \big\{x^* + D^*G(\bar{x} \mid \bar{y})\big(y^*\big) \bigm| \big(x^*, y^*\big) \in \partial\varphi(\bar{x}, \bar{y})\big\}. \qquad \square \end{aligned}$$

We remark that [12, Section 10H] and [9, Section 3.2] contain other coderivative calculus rules that can be easily extended for the convexified limiting coderivative. As we have remarked after Theorem 3.1, the convexified limiting coderivative of G in Lemma 3.3 is sufficient for obtaining the Clarke subdifferential of f.

Remark 3.4 *(Alternative view of marginal functions).* A different view useful for later discussions is to consider

$$\begin{aligned} \min_{(x,y)} \quad & \varphi(x, y) \\ \text{s.t.} \quad & (x, y) \in \mathrm{Graph}(G). \end{aligned}$$

As is well known in nonlinear programming, if the point $(\bar{x}, \bar{y})$ is optimal, then $0 \in \partial\varphi(\bar{x}, \bar{y}) + N_{\mathrm{Graph}(G)}(\bar{x}, \bar{y})$. Recall that through the definition of coderivatives, $N_{\mathrm{Graph}(G)}(\bar{x}, \bar{y})$ is related to $\mathrm{Graph}(D^*G(\bar{x} \mid \bar{y}))$ by a linear transformation.

4. Subdifferential analysis of discretized inclusions

In this section, we consider the discretized inclusion and calculate the coderivatives of its reachable map. One can then obtain the subdifferential dependence of the differential inclusion in terms of its initial conditions. We can then obtain a necessary optimality condition of the discretized inclusion similar to the Euler–Lagrange and Transversality Conditions. Finally, we discuss the limitations of obtaining a discretized version of the Weierstrass–Pontryagin Maximum Principle.

We consider the following discrete inclusion as an analogue to the differential inclusion (1.1):

$$\begin{aligned} \min_{x_k\in\mathbb{R}^n \text{ for } k\in\overline{0,N}} \quad & \varphi(x_0,x_N) \\ \text{s.t.} \quad & x_k \in x_{k-1}+(\Delta t)F\big((k-1)(\Delta t),x_{k-1}\big). \end{aligned} \tag{4.1}$$

Here, $\Delta t = T/N$. The inclusion systems above can be further modified to one defined in terms of the reachable map. The discretized version of the reachable map $R_N:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ can be defined by

$$\begin{aligned} R_N(x_0)=\big\{x_N\colon\ \exists x_k\in\mathbb{R}^n \text{ for } k=\overline{1,N-1} \text{ s.t.}\\ x_k\in x_{k-1}+(\Delta t)F\big((k-1)(\Delta t),x_{k-1}\big) \text{ for all } k=\overline{1,N}\big\}. \end{aligned} \tag{4.2}$$

Then (4.1) can be rewritten as

$$\begin{aligned} \min_{x_0,x_N} \quad & \varphi(x_0,x_N) \\ \text{s.t.} \quad & x_N\in R_N(x_0)\subset\mathbb{R}^n. \end{aligned} \tag{4.3}$$

Theorem 4.1 *(Coderivatives of discretized reachable map). Recall the map $R_N:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ as defined in (4.3). Let $\Delta t = T/N$, and define $F_{k,N}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and $M_{k,N}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by*

$$F_{k,N}(\cdot):=F\big(k(\Delta t),\cdot\big) \quad \text{and} \quad M_{k,N}(x):=x+(\Delta t)F_{k,N}(x). \tag{4.4}$$

Note that $R_N = M_{N-1,N}\circ M_{N-2,N}\circ\cdots\circ M_{0,N}$. Assume that $F_{k,N}(\cdot)$ are osc and locally Lipschitz, and for all x, k and N, $F_{k,N}(\cdot)$ is locally bounded at x, i.e., there exists a neighborhood U of x and finite R such that $F_{k,N}(x')\subset R\mathbb{B}$ for all $x'\in U$.

(1) *For $x_N\in R(x_0)$, the coderivative of $R_N:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisfies*

$$D^*R_N(x_0\mid x_N)\subset \bigcup_{\{\tilde x_i\}_{i=0}^N\in\mathscr{X}_N} \underbrace{D^*M_{0,N}(\tilde x_0\mid\tilde x_1)\circ\cdots\circ D^*M_{N-1,N}(\tilde x_{N-1}\mid\tilde x_N)}_{G_{\{\tilde x_i\}_{i=0}^N}} \tag{4.5}$$

where

$$\mathscr{X}_N=\big\{\{\tilde x_i\}_{i=0}^N\colon\ \tilde x_k\in M_{k-1,N}(\tilde x_{k-1}) \text{ for all } k\in\overline{1,N},\ \tilde x_0=x_0 \text{ and } \tilde x_N=x_N\big\}. \tag{4.6}$$

(2) *If in addition $F_{k,N}(\cdot)$ are all graph convex, then*

$$D^*R_N(x_0\mid x_N)=D^*M_{0,N}(\tilde x_0\mid\tilde x_1)\circ\cdots\circ D^*M_{N-1,N}(\tilde x_{N-1}\mid\tilde x_N)$$

for any $\{\tilde x_i\}_{i=0}^N\in\mathscr{X}_N$.

(3) *Consider the conditions:*
  (a) *$p_0\in D^*R_N(x_0\mid x_N)(p_N)$.*
  (b) *There are $\{\tilde x_i\}_{i=0}^N\in\mathscr{X}_N$ and $\{\tilde p_i\}_{i=0}^N$ such that $p_0=\tilde p_0$, $p_N=\tilde p_N$ and*

$$\frac{p_k-p_{k-1}}{\Delta t}\in -D^*F_{k-1,N}\left(\tilde x_{k-1}\ \middle|\ \frac{1}{\Delta t}(\tilde x_k-\tilde x_{k-1})\right)(p_k) \quad \text{for all } k\in\overline{1,N}. \tag{4.7}$$

We have (a) *implies* (b)*, and in the case where each* $F_{k,N}(\cdot)$ *is graph convex for all* $k \in \overline{0,(N-1)}$*, the converse holds as well.*

Proof. For (1), the case where $N=2$ follows directly from Theorem 3.1(2). The local boundedness of $F_{k,N}(\cdot)$ ensures that $S(\cdot,\cdot)$ in Theorem 3.1(2) is inner semicompact, and the local Lipschitz continuity of $F_{k,N}(\cdot)$ implies the constraint qualification in (3.1) holds. The case for general N is easily deduced from the case where $N=2$. For (2), we follow the similar steps and apply Theorem 3.1(3).

To prove that (3)(a) implies (3)(b), let $G_{\{\tilde{x}_i\}_{i=0}^N}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be the formula as marked in (4.5). For $p_N, p_0 \in \mathbb{R}^n$, we have $p_0 \in G_{\{\tilde{x}_i\}_{i=0}^N}(p_N)$ if and only if there exists some $\{\tilde{p}_i\}_{i=0}^N$ such that $p_0=\tilde{p}_0$, $p_N=\tilde{p}_N$ and

$$p_{k-1} \in D^*M_{k-1}(\tilde{x}_{k-1} \mid \tilde{x}_k)(p_k) \quad \text{for all } k \in \overline{1,N}. \tag{4.8}$$

From the definition of $M_{k-1,N}$ and calculus rules for coderivatives in [12, Section 10H], we have

$$D^*M_{k-1,N}(\tilde{x}_{k-1} \mid \tilde{x}_k) = I + (\Delta t) D^*F_{k-1,N}\left(\tilde{x}_{k-1} \,\middle|\, \frac{1}{\Delta t}(\tilde{x}_k - \tilde{x}_{k-1})\right). \tag{4.9}$$

The formula (4.7) follows easily from (1). The converse holds due to (2). □

Putting together the previous results, we have the following necessary optimality condition for the discrete inclusion problem.

Theorem 4.2 *(Subdifferential analysis of discrete inclusions). For the discrete inclusion* (4.1)*, define* $F_{k,N}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *and* $M_{k,N}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *as in* (4.4)*,* $R_N:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *by* (4.2)*, and* $f:\mathbb{R}^n \to \mathbb{R}$ *by*

$$\begin{aligned} f(x_0) := \min_{x_N} \quad & \varphi(x_0, x_N) \\ \text{s.t.} \quad & x_N \in R_N(x_0) \subset \mathbb{R}^n. \end{aligned}$$

Suppose

- *for all k,* $F_{k,N}(\cdot)$ *is osc and locally Lipschitz,*
- *for all k, there is some finite* $b_{k,N}$ *such that* $F_{k,N}(x) \subset b_{k,N}\mathbb{B}$ *for all x, and*
- *the function* $\varphi(\cdot,\cdot)$ *is lsc.*

Then

$$\partial f(x_0) \subset \bigcup_{\substack{\{\tilde{x}_i\}_{i=0}^N \in \mathcal{X}_N,\, x_N \in R_N(x_0) \\ \varphi(x_0,x_N)=f(x_0)}} \left\{x^* + D^*G_{\{\tilde{x}_i\}_{i=0}^N}(x_0 \mid x_N)\big(y^*\big) \,\middle|\, \big(x^*, y^*\big) \in \partial\varphi(x_0, x_N)\right\},$$

where $G_{\{\tilde{x}_i\}_{i=0}^N}:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *and* $\mathcal{X}_N$ *are defined as in* (4.5) *and* (4.6)*. If in addition all the* $F_{k,N}(\cdot)$ *are all graph convex and* $\varphi(\cdot,\cdot)$ *is convex, we have*

$$\begin{aligned} \partial f(x_0) = {} & \left\{x^* + D^*G_{\{\tilde{x}_i\}_{i=0}^N}(x_0 \mid x_N)\big(y^*\big) \,\middle|\, \big(x^*, y^*\big) \in \partial\varphi(x_0, x_N)\right\} \\ & \text{for any } \{\tilde{x}_i\}_{i=0}^N \in \mathcal{X}_N \text{ s.t. } f(x_0) = \varphi(x_0, x_N). \end{aligned}$$

In particular (*not assuming the convexity of* Graph($F_{k,N}$) *and* $\varphi(\cdot,\cdot)$)*, a necessary condition for the optimality of the path* $\{\tilde{x}_i\}_{i=0}^N \in \mathcal{X}_N$ *is the existence of* $\{p_i\}_{i=0}^N$ *such that*

(1) $(-p_0, p_N) \in \partial\varphi(\tilde{x}_0, \tilde{x}_N)$, *and*
(2) $\frac{p_k - p_{k-1}}{\Delta t} \in -D^* F_{k-1,N}(\tilde{x}_{k-1} \mid \frac{1}{\Delta t}(\tilde{x}_k - \tilde{x}_{k-1}))(p_k)$ *for all* $k \in \overline{1,N}$.

Proof. Apply Theorem 4.1 and Lemma 3.3. □

Remark 4.3 *(Discrete Euler–Lagrange and Transversality Conditions).* Condition (1) in Theorem 4.2 is the discrete analogue of the Transversality Condition (TC), while condition (2) is the analogue of the Euler–Lagrange Condition (EL). A direct analysis without using coderivatives of the discretized reachable map allows us to drop the assumption that F is locally Lipschitz. See for example [8, Theorem 5.2], or [9, Theorem 6.17] for an infinite dimensional extension. For the continuous problem considered in Section 5, it seems that one cannot weaken the condition that $F(t,\cdot)$ is locally Lipschitz for almost every $t \in [0,T]$ to prove the Euler–Lagrange and Transversality Conditions.

Finally, we make a remark on the Weierstrass–Pontryagin Maximum Principle (WP). Before we do so, we recall that for $F : [0,T] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, the reachable map of the *relaxed* differential inclusion is defined by

$$\begin{aligned} R_{\overline{\text{co}}\, F}(x_0) := \big\{ y\colon\ \exists x(\cdot) \in AC\big([0,T],\mathbb{R}^n\big) \text{ s.t.} \\ x'(t) \in \overline{\text{co}}\, F\big(t, x(t)\big) \text{ for } t \in [0,T] \text{ a.e.}, \\ x(0) = x_0 \text{ and } x(T) = y \big\}. \end{aligned}$$

It is well known that under mild conditions, we have $\operatorname{cl} R(x) = R_{\overline{\text{co}}\, F}(x)$ for all $x \in \mathbb{R}^n$.

Remark 4.4 *(Discrete analogue of the Weierstrass–Pontryagin Maximum Principle).* Recall the chain rule for set-valued maps $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ as presented in Theorem 3.1. If the conclusion of the chain rule had been that for all $r \in \mathbb{R}^n$,

$$\begin{aligned} & D^*(F \circ G)(\bar{x} \mid \bar{z})(r) \\ & \quad \subset \bigcup_{\bar{y} \in F^{-1}(\bar{z}) \cap G(\bar{x})} \big\{ \overline{\text{co}}\, D^* G(\bar{x} \mid \bar{y})(q) \;\big|\; q \in \overline{\text{co}}\, D^* F(\bar{x} \mid \bar{y})(r),\ \langle q, \bar{y} - y' \rangle \leqslant 0 \text{ for all } y' \in G(\bar{x}) \big\}, \end{aligned} \tag{4.10}$$

then we can repeatedly apply this chain rule like in Theorem 4.1 so that under the conditions of Theorem 4.1, $p_0 \in D^* R_N(x_0 \mid x_N)(p_N)$ implies that there are $\{\tilde{x}_i\}_{i=0}^N$ and $\{\tilde{p}_i\}_{i=0}^N$ such that

$$\tilde{x}_0 = x_0, \qquad \tilde{x}_N = x_N, \qquad \tilde{p}_0 = p_0, \qquad \tilde{p}_N = p_N, \tag{4.11a}$$

$$\frac{\tilde{p}_k - \tilde{p}_{k-1}}{\Delta t} \in -D^* F_{k-1,N}\left(\tilde{x}_{k-1} \;\middle|\; \frac{1}{\Delta t}(\tilde{x}_k - \tilde{x}_{k-1})\right)(\tilde{p}_k) \quad \text{for all } k \in \overline{1,N}, \quad \text{and} \tag{4.11b}$$

$$\left\langle -\tilde{p}_k, v - \frac{1}{\Delta t}(\tilde{x}_k - \tilde{x}_{k-1}) \right\rangle \leqslant 0 \quad \text{for all } v \in F_{k-1,N}(\tilde{x}_{k-1}) \text{ and } k \in \overline{1,N}. \tag{4.11c}$$

Such a formula would be appealing because (4.11b) corresponds to the Euler–Lagrange Condition (EL) and (4.11c) corresponds to the Weierstrass–Pontryagin Maximum Principle (WP). However, (4.10) is not true in general. Consider the maps $G : \mathbb{R} \rightrightarrows \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ defined by

$$\begin{aligned} G(x) &:= [x+1, x+2] \cup [x-2, x-1], \quad \text{and} \\ f(x) &:= -|x - 0.5|. \end{aligned}$$

Then $f\circ G(0)=[-2.5,-0.5]$, and $f^{-1}(-0.5)\cap G(0)=\{1\}$. We can calculate that

$$\begin{aligned} D^*G(0\mid 1)(1)&=\{1\},\\ D^*f(1\mid -0.5)(-1)&=\{1\},\quad\text{and}\\ D^*(f\circ G)(0\mid -0.5)(-1)&=\{1\}. \end{aligned}$$

However, since we do not have $\langle 1,1-v\rangle\leqslant 0$ for all $v\in G(0)=[1,2]\cup[-2,-1]$, the right-hand side of (4.10) is empty, showing us that (4.10) cannot be true.

We now consider the case where $F_{k,N}:\mathbb{R}^n\rightrightarrows\mathbb{R}^n$ are convex-valued (so that we are considering the relaxed differential inclusion). It follows easily from the definitions that (4.11b) is equivalent to

$$\left(-\frac{\tilde{p}_k-\tilde{p}_{k-1}}{\Delta t},-\tilde{p}_k\right)\in N_{\mathrm{Graph}(F_{k-1,N})}\left(\tilde{x}_{k-1}\;\middle|\;\frac{1}{\Delta t}(\tilde{x}_k-\tilde{x}_{k-1})\right). \tag{4.12}$$

If $F_{k-1,N}(\tilde{x}_{k-1})$ is convex and $F_{k,N}$ is inner semicontinuous, then (4.11c) follows easily from (4.12). (See for example [8, Proposition 4.7] or [9, Theorem 1.34].)

5. Subdifferential analysis of differential inclusions

In this section, we make use of the work in Section 4 to calculate estimates of the convexified limiting coderivative of the (continuous) reachable map, and explain how this new formula gives a new way to interpret the Euler–Lagrange and Transversality Conditions.

We first simplify the notation. Define $\mathscr{F}(\bar{x},\bar{y})$ to be the set of feasible paths with end points $\bar{x}$ and $\bar{y}$, i.e.,

$$\begin{aligned}\mathscr{F}(\bar{x},\bar{y}):=\bigl\{x(\cdot)\;\big|\;x(\cdot)\in AC\bigl([0,T],\mathbb{R}^n\bigr),\ x(0)=\bar{x},\ x(T)=\bar{y},\\ \text{and } x'(t)\in F\bigl(t,x(t)\bigr)\text{ for } t\in[0,T]\text{ a.e.}\bigr\}.\end{aligned} \tag{5.1}$$

Define $\Pi:\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n\rightrightarrows\mathbb{R}^n$ by

$$\begin{aligned}\Pi(x,y,v):=\bigl\{u\;\big|\;&\exists x(\cdot)\in\mathscr{F}(x,y),\ p(\cdot)\in AC\bigl([0,T],\mathbb{R}^n\bigr)\\ &\text{s.t. } p(0)=u,\ p(T)=v\text{ and}\\ &p'(t)\in-\overline{\mathrm{co}}\,D_x^*F\bigl(t,x(t)\;\big|\;x'(t)\bigr)\bigl(p(t)\bigr)\text{ for } t\in[0,T]\text{ a.e.}\bigr\}.\end{aligned}$$

Here, $\overline{\mathrm{co}}\,D_x^*F(t,x(t)\mid x'(t)):\mathbb{R}^n\rightrightarrows\mathbb{R}^n$ is to be understood as $\overline{\mathrm{co}}\,D^*F_t(x(t)\mid x'(t)):\mathbb{R}^n\rightrightarrows\mathbb{R}^n$, where $F_t(\cdot)=F(t,\cdot)$. Corresponding to $\Pi(x,y,v)$ is its discretized version:

$$\begin{aligned}\Pi_N(x,y,v):=\Bigl\{u\;\Big|\;&\exists\{x_i\}_{i=0}^N,\{p_i\}_{i=0}^N\text{ s.t. } x_0=x,\ x_N=y,\ p_0=u,\ p_N=v,\\ &\frac{1}{\Delta t}(x_k-x_{k-1})\in F\bigl((k-1)(\Delta t),x_{k-1}\bigr)\text{ for all } k\in\overline{1,N},\text{ and}\\ &\frac{1}{\Delta t}(p_k-p_{k-1})\in-D^*F_{k-1,N}\left(x_{k-1}\;\middle|\;\frac{1}{\Delta t}(x_k-x_{k-1})\right)(p_k)\text{ for all } k\in\overline{1,N}\Bigr\}.\end{aligned}$$

We make the following conjecture.

Conjecture 5.1 *(Upper estimate of discretized coderivative of reachable map). For the reachable map $R:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined in* (1.2)*, the convexified coderivative $\overline{\mathrm{co}}\, D^*R(\bar{x}\mid\bar{y})$ satisfies*

$$\begin{aligned} D^*R(\bar{x}\mid\bar{y})(v) \subset \big\{u\colon \exists x(\cdot)\in\mathscr{F}(\bar{x},\bar{y}),\ p(\cdot)\in AC\big([0,T],\mathbb{R}^n\big) \text{ s.t.}\\ p'(t)\in -\overline{\mathrm{co}}\, D^*_x F\big(t,x(t)\,\big|\,x'(t)\big)\big(p(t)\big),\\ p_0=u \text{ and } p_T=v\big\} \quad \text{for all } v\in\mathbb{R}^n. \end{aligned} \tag{5.2}$$

Remark 5.2 *(Consequence of Conjecture 5.1).* Consider the problem

$$\begin{aligned} \min_{(x,y)}\quad & \varphi(x,y)\\ \text{s.t.}\quad & (x,y)\in \operatorname{Graph}(R). \end{aligned}$$

Recall the discussion in Remark 3.4. Provided (5.2) holds, if the point $(\bar{x},\bar{y})$ is optimal, then $0\in \partial\varphi(\bar{x},\bar{y})+N_{\operatorname{Graph}(R)}(\bar{x},\bar{y})$. We have

$$\begin{aligned} \partial\varphi(\bar{x},\bar{y})+N_{\operatorname{Graph}(R)}(\bar{x},\bar{y}) &= \partial\varphi(\bar{x},\bar{y})+L\operatorname{Graph}\big(D^*R(\bar{x},\bar{y})\big)\\ &\subset \partial\varphi(\bar{x},\bar{y})+L\operatorname{Graph}\big(\overline{\mathrm{co}}\, D^*R(\bar{x}\mid\bar{y})\big), \end{aligned}$$

where $L:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n\times\mathbb{R}^n$ is the linear map represented by the matrix $\left(\begin{smallmatrix}0 & I\\ -I & 0\end{smallmatrix}\right)$. Unrolling the definition of $D^*R(\bar{x}\mid\bar{y})$ gives the following optimality condition: If $(\bar{x},\bar{y})$ is optimal, then there are paths $x(\cdot), p(\cdot)\in AC([0,T],\mathbb{R}^n)$ such that $x(\cdot)$ is feasible for the differential inclusion, $x(0)=\bar{x}$, $x(T)=\bar{y}$, and satisfies the Transversality Condition (TC) and the Euler–Lagrange Condition (EL).

We will prove the following weaker result instead:

$$\begin{aligned} \overline{\mathrm{co}}\, D^*R(\bar{x}\mid\bar{y})(v) \subset \overline{\mathrm{co}}\,\big\{u\colon \exists x(\cdot)\in\mathscr{F}(\bar{x},\bar{y}),\ p(\cdot)\in AC\big([0,T],\mathbb{R}^n\big) \text{ s.t.}\\ p'(t)\in -\overline{\mathrm{co}}\, D^*_x F\big(t,x(t)\,\big|\,x'(t)\big)\big(p(t)\big),\\ p_0=u \text{ and } p_T=v\big\} \quad \text{for all } v\in\mathbb{R}^n. \end{aligned} \tag{5.3}$$

Our strategy is to prove the following three inclusions:

$$\overline{\mathrm{co}}\, D^*R(\bar{x}\mid\bar{y}) \subset \bigcap_{\substack{N\in\mathbb{N}\\ \delta>0}} \overline{\mathrm{co}} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} D^*R_i(x\mid y), \tag{5.4a}$$

$$\bigcap_{\substack{N\in\mathbb{N}\\ \delta>0}} \overline{\mathrm{co}} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} D^*R_i(x\mid y)(v) \subset \bigcap_{\substack{N\in\mathbb{N}\\ \delta>0}} \overline{\mathrm{co}} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} \Pi_i(x,y,v), \quad \text{and} \tag{5.4b}$$

$$\bigcap_{\substack{N\in\mathbb{N}\\ \delta>0}} \overline{\mathrm{co}} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} \Pi_i(x,y,v) \subset \overline{\mathrm{co}}\, \Pi(\bar{x},\bar{y},v), \tag{5.4c}$$

where (5.4b) and (5.4c) hold for all $v\in\mathbb{R}^n$. Conditions for $D^*R_i(x\mid y)(v)\subset \Pi_{p,i}(x,y,v)$, which addresses (5.4b), were discussed in Theorem 4.1. The same steps used to prove that (5.4b) and (5.4c) hold for all $v\in\mathbb{R}^n$ yield the following stronger statements: For all $v\in\mathbb{R}^n$,

$$\bigcap_{\substack{N\in\mathbb{N}\\ \delta>0}} \operatorname{cl} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} D^*R_i(x \mid y)(v) \subset \bigcap_{\substack{N\in\mathbb{N}\\ \delta>0}} \operatorname{cl} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} \Pi_i(x,y,v), \quad \text{and}$$

$$\bigcap_{\substack{N\in\mathbb{N}\\ \delta>0}} \operatorname{cl} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} \Pi_i(x,y,v) \subset \Pi(\bar{x},\bar{y},v).$$

Notice that if (5.4a) were strengthened to be

$$D^*R(\bar{x} \mid \bar{y}) \subset \bigcap_{\substack{N\in\mathbb{N}\\ \delta>0}} \operatorname{cl} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} D^*R_i(x \mid y)$$

instead, then piecing the last three formulas together gives (5.2). We continue with some lemmas.

Lemma 5.3 *(Coderivatives around $(\bar{x},\bar{y})$). Let $\delta > 0$, and $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be osc. Suppose $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a prefan such that*

$$H \in \mathscr{H}\Bigg(\overline{\text{co}} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} D^*S(x \mid y)\Bigg).$$

Let $\delta' := \min(\delta, \frac{\delta}{5\|H\|^+})$. Then

$$S(x') \cap \mathbb{B}_{\delta/2}(\bar{y}) \subset S(x'') + H(x' - x'') \quad \textit{for all } x', x'' \in \mathbb{B}_{\delta'}(\bar{x}).$$

Proof. For any $x \in \mathbb{B}_\delta(\bar{x})$ and $y \in \mathbb{B}_\delta(\bar{y})$, we have $H \in \mathscr{H}(D^*S(x \mid y))$. Choose any $\theta > 0$. There exists some $\epsilon_{x,y,\theta} > 0$ such that

$$S(x') \cap \mathbb{B}_{\epsilon_{x,y,\theta}}(y) \subset S(x'') + (H+\theta)(x' - x'') \quad \text{for all } x', x'' \in \mathbb{B}_{\epsilon_{x,y,\theta}}(x).$$

For each $x \in \mathbb{B}_\delta(\bar{x})$, we can find a finite number of elements in $\mathbb{B}_\delta(\bar{y})$, say $\{y_j\}_{j=1}^J$, such that $\mathbb{B}_\delta(\bar{y}) \subset \bigcup_{j=1}^J \mathbb{B}_{\epsilon_{x,y,\theta}}(y)$. Letting $\epsilon_{x,\theta} := \min_{j\in\overline{1,J}}(\epsilon_{x,y_j,\theta})$, we have

$$S(x') \cap \mathbb{B}_\delta(\bar{y}) \subset S(x'') + (H+\theta)(x' - x'') \quad \text{for all } x', x'' \in \mathbb{B}_{\epsilon_{x,\theta}}(x).$$

For any line segment $[x', x'']$ in $\mathbb{B}_\delta(\bar{x})$, we can find finitely many x in $\mathbb{B}_\delta(\bar{x})$, say $\{x_k\}_{k=1}^K$ such that $[x', x''] \subset \bigcup_{k=1}^K \mathbb{B}_{\epsilon_{x_k,\theta}}(x_k)$. We can break up the line segment $[x', x'']$ to a union of line segments $\bigcup_{j=1}^{J-1}[\tilde{x}_j, \tilde{x}_{j+1}]$ so that $\{\tilde{x}_j\}_{j=1}^J$ line up in that order, each $[\tilde{x}_j, \tilde{x}_{j+1}]$ is inside some $\mathbb{B}_{\epsilon_{x_k,\theta}}(x_k)$, $\tilde{x}_1 = x'$ and $\tilde{x}_J = x''$. Then

$$\begin{aligned} &S(\tilde{x}_j) \cap \mathbb{B}_\delta(\bar{y}) \subset S(\tilde{x}_{j+1}) + (H+\theta)(\tilde{x}_j - \tilde{x}_{j+1}) \\ &\quad\Rightarrow \quad \big[S(\tilde{x}_j) \cap \mathbb{B}_\delta(\bar{y})\big] + (H+\theta)(\tilde{x}_1 - \tilde{x}_j) \subset S(\tilde{x}_{j+1}) + (H+\theta)(\tilde{x}_1 - \tilde{x}_{j+1}). \end{aligned}$$

We write $\kappa = \|H\|^+$ to simplify notation. This gives

$$\begin{aligned} \big[S(\tilde{x}_j) + (H+\theta)(\tilde{x}_1 - \tilde{x}_j)\big] \cap \mathbb{B}_{\delta-(\kappa+\theta)|x'-x''|}(\bar{y}) &\subset \big[S(\tilde{x}_j) \cap \mathbb{B}_\delta(\bar{y})\big] + (H+\theta)(\tilde{x}_1 - \tilde{x}_j) \\ &\subset S(\tilde{x}_{j+1}) + (H+\theta)(\tilde{x}_1 - \tilde{x}_{j+1}), \end{aligned}$$

which implies

$$\begin{aligned}&\big[S(\tilde{x}_j)+(H+\theta)(\tilde{x}_1-\tilde{x}_j)\big]\cap\mathbb{B}_{\delta-(\kappa+\theta)|x'-x''|}(\bar{y})\\&\quad\subset\big[S(\tilde{x}_{j+1})+(H+\theta)(\tilde{x}_1-\tilde{x}_{j+1})\big]\cap\mathbb{B}_{\delta-(\kappa+\theta)|x'-x''|}(\bar{y}).\end{aligned}\tag{5.5}$$

Consider the case where $\theta<\kappa/4$ so that $4(\kappa+\theta)<5\kappa$. If $x',x''\in\mathbb{B}_{\delta'}(\bar{x})$, where $\delta'=\min(\delta,\frac{\delta}{5\kappa})$, then

$$(\kappa+\theta)\big|x'-x''\big|\leqslant\frac{5}{4}\kappa\frac{2\delta}{5\kappa}\leqslant\frac{\delta}{2}.$$

Recalling that $\tilde{x}_1=x'$ and $\tilde{x}_J=x''$ and applying (5.5) repeatedly, we have

$$S\big(x'\big)\cap\mathbb{B}_{\delta/2}(\bar{y})\subset S\big(x''\big)+(H+\theta)\big(x'-x''\big).$$

The above holds for all $x',x''\in\mathbb{B}_{\delta'}(\bar{x})$ and for all $\theta>0$, and hence for $\theta=0$, giving us the conclusion we need. □

This result gives a handle on the left-hand bound.

Lemma 5.4 *(On (5.4a)). Let $S:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ be a closed set-valued map. Suppose $\{S_i(\cdot)\}_{i=1}^{\infty}$, where $S_i:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$, are osc set-valued maps such that for any $\epsilon>0$ and $x\in\mathbb{R}^n$, there is some I such that*

$$\mathbf{d}\big(S(x),S_i(x)\big)<\epsilon\quad\text{for all } i>I.\tag{5.6}$$

Then for any $\delta>0$ and positive integer N, we have

$$\overline{\text{co}}\,D^*S(\bar{x}\mid\bar{y})\subset\overline{\text{co}}\bigcup_{i>N}\bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\y\in\mathbb{B}_\delta(\bar{y})}}D^*S_i(x\mid y).$$

Proof. By Lemma 2.8, we can prove that the following holds for all $\delta>0$ and positive integers N instead:

$$\mathscr{H}\big(\overline{\text{co}}\,D^*S(\bar{x}\mid\bar{y})\big)\supset\mathscr{H}\Bigg(\overline{\text{co}}\bigcup_{i>N}\bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\y\in\mathbb{B}_\delta(\bar{y})}}D^*S_i(x\mid y)\Bigg).$$

Suppose $H:\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ is a prefan in the RHS. Then for any $i>N$,

$$H\in\mathscr{H}\Bigg(\overline{\text{co}}\bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\y\in\mathbb{B}_\delta(\bar{y})}}D^*S_i(x\mid y)\Bigg).$$

By Lemma 5.3, if $\delta'=\min(\delta,\frac{\delta}{5\|H\|^+})$, then

$$S_i\big(x'\big)\cap\mathbb{B}_{\delta/2}(\bar{y})\subset S_i\big(x''\big)+H\big(x'-x''\big)\quad\text{for all } x',x''\in\mathbb{B}_{\delta'}(\bar{x}).$$

For all $x', x'' \in \mathbb{B}_{\delta'}(\bar{x})$ and $\epsilon > 0$, we can find i large enough so that

$$\begin{aligned} S(x') \cap \mathbb{B}_{\delta/2}(\bar{y}) &\subset \left[S_i(x') + \epsilon\mathbb{B}\right] \cap \mathbb{B}_{\delta/2}(\bar{y}) \\ &\subset S_i(x'') + H(x' - x'') + \epsilon\mathbb{B} \\ &\subset S(x'') + H(x' - x'') + 2\epsilon\mathbb{B}. \end{aligned}$$

The above holds for all $\epsilon > 0$, and we have

$$S(x') \cap \mathbb{B}_{\delta/2}(\bar{y}) \subset S(x'') + H(x' - x'') \quad \text{for all } x', x'' \in \mathbb{B}_{\delta'}(\bar{x}).$$

This implies that $H \in \mathscr{H}(\overline{\text{co}}\, D^*S(\bar{x} \mid \bar{y}))$ as needed. □

Remark 5.5 *(On formula (5.6)).* We note that conditions for $\mathbf{d}(S(x), S_i(x)) < \epsilon$ were given in [4], and in particular, conditions for $S(x) \subset S_i(x) + \epsilon\mathbb{B}$ were given in [9, Theorem 6.4] for example.

Note that Theorem 4.1 says that $D^*S_i(x \mid y)(v) \subset \Pi_i(x, y, v)$. To find suitable conditions for (5.4c), we need the following result.

Lemma 5.6 *(Convexification of intersection of nested sets). Suppose $\{A_i\}_{i=1}^{\infty} \subset \mathbb{R}^n$ are nested compact sets such that $A_{i+1} \subset A_i$. Then $\overline{\text{co}} \bigcap_i A_i = \bigcap_i \overline{\text{co}}\, A_i$.*

Proof. Suppose x is in the LHS. Then $x \in \overline{\text{co}}\, A_i$ for all i, so $x \in \bigcap_i \overline{\text{co}}\, A_i$, establishing $\overline{\text{co}} \bigcap_i A_i \subset \bigcap_i \overline{\text{co}}\, A_i$.

Next, suppose x is in the RHS. Then $x \in \overline{\text{co}}\, A_i$ for all i. Consider any $v \in \mathbb{R}^n \backslash \{0\}$. Since $x \in \overline{\text{co}}\, A_i$, we have $v^T x \leqslant \sup_{a \in A_i} v^T a$. By the compactness of A_i, let $\bar{a}_i$ be such that $v^T \bar{a}_i = \sup_{a \in A_i} v^T a$. Since $\bigcap_j A_j \subset A_i$ for all i, it is clear that $\sup_{a \in \bigcap_j A_j} v^T a \leqslant \sup_{a \in A_i} v^T a$ for all i, so $\sup_{a \in \bigcap_j A_j} v^T a \leqslant \inf_i \sup_{a \in A_i} v^T a$. By the compactness of A_i, the limit $\bar{a} = \lim_{j \to \infty} \bar{a}_j$ exists and lies in $\bigcap_j A_j$. This shows that

$$\begin{aligned} \inf_i \sup_{a \in A_i} v^T a &= \inf_i v^T \bar{a}_i \\ &= v^T \bar{a} \\ &\leqslant \sup_{a \in \bigcap_j A_j} v^T a. \end{aligned}$$

Then $v^T x \leqslant \sup_{a \in \bigcap_j A_j} v^T a$, which holds for all v. Thus we have $x \in \overline{\text{co}} \bigcap_i A_i$, so $\overline{\text{co}} \bigcap_i A_i = \bigcap_i \overline{\text{co}}\, A_i$ as needed. □

Here is a lemma useful for proving our next result. We take our result from [14, Lemma 4.4].

Lemma 5.7 *(Continuous solutions from discrete solutions). Assume that a set-valued map $F : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ has closed convex values. Let the set-valued map $(x, y) \mapsto F(t, x, y)$ be upper semicontinuous for almost all $t \in [0, T]$, and let $F(t, x, y) \subset b(t)\mathbb{B}$ for all $(t, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$, where $b(\cdot) \in L_1([0, T], \mathbb{R})$. Assume that functions $x_k(\cdot) \in AC([0, T], \mathbb{R}^n)$, $k = 0, 1, \ldots,$ satisfy*

$$x_k'(t) \in \overline{\text{co}}\, F(t, x_k(t), \eta_k(t)\mathbb{B}_m) + \eta_k(t)\mathbb{B}_n,$$

where $\eta_k \geqslant 0$, $\lim_{k\to\infty} \eta_k(t) = 0$ almost everywhere, and $|\eta_k(t)| \leqslant \eta(t)$, $k = 1, 2, \ldots$, $\eta(\cdot) \in L_1([0,T],\mathbb{R})$. Then the functions $x_k(\cdot)$ are equicontinuous on $[0,T]$; and if a subsequence $x_{k_p}(\cdot)$ uniformly converges to a function $x(\cdot)$, then $x(\cdot)$ is a solution of the differential inclusion

$$x'(t) \in F\big(t, x(t), 0\big) \quad \text{for } t \in [0,T] \text{ a.e.}$$

Before we state our main result, we describe in detail the paths produced by discrete approximations in the remark below.

Remark 5.8 *(Discrete approximations).* For $\{x_{i,j}\}_{\substack{0\leqslant j\leqslant i\\ 1\leqslant i\leqslant\infty}}$ and $\{p_{i,j}\}_{\substack{0\leqslant j\leqslant i\\ 1\leqslant i\leqslant\infty}}$, let $\Delta t = T/i$, and construct the following path $x_i : [0,T] \to \mathbb{R}^n$ defined by

$$x_i(t) = \frac{t - j\Delta t}{\Delta t} x_{i,j+1} + \frac{(j+1)\Delta t - t}{\Delta t} x_{i,j} \quad \text{whenever } t \in \big[j\Delta t, (j+1)\Delta t\big], \quad \text{and}$$
$$p_i(t) = \frac{t - j\Delta t}{\Delta t} p_{i,j+1} + \frac{(j+1)\Delta t - t}{\Delta t} p_{i,j} \quad \text{whenever } t \in \big[j\Delta t, (j+1)\Delta t\big].$$

It is clear that $x_i(\cdot)$ and $p_i(\cdot)$ are piecewise differentiable at all points other than integer multiples of Δt, and the derivatives satisfy

$$x_i'(t) = \frac{1}{\Delta t}(x_{i,j+1} - x_{i,j}) \quad \text{whenever } t \in \big(j\Delta t, (j+1)\Delta t\big), \quad \text{and} \tag{5.7a}$$
$$p_i'(t) = \frac{1}{\Delta t}(p_{i,j+1} - p_{i,j}) \quad \text{whenever } t \in \big(j\Delta t, (j+1)\Delta t\big). \tag{5.7b}$$

We also need the following condition for Lemma 5.9, which was one of the conclusions in Theorem 4.2:

$$\frac{p_k - p_{k-1}}{\Delta t} \in -D^* F_{k-1,N}\left(x_{k-1} \;\middle|\; \frac{1}{\Delta t}(x_k - x_{k-1})\right)(p_k) \quad \text{for all } k \in \overline{1,N}. \tag{5.8}$$

We now prove our result on (5.4c). Note that (5.4c) represents a closedness property, and we shall show that Lemma 5.7 provides some reasonable conditions for (5.4c) to hold.

Lemma 5.9 *(On (5.4c)). Suppose $F : [0,T] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is osc. Assume further that there is some $b(\cdot) \in L_1([0,T],\mathbb{R}^n)$ such that $b(t)$ is finite for all t and $\overline{\text{co}}\, D_x^* F(t, x \mid y)(p) \subset b(t)\|p\|\mathbb{B}$ for all $(t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$. Suppose also that the following assumption holds*:

(1) *If $u \in \bigcap_{N\in\mathbb{N},\,\delta>0} \text{cl} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} \Pi_i(x,y,v)$, then there exist $x(\cdot) \in AC([0,T],\mathbb{R}^n)$ and subsequences of paths $\{x_{i_k}(\cdot)\}_{k=1}^\infty$ and $\{p_{i_k}(\cdot)\}_{k=1}^\infty$ of $\{x_i(\cdot)\}_{i=1}^\infty$ and $\{p_i(\cdot)\}_{i=1}^\infty$ respectively such that*
   - *$\{x_i(\cdot)\}_{i=1}^\infty$ and $\{p_i(\cdot)\}_{i=1}^\infty$ are constructed based on discrete approximations $\{x_{i,j}\}_{\substack{0\leqslant j\leqslant i\\ 1\leqslant i\leqslant\infty}}$ and $\{p_{i,j}\}_{\substack{0\leqslant j\leqslant i\\ 1\leqslant i\leqslant\infty}}$ satisfying* (5.8) *as described in Remark* 5.8,
   - *$x_{i_k}(0) \to \bar{x}$, $x_{i_k}(T) \to \bar{y}$, $p_{i_k}(0) \to u$ and $p_{i_k}(T) \to v$ as $k \to \infty$,*
   - *$x_{i_k}(\cdot)$ converges uniformly to $x(\cdot)$,*
   - *$x_{i_k}'(\cdot)$ converges pointwise almost everywhere to $x'(\cdot)$,*

- *$x(\cdot)$ satisfies the differential inclusion*

$$x'(t) \in F\big(t, x(t)\big) \quad a.e.,$$
$$x(0) = \bar{x} \quad and \quad x(T) = \bar{y},$$

- *and $\{p_{i_k}(\cdot)\}_{k=1}^{\infty}$ converges uniformly to some $p(\cdot)$.*

Then we have

$$\bigcap_{N\in\mathbb{N},\,\delta>0} \operatorname{cl} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} \Pi_i(x, y, v) \subset \Pi(\bar{x}, \bar{y}, v), \quad and \tag{5.9}$$

$$\bigcap_{N\in\mathbb{N},\,\delta>0} \overline{\operatorname{co}} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} \Pi_i(x, y, v) \subset \overline{\operatorname{co}}\, \Pi(\bar{x}, \bar{y}, v). \tag{5.10}$$

Proof. First, we note that (5.9) implies (5.10). If (5.9) holds, then by Lemma 5.6 we have

$$\begin{aligned}\bigcap_{N\in\mathbb{N},\,\delta>0} \overline{\operatorname{co}} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} \Pi_i(x, y, v) &= \overline{\operatorname{co}} \bigcap_{N\in\mathbb{N},\,\delta>0} \operatorname{cl} \bigcup_{i>N} \bigcup_{\substack{x\in\mathbb{B}_\delta(\bar{x})\\ y\in\mathbb{B}_\delta(\bar{y})}} \Pi_i(x, y, v)\\ &\subset \overline{\operatorname{co}}\, \Pi(\bar{x}, \bar{y}, v).\end{aligned}$$

Proving (5.9) is equivalent to proving the following: If $u_i \in \Pi_i(\bar{x}_i, \bar{y}_i, v)$ and $u_i \to u$, $\bar{x}_i \to \bar{x}$ and $\bar{y}_i \to \bar{y}$ as $i \to \infty$, then $u \in \Pi(\bar{x}, \bar{y}, v)$. Consider $v \in \mathbb{R}^n$ and the sequences of functions $\{x_i(\cdot)\}_{i=1}^{\infty}$ and $\{p_i(\cdot)\}_{i=1}^{\infty}$ constructed from $\{x_{i,j}\}_{\substack{0\leqslant j\leqslant i\\ 1\leqslant i\leqslant\infty}}$ and $\{p_{i,j}\}_{\substack{0\leqslant j\leqslant i\\ 1\leqslant i\leqslant\infty}}$ such that $x_i(0) = \bar{x}_i$, $x_i(T) = \bar{y}_i$, $p_i(0) = u_i$ and $p_i(T) = v$. We therefore need to show that $u \in \Pi(\bar{x}, \bar{y}, v)$.

For a fixed $t \in [0, T]$, the map

$$(p, \tilde{x}, \tilde{y}, \tilde{p}) \mapsto -D_x^* F\big(t, x(t) + \tilde{x} \,\big|\, x'(t) + \tilde{y}\big)(p + \tilde{p})$$

can be checked to be osc (at where $x(t)$ and $x'(t)$ are defined) from the definition of the coderivatives and the fact that the map $(x, y) \mapsto N_{\operatorname{Graph}(F(t,\cdot))}(x, y)$ is osc. The map

$$(p, \tilde{x}, \tilde{y}, \tilde{p}) \mapsto -\overline{\operatorname{co}}\, D_x^* F\big(t, x(t) + \tilde{x} \,\big|\, x'(t) + \tilde{y}\big)(p + \tilde{p})$$

is osc since the convex hull operation preserves outer semicontinuity. (The proof is elementary, and the steps are shown in [11] for example.)

Suppose $x(\cdot)$ is such that assumption (1) in the statement holds. Our problem can be solved if we can show that $p(\cdot)$ satisfies the differential inclusion

$$p'(t) \in -\overline{\operatorname{co}}\, D_x^* F\big(t, x(t) \,\big|\, x'(t)\big)\big(p(t)\big).$$

We try to find $\eta_k : [0, T] \to [0, \infty)$ such that for all $t \in [0, T]$,

$$\lim_{k\to\infty} \eta_k(t) = 0 \quad \text{and}$$
$$p'_{i_k}(t) \in -\overline{\operatorname{co}}\, D_x^* F\big(t + \eta_k(t)\mathbb{B}, x(t) + \eta_k(t)\mathbb{B} \,\big|\, x'(t) + \eta_k(t)\mathbb{B}\big)\big(p_{i_k}(t) + \eta_k(t)\mathbb{B}\big). \tag{5.11}$$

For each $t \in [0,T]$ and $k \in \overline{1,\infty}$, we have $\lfloor t/(\Delta t)\rfloor(\Delta t) \leqslant t \leqslant \lfloor t/(\Delta t)+1\rfloor(\Delta t)$, where $\Delta t = T/i_k$ and $\lfloor\alpha\rfloor$ is the greatest integer not more than α. For simplicity, we consider the case where t/T is irrational. From the definitions of $x_{i_k}(\cdot)$ and $p_{i_k}(\cdot)$ and (5.8), we have

$$p'_{i_k}(t) \in -\overline{\mathrm{co}}\, D^*_x F\big(t_k, x_{i_k}(t_k) \bigm| x'_{i_k}(t)\big)\big(p_{i_k}(t_k+\Delta t)\big),$$

where $t_k = \lfloor t/(\Delta t)\rfloor(\Delta t)$. To establish the existence of $\eta_k(\cdot)$ in (5.11), it suffices to show that for each t,

$$\max\big(|t_k - t|, \big\|x_{i_k}(t_k) - x(t)\big\|, \big\|x'_{i_k}(t) - x'(t)\big\|, \big\|p_{i_k}(t_k+\Delta t) - p_{i_k}(t)\big\|\big) \searrow 0 \quad \text{as } k \nearrow \infty.$$

We first have $x'_{i_k}(t) \to x'(t)$ and $t_k \to t$ as $k \to \infty$. Next, since $p_{i_k}(\cdot)$ converges uniformly to $p(\cdot)$, we have

$$\begin{aligned}&\big\|p_{i_k}(t_k+\Delta t) - p_{i_k}(t)\big\| \\ &\quad \leqslant \underbrace{\big\|p_{i_k}(t_k+\Delta t) - p(t_k+\Delta t)\big\|}_{(1)} + \underbrace{\big\|p(t_k+\Delta t) - p(t)\big\|}_{(2)} + \underbrace{\big\|p(t) - p_{i_k}(t)\big\|}_{(3)},\end{aligned} \tag{5.12}$$

so the term on the LHS converges to zero as $k \to \infty$. A similar argument with $x_{i_k}(t_k) - x(t)$ shows that its norm goes to zero as $k \to \infty$. So the presence of $\eta_k(t)$ satisfying (5.11) is established.

Since $p(\cdot)$ is continuous on the compact set $[0,T]$, it is uniformly continuous. This implies that for any $\epsilon > 0$, we can find K such that term (2) in (5.12) has norm less than ϵ for all $k > K$. The condition that $\eta_k(t) \leqslant \eta(t)$ for all $t \in [0,T]$ for some $\eta(\cdot) \in L_1([0,T],\mathbb{R}^n)$ (in fact, $\eta(\cdot) \in L_\infty([0,T],\mathbb{R}^n)$) follows easily. All the conditions for Lemma 5.7 are satisfied, and we have $u \in \Pi(\bar{x},\bar{y},v)$ as needed. □

Though condition (1) may look more complicated than (5.4c) alone, it can be understood as a measurability condition on $x(\cdot)$ and $p(\cdot)$. We collect the previous results to obtain an estimate of the convexified limiting coderivative of the reachable map.

Theorem 5.10 *(Convexified coderivative of reachable map). The formula* (5.3) *holds provided*:

(a) *For any* $\epsilon > 0$ *and* $x \in \mathbb{R}^n$, *there is some* I *such that* $\mathbf{d}(R(x), R_i(x)) < \epsilon$ *for all* $i > I$.
(b) *For all* x *and* t, $F(t,\cdot)$ *is osc and locally bounded at* x.
(c) *There is some* $b(\cdot) \in L_1([0,T],\mathbb{R}^n)$ *such that* $b(t)$ *is finite for all* t, *and either one of the following conditions equivalent under condition* (b) *holds*:
  - $\|D^*_x F(t,x \mid y)\|^+ \leqslant b(t)$ *for all* $(t,x,y) \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^n$,
  - $F(t,\cdot)$ *is locally Lipschitz with modulus at most* $b(t)$ *for all* $t \in [0,T]$.

(d) *Assumption* (1) *of Lemma* 5.9 *holds.*

Proof. This combines Lemma 5.4, Theorem 4.1 and Lemma 5.9. We can check that the requirements for Theorem 4.1 are satisfied. The condition $\|D^*_x F(t,x \mid y)\|^+ \leqslant b(t)$ in (c) is equivalent to the condition on $\overline{\mathrm{co}}\, D^*_x F(t,x \mid y)$ in Lemma 5.9. Under condition (b), the two conditions in (c) are equivalent due to the Mordukhovich criterion. (See for example, [12, Chapter 9].) □

Conditions (b) and (c) are typical assumptions for (EL), (TC) and (WP) to hold. Condition (a) is a mild assumption on how the discretized reachable map can approximate the continuous reachable map, and condition (d) relates the discretized paths to continuous paths. The procedure of passing a sequence of discrete problems to the limit seems to make it unavoidable that assumption (d) has to hold, and that the conclusion can only be expressed in terms of convexified limiting coderivatives. The conditions (EL), (TC) and (WP) are usually proved with direct methods in analysis rather than through discrete approximations, so it remains to be seen whether Theorem 5.10 can be further strengthened with such techniques.

Remark 5.11 *(Graph convex $F(t,\cdot)$).* The discrete case suggests that when $F(t,\cdot)$ is graphically convex for all t, then (5.2) is actually an equation. For the continuous case, we study (5.3) instead, and ask whether (5.3) is an equation when $F(t,\cdot)$ is graphically convex for all t. In this case, (5.4b) is an equation, but equality for (5.4a) requires further assumptions. The reverse inclusion for (5.4c) holds if every continuous path on the RHS can be described as a limit of sequences on the left-hand side. Such results may already be in the literature. We cite [14, Theorem 4.16] for example, which states that the reverse inclusion in (5.4c) holds when $F(\cdot,\cdot)$ is independent of its first argument t and is Lipschitz.

6. Conclusion

In this paper, we study how discrete and differential inclusions depend on the initial conditions. The advantage of such results over necessary optimality conditions is that such results give an indication of how to perturb the initial point to optimality. The results for discrete inclusions seem quite satisfactory, but the results for differential inclusions still require further improvement.

Acknowledgments

The author thanks the referee for helpful suggestions and historical remarks which have improved the paper. He also thanks Alexander Ioffe and Boris Mordukhovich for conversations in October 2011 on the topic in this paper, as well as the co-editor-in-chief Alberto Bressan for the prompt handling of the paper. Finally, he thanks the Department of Mathematics at MIT for providing an excellent environment for working on this paper.

References

[1] J.-P. Aubin, A. Cellina, Differential Inclusions: Set-Valued Maps and Viability Theory, Grundlehren Math. Wiss., vol. 264, Springer, New York, 1984.
[2] J.-P. Aubin, H. Frankowska, Set-Valued Analysis, Birkhäuser, Boston, 1990; Republished as a Modern Birkhäuser Classic, 2009.
[3] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, Philadelphia, 1983; Republished as a SIAM Classic Appl. Math., 1990.
[4] A.L. Dontchev, F. Lempio, Difference methods for differential inclusions: A survey, SIAM Rev. 34 (2) (1992) 263–294.
[5] A.L. Dontchev, M. Quincampoix, N. Zlateva, Aubin criterion for metric regularity, J. Convex Anal. 3 (2006) 45–63.
[6] A.D. Ioffe, Nonsmooth analysis: differential calculus of non-differentiable mappings, Trans. Amer. Math. Soc. 266 (1981) 1–56.
[7] B.S. Mordukhovich, Metric approximations and necessary optimality conditions for general classes of extremal problems, Soviet Math. Dokl. 22 (1980) 526–530.
[8] B.S. Mordukhovich, Discrete approximations and refined Euler–Lagrange conditions for nonconvex differential inclusions, SIAM J. Control Optim. 33 (1995) 882–915.
[9] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation I, Grundlehren Math. Wiss., vol. 330, Springer, Berlin, 2006;
B.S. Mordukhovich, Variational Analysis and Generalized Differentiation II, Grundlehren Math. Wiss., vol. 331, Springer, Berlin, 2006.
[10] C.H.J. Pang, Generalized differentiation with positively homogeneous maps: Applications in set-valued analysis and metric regularity, Math. Oper. Res. 36 (3) (2011) 377–397.
[11] C.H.J. Pang, Characterizing generalized derivatives of set-valued maps: Extending tangential and normal approaches, 2012; Available in http://arxiv.org/abs/1106.2338.
[12] R.T. Rockafellar, R.J.-B. Wets, Variational Analysis, Grundlehren Math. Wiss., vol. 317, Springer, Berlin, 1998.
[13] G. Smirnov, Discrete approximations and optimal solutions to differential inclusions, Cybernetics 27 (1991) 101–107.
[14] G. Smirnov, Introduction to the Theory of Differential Inclusions, Grad. Stud. Math., vol. 41, Amer. Math. Soc., Providence, RI, 2002.
[15] R.B. Vinter, Optimal Control, Birkhäuser, Boston, 2000; Republished as a Modern Birkhäuser Classic, 2010.