Серия «Математика»
2019. Т. 30. С. 45—58

Онлайн-доступ к журналу:
http://mathizv.isu.ru

ИЗВЕСТИЯ
Иркутского
государственного
университета

УДК 517.97
MSC 49J52, 49J53, 49K24, 49M25, 90C30
DOI https://doi.org/10.26516/1997-7670.2019.30.45

Optimal Control of Differential Inclusions, I: Lipschitzian case *

B. S. Mordukhovich
Wayne State University, Detroit, USA

In memory of Oleg Vasiliev, a colleague and friend

Abstract. We develop the method of discrete approximations to study optimal control problems for differential inclusions by using advanced tools of variational analysis and generalized differentiation. The first part describes the method, appropriate machinery of variational analysis and then presents the main result on necessary optimality conditions in maximum principle form for Lipschitzian differential inclusions.

Keywords: optimal control, Lipschitzian differential inclusions, variational analysis, discrete approximations, generalized differentiation.

1. Introduction

Classical optimal control theory deals with dynamical systems governed by controlled ordinary differential equations

$$\dot{x}(t) = f\big(t, x(t), u(t)\big), \quad u(t) \in U \quad \text{a.e.} \quad t \in [a, b], \tag{1.1}$$

in the class of measurable controls $u(\cdot)$, where $f\colon [a,b] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ is a vector function that is continuously differentiable in x, and where U is a compact set. The main result there is the *Pontryagin maximum principle* (PMP), which provides necessary optimality conditions for strong local minimizers via the maximization of a certain Hamiltonian function; see

* This research was partly supported by the USA National Science Foundation under grants DMS-1512846 and DMS-1808978, by the USA Air Force Office of Scientific Research under grant 15RT04, and by Australian Research Council, Discovery Project under grant DP-190100555.

[14] and further developments in [1; 4; 6; 17; 18] with the references therein, where the reader can also find extensions of the classical PMP to various hereditary systems, nonsmooth problems, partial differential equations of parabolic and hyperbolic types, and other controlled dynamical systems.

More recently, optimal control theory has been extended to dynamical systems without explicit control parameterizations by considering *differential inclusions* of the type

$$\dot{x}(t) \in F\big(t, x(t)\big) \quad \text{a.e.} \quad t \in [a, b] \tag{1.2}$$

in the class of absolutely continuous trajectories $x(\cdot)$, where $F\colon [a, b] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued mapping/multifunction acting in finite-dimensional spaces. We refer the reader to the books [1;6;10;18] with the vast bibliographies and commentaries therein for various results on optimization problems for differential inclusions obtained under Lipschitzian assumptions on the set-valued velocity mapping $F(t, \cdot)$ in (1.2). Observe that optimization problems for Lipschitzian differential inclusions are *intrinsically nonsmooth*, and thus their study requires appropriate tools of generalized differentiation. Necessary optimality conditions for such problems were obtained in extended Euler-Lagrange and Hamiltonian forms (including maximization conditions of the Weierstrass-Pontryagin type); see the aforementioned monographs for more details.

It is worth mentioning that the differential inclusion framework (1.2) covers (via measurable selection theorems) not only the standard optimal control setting (1.2) with constant control sets U (which may evolve in time), but also much more challenging situations where control sets depend on state variables $U = U(t, x)$. The latter setting corresponds to the representation $F(t, x) = f(t, x, U(t, x))$ in (1.2) while reflecting a certain *feedback control* effect that is crucial, in particular, for engineering design. Observe also that the differential inclusion formalism arises not only in describing the parameterized control systems of type (1.1) with $U = U(t, x)$, but in other numerous applications to economic, mechanic, and behavioral science models that do not involve any control parametrization.

In this paper we mainly discuss a constructive approach to the study and solving of optimization problems for differential inclusions that is based on the *method of discrete approximations*. This approach clearly has a computational flavor to justify the possibility of the numerical solution of infinite-dimensional optimization problems by optimizing their finite-dimensional discrete-time counterparts. But our major goal here is to derive necessary optimality conditions for the original infinite-dimensional control problems by reducing them to finite-dimensional ones and employing optimality conditions in mathematical programming. The original idea of this approach goes back to Euler [3] who used it to obtain a necessary optimality condition ("Euler equation") for a specific ("simplest") problem of the calculus of variations on minimizing a particular integral functional depending on the velocity variable.

The development of this idea in problems with dynamic constraints of type (1.2), or even of the standard optimal control type (1.1) with smooth dynamics, is significantly more challenging. The reader is referred to the author's books [6;10] with the extensive bibliographies and commentaries therein for the implementation of this approach in various classes of dynamical systems: ordinary differential equations and inclusions, delay-differential and neutral-type inclusions, partial differential equations and inclusions of the parabolic type, etc.

In what follows we discuss the current stage of the method of discrete approximations married to appropriate tools of variational analysis to study optimal control problems for differential inclusions of type (1.2), where set-valued mappings F are *Lipschitz continuous* with respect to state variables. In this way we derive necessary optimality conditions in the extended Euler-Lagrange form accompanied by the Weierstrass-Pontryagin maximum condition. The second part of the paper [13] deals with new classes of control systems governed by highly *discontinuous* inclusions with a controlled sweeping process dynamics.

The rest of the paper is organized as follows. In Section 2 we present and discuss some basic robust constructions of *generalized differentiation* in variational analysis that are appropriate to study differential inclusions while being widely used in all the subsequent sections.

Section 3 deals with *discrete approximations* of Lipschitzian differential inclusions (1.2), without considering their optimization so far. The main result here shows the possibility of an appropriate *strong approximation* of *any* feasible trajectory of (1.2) by feasible trajectories of discrete-time systems that are piecewise linearly extended to the continuous-time interval ("Euler broken lines"). The developed approximation procedure can be viewed as a *numerical scheme* of finite-dimensional approximations of infinite-dimensional problems.

In Section 4 we formulate the Bolza-type optimization problem for differential inclusions under consideration and construct a sequence of its finite-dimensional approximations by problems with discrete time. The main result here verifies the *well-posedness* of such a discrete approximation in the sense of *strong* $W^{1,2}$*-convergence* of optimal solutions to discrete-time problems to a given local minimizer of the original continuous-time Bolza problem under the local Lipschitz continuity of the velocity map.

The final Section 5 is devoted to deriving *necessary optimality conditions* for Bolza-type optimization problems governed by Lipschitzian differential inclusions. Our approach is based on the method of discrete approximations with the usage of the convergence results from Sections (3) and (4) together with the tools of generalized differentiation in variational analysis discussed in Section 2. In this way we construct a well–posed family of discrete-time optimization problems, which strongly converge to the given local minimizer of the original continuous-time problem, obtain necessary optimality conditions for discrete approximations, and then establish optimality conditions for differential inclusions of the *extended Euler-Lagrange* type accompanied by the *Weierstrass-Pontryagin maximization condition.*

Throughout the paper we use the standard notation of variational analysis, generalized differentiation and control theory; see, e.g., [11; 15; 18]. We specified them in the places where they appear for the first time in the paper. Among other symbols, recall that A^* signifies for the transposed/adjoint matrix to A and that $\mathbb{N} := \{1, 2, \ldots\}$. We also mention that $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ indicates that F may be a set-valued mapping, in contrast to the usual notation $F\colon \mathbb{R}^n \to \mathbb{R}^m$ for single-valued ones.

2. Tools of Generalized Differentiation

In this section we provide a brief overview of those constructions of generalized differentiation for nonsmooth functions, nonconvex sets, and set-valued mappings that are used in the paper. These constructions have been initiated by the author in [5] while now being major in variational analysis and its applications to optimization, control theory, and numerous applications; see, e.g., the books [6; 9–11; 15; 18] and the references therein for more details.

We start with extended-real-valued functions $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}} := (-\infty, \infty]$, which is a standard and convenient framework in convex and variational analysis. Given $\bar{x} \in \operatorname{dom}\varphi := \{x \in \mathbb{R}^m |\ \varphi(x) < \infty\}$, the (first-order) *subdifferential* (or the set of subgradients) of φ at $\bar{x}$ is defined by

$$\partial\varphi(\bar{x}) := \Big\{ v \in \mathbb{R}^n \Big|\ \exists\, x_k \to \bar{x}, \exists\, v_k \to v \text{ with } \varphi(x_k) \to \varphi(\bar{x}) \text{ and } \liminf_{x \to x_k} \frac{\varphi(x) - \varphi(x_k) - \langle v_k, x - x_k\rangle}{\|x - x_k\|} \geq 0 \Big\}, \tag{2.1}$$

where $k \to \infty$. This construction reduces to the gradient $\partial\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$ for smooth functions and to the classical subdifferential of convex analysis if φ is convex. The subgradient set (2.1) is nonempty for any function φ that is locally Lipschitzian around $\bar{x}$ while may be nonconvex even for simple Lipschitzian functions; e.g., $\partial\varphi(0) = \{-1, 1\}$ for $\varphi(x) := -|x|$ on $\mathbb{R}$. Nevertheless the subdifferential (2.1) and associated constructions for sets and set-valued mappings enjoy comprehensive calculus rules, which are based on *variational/extremal principles* of variational analysis.

Given a set $\Omega \subset \mathbb{R}^n$, consider its indicator function $\delta_\Omega(x)$, which equals 0 for $x \in \Omega$ and ∞ otherwise, and define the *normal cone* to Ω at $\bar{x}$ by

$$N_\Omega(\bar{x}) = N(\bar{x}; \Omega) := \partial\delta_\Omega(\bar{x}) \text{ for } x \in \Omega,\ \ N(\bar{x}; \Omega) := \emptyset \ \text{ for } \ \bar{x} \notin \Omega. \tag{2.2}$$

Considering then a set-valued mapping/multifunction $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with the domain $\operatorname{dom} F := \{x \in \mathbb{R}^n |\ F(x) \neq \emptyset\}$ and the graph $\operatorname{gph} F := \{(x, y) \in$

$\mathbb{R}^n \times \mathbb{R}^m |\ y \in F(x))\}$, the *coderivative* of F at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ is defined by

$$D^*F(\bar{x}, \bar{y})(u) := \big\{v \in \mathbb{R}^n \big|\ (v, -u) \in N\big((\bar{x}, \bar{y}); \operatorname{gph} F\big)\big\},\ u \in \mathbb{R}^m, \quad (2.3)$$

while we drop $\bar{y}$ in (2.3) when F is single-valued. In the case where F is smooth (C^1) around $\bar{x}$ we have

$$D^*F(\bar{x})(u) = \big\{\nabla F(\bar{x})^* u\big\} \text{ for all } u \in \mathbb{R}^m$$

via the transpose Jacobian matrix, but in general the coderivative (2.3) is a positively homogeneous set-valued mapping enjoying full calculus rules and providing complete characterizations (called "Mordukhovich criteria" in [15]) of the major well-posedness properties in nonlinear analysis related to Lipschitzian stability, metric regularity, and linear openness/covering of multifunctions; see [7] and then [9;11;15] for different proofs and numerous applications. Let us present the corresponding characterization of the local Lipschitzian property of set-valued (and single-valued) mappings taken from [7, Theorem 5.11]. Recall that a multifunction $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *locally Lipschitzian* around ($\bar{x} \in \operatorname{dom} F$ is there exist a constant $\kappa \geq 0$ and a neighborhood U of $\bar{x}$ such that

$$F(x) \subset F(u) + \kappa\|x - u\|\mathbb{B} \text{ for all } x, u \in U,$$

where $\mathbb{B}$ stands for the closed unit ball of the space in question.

Theorem 1. *Let F be locally bounded around $\bar{x}$. Then it is locally Lipschitzian around this point if and only if we have*

$$D^*F(\bar{x}, \bar{y})(0) = \{0\} \text{ for all } \bar{y} \in F(\bar{x}).$$

This criterion plays a crucial role in deriving necessary optimality conditions for Lipschitzian differential inclusions in Section 4.

3. Discrete Approximations of Differential Inclusions

This section concerns discrete approximations of continuous-time dynamical systems governed by differential inclusions of type (1.2). For simplicity we consider the case of *autonomous differential inclusions*

$$\dot{x}(t) \in F\big(x(t)\big) \text{ a.e. } t \in [0, T] \quad (3.1)$$

in the class of absolutely continuous trajectories $x\colon [0, T] \to \mathbb{R}^n$, where the terminal time $T > 0$ is fixed. Our goal is the show that any feasible trajectory of (3.1) can be strong approximate in the norm of the classical Sobolev space $W^{1,2}([0, T]; \mathbb{R}^n)$ by feasible trajectories of discrete-time systems that are piecewise linearly extended to the continuous time interval $[0, T]$. Note that the $W^{1,2}([0,T];\mathbb{R}^n)$-*norm convergence* of a functional sequence yields the uniform convergence of the functions on $[0,T]$ and the a.e. pointwise convergence of their derivatives along some subsequence.

In what follows we confine ourselves to the *uniform Euler scheme* for the finite-difference replacement of the derivative

$$\dot{x}(t) \approx \frac{x(t+h)-x(t)}{h}, \quad h \downarrow 0,$$

while it does not actually restrict the generality. The discrete approximation process for (3.1) is formalized as follows. Take any natural number $k \in \mathbb{N}$ and form the *discrete mesh/grid*

$$T_k := \{0, h_k, \ldots, T-h_k, T\} \text{ with } h_k := T/k.$$

Denoting the mesh points $t+jh_k$ as $j=0,\ldots,k$ with $t_0=0$ and $t_k=T$, construct the family of *discrete-time inclusions* by

$$x_k(t_{j+1}) \in x_k(t_j) + h_k F\big(x_k(t_j)\big), \quad j=0,\ldots,k-1. \tag{3.2}$$

Given a feasible trajectory $\bar{x}(t)$ to (3.1), assume from now on that:

(H1) F is locally bounded and locally Lipschitzian around $\bar{x}(t)$ uniformly in t on the entire interval $[0,T]$.

The following result provides the desired discrete approximation of *any feasible trajectory* to the differential inclusion (3.1).

Theorem 2. *Let $\bar{x}(t)$ be a feasible trajectory to the differential inclusion (3.1), and let the assumptions in* (H1) *be satisfied around $\bar{x}(t)$. Then there exists a sequence of discrete trajectories $z_k(t_j)$, $t_j \in T_k$, to (3.2) with $z_k(t_0)=\bar{x}(0)$ such that their piecewise linear extensions $z_k(t)$, $t \in [0,T]$, converge to $\bar{x}(t)$ in the $W^{1,2}([0,T];\mathbb{R}^n)$-norm as $k \to \infty$.*

Sketch of Proof. We first approximate $\dot{\bar{x}}(t)$ in the strong topology of $L^1([0,T];$
$\mathbb{R}^n$ by a sequence of step functions $w_k(t)$, which are constant on $[t_j,t_{j+1})$ for all $j=0,\ldots,k-1$. Then we construct $z_k(t_j)$ recurrently by the following *proximal algorithm*:

$$\begin{array}{l} z_k(t_0)=\bar{x}(0),\ z_k(t_{j+1})=z_k(t_j)+h_k v_k(t_j), \\ \text{where } \ v_k \in F\big(z_k(t_j)\big) \ \text{ with} \\ \|v_k(t_j)-w_k(t_j)\| = \operatorname{dist}\big(w_k(t_j);F\big(z_k(t_j)\big),\ j=0,\ldots,k-1 \end{array}$$

The imposed local Lipschitz continuity of $F(\cdot)$ allows us to verify the claimed strongly approximation with efficient numerical estimates.

We refer the reader to [10, Theorem 6.4] for a detailed proof of the approximation result under more general assumptions for nonautonomous differential/evolution inclusions with the right-hand side $F\colon [0,T]\times X \rightrightarrows X$ defined on an arbitrary Banach space X. However, the crucial *Lipschitz continuity* property of F with respect to the state variable is essentially used in the given proof. Observe also that the given proof is constructive with establishing efficient error bounds and estimates, and hence Theorem 2 is of a certain *numerical value*. Nevertheless, the main goal of the developed approach is utilizing Theorem 2 to derive necessary optimality conditions in optimal control problems for differential inclusions by using the method of discrete approximations married to the powerful machinery of variational analysis and generalized differentiation.

4. Strong Convergence of Discrete Optimal Solutions

In this section we consider the following Bolza-type *optimal control problem* (P) for differential inclusions:

$$\text{minimize } \; J[x] := \varphi\big(x(0), x(T)\big) + \int_0^T \ell\big(x(t), \dot{x}(t)\big)dt \tag{4.1}$$

over absolutely continuous trajectories $x\colon [0,T] \to \mathbb{R}^n$ of the autonomous differential inclusion (3.1) subject to the geometric endpoint constraints

$$\big(x(0), x(T)\big) \in \Omega. \tag{4.2}$$

In [1; 10; 18] the reader can find more general versions of this problem for nonautonomous differential inclusions without any convexity assumptions on $F(x)$ and $\ell(x,\cdot)$ and Lipschitzian assumptions on the terminal and running costs φ and ℓ, respectively. We choose here the model in (3.1), (4.1), (4.2) for simplicity and better comparison with controlled sweeping processes considered in Part II [13]. The crucial assumption in the necessary optimality conditions of Theorem 4 is the *Lipschitzian* dependence of the set-valued mapping F on the state variable x.

As mentioned in Section 1, our approach to deriving necessary optimality conditions for local minimizers of the above problem (P) is based on the method of discrete approximations and appropriate machinery of variational analysis. The main issues of this approach are as follows:

- Construct a family of discrete approximations of the differential inclusion (3.1) involving a finite-difference replacement of the derivative in (3.1) and a consistent perturbation of the endpoint constraints in (4.2). Then approximate any feasible trajectory of (3.1) by feasible trajectories of discrete systems in a topology implying the a.e. convergence of the discrete derivatives that are piecewise constantly extended on the continuous-time interval $[0,T]$. In this step (see Section 3) we address not only qualitative aspects of well-posedness but also numerical ones with estimating error bounds, convergence rates, etc. Achieving it leads us to the $W^{1,2}$-norm approximation of a given local minimizer for the continuous-time problem (P) by a sequence of optimal solutions to the discrete-time problems that are piecewise linearly extended to the entire interval $[0,T]$. In [10] it was done for a class of the so-called "intermediate local minimizers" introduced in [8]. This class includes strong local minimizers while occupying an intermediate position between the latter and weak local minimizers in dynamic optimization; see Definition 1 and subsequent discussions.

• Each discrete-time problem that approximates the original one can be reduced to a nondynamic problem of mathematical programming in finite dimensions with increasingly many geometric constraints of the graphical type. We employ the powerful tools of generalized differentiation discussed in Section 2 for deriving necessary optimality conditions in the approximating discrete-time problems. It can be done without any Lipschitzian and convexity assumptions by applying the well-developed generalized differential calculus for them. Note that dealing with the graphical structure of the geometric constraints requires that the used generalized differential constructions should be subtle and small enough to handle graphical sets. In particular, the convexified normal cone by Clarke [1] cannot be employed for these purposes since applying it to graphical sets often gives us the whole space or its subspace of maximal dimension; see [9; 11; 15] for more details. On the other hand, our constructions discussed in Section 2 satisfy all the required properties and thus can be successfully implemented.

• Finally, we derive necessary optimality conditions for local minimizers of (P) by passing to the limit from those for discrete approximations. This part is the most challenging while requiring the clarification and justification of an appropriate convergence of dual arcs. For the case of Lipschitzian differential inclusions it is done by using the coderivative criterion from Theorem 1 for the Lipschitz continuity of set-valued mappings.

The necessary optimality conditions formulated below concerns the following notion of local minimizers for optimization problems of type (P) from (3.1), (4.1), (4.2) that first appeared in [8].

Definition 1. *Let $\bar{x}(\cdot)$ be a feasible solution to problem (P). We say that $\bar{x}(\cdot)$ is an* INTERMEDIATE LOCAL MINIMIZER *of rank $p \in [1,\infty)$ for this problem if there are $\varepsilon > 0$ and $\alpha \geq 0$ such that $J[\bar{x}] \leq J[x]$ for any feasible solution to (P) satisfying the localizing constraints*

$$\|x(t) - \bar{x}(t)\| < \varepsilon \ \textit{ for all } \ t \in [0,T] \ \textit{ and } \ \alpha \int_0^T \|\dot{x}(t) - \dot{\bar{x}}(t)\|^p dt < \varepsilon. \quad (4.3)$$

The localization in (4.3) means in fact that a neighborhood of $\bar{x}(\cdot)$ in the space $W^{1,p}\big([0,T];\mathbb{R}^n\big)$ is considered. If $\alpha = 0$ in (4.3), we get the classical *strong* local minimum corresponding to a neighborhood of $\bar{x}$ in the norm topology of $C([0,T];\mathbb{R}^n)$. If (4.3) is replaced by

$$\|\dot{x}(t)-\dot{\bar{x}}(t)\|<\varepsilon \ \text{ a.e. } \ t\in[0,T],$$

we get the classical *weak* local minimum in the framework of Definition 1, which corresponds to considering a neighborhood of $\bar{x}(\cdot)$ in the norm topology of $W^{1,\infty}([0,T];\mathbb{R}^n)$. The reader is referred to [8;18] for various examples showing that the intermediate notion of Definition 1 is strictly different from both strong and weak local minimizers of (P) even for convex autonomous differential inclusions considered in this paper for simplicity. More precisely, given a reference trajectory $\bar{x}(t)$, we suppose that:

(H2) The velocity mapping F in (3.1) is convex-valued and the running cost ℓ in (4.1) is convex with respect top the velocity variable $\dot{x}$.

Note that the convexity assumptions in $(H2)$ can be *dismissed* by using a *relaxation procedure* of the Bogolyubov-Young type. The reader can find more details on such relaxations in [1; 2; 10; 16; 18] and the references therein. This approach leads us to study the so-called *relaxed intermediate local minimizers* of (P) as in [10, Definition 6.7]. Observe also that the assumptions imposed in the results below ensure that we can consider the case of $p=2$ without loss of generality and thus refer to $\bar{x}(\cdot)$ as to an intermediate local minimizer for problem (P).

We proceed further with the construction of the family of discrete approximation problems to derive necessary optimality conditions for the *given* intermediate local minimizer $\bar{x}(t)$ of the original problem (P). Note that the family (sequence) of discrete-time problems (P_k), $k\in\mathbb{N}$, constructed below explicitly involves the given minimizer under consideration, while we do *not* employ *any variations* of it as in the conventional methods of the calculus of variations and optimal control. Given $\bar{x}(t)$, for each $k\in\mathbb{N}$ problem (P_k) is defined as follows: minimize

$$\begin{aligned}J_k[x_k]:=\varphi\big(x_k(t_0),x_k(t_k)\big)+\|x_k(t_0)-\bar{x}(0)\|^2+\\ +\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}}\ell\Big(x_k(t_j),\frac{x_k(t_{j+1})-x_k(t_j)}{h_k}\Big)dt+\\ +\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}}\Big\|\frac{x_k(t_{j+1})-x_k(t_j)}{h_k}-\dot{\bar{x}}(t)\Big\|^2dt \quad (4.4)\end{aligned}$$

over discrete trajectories $x_k=x_k(\cdot)=(x_k(t_0),\ldots,x_k(t_k))$ for the difference inclusions (3.2) subjects to the constraints

$$\big(x_k(t_0),x_k(t_k)\big)\in\Omega+\gamma_k\mathbb{B}, \quad (4.5)$$

$$\|x_k(t_j)-\bar{x}(t_j)\|\le\frac{\varepsilon}{2} \ \text{ for all } \ j=1,\ldots,k, \quad \text{and} \quad (4.6)$$

$$\sum_{j=0}^{k-1}\int_{t_j}^{t_{j+1}}\left\|\frac{x_k(t_{j+1})-x_k(t_j)}{h_k}-\dot{\bar{x}}(t)\right\|^2 dt\le\frac{\varepsilon}{2}, \tag{4.7}$$

where ε is taken from Definition 1 of the intermediate local minimizer $\bar{x}(t)$, and where $\gamma_k := \|z_k(t_k)-\bar{x}_k(T)\|$ for the sequence $\{z_k(t)\}$ that approximates $\bar{x}(t)$ by Theorem 2.

The next theorem establishes the *well-posedness* of the method of discrete approximations concerning the existence of optimal solutions $\bar{x}_k(\cdot)$ to each problem (P_k) and the strong $W^{1,2}([0,T];\mathbb{R}^n)$ convergence of the extended sequence of the discrete trajectories $\bar{x}_k(t)$, $t\in[0,T]$, to the given intermediate local minimizer $\bar{x}(t)$ to (P).

Theorem 3. *Let $\bar{x}(\cdot)$ be an intermediate local minimizer for problem (P) under the assumptions in* (H1) *and* (H2)*, and let (P_k), $k\in\mathbb{N}$ be a sequence of the discrete-time optimization problems constructed in (3.2), (4.4)–(4.7). Suppose in addition that the cost functions φ and ℓ in (4.1) are continuous around $\bar{x}(\cdot)$. Then the following assertions hold:*

(i) *Each problem (P_k), when $k\in\mathbb{N}$ is sufficiently large, admits an optimal solution $\bar{x}_k(\cdot)$.*

(ii) *Any sequence $\{\bar{x}_k(t)\}$ of optimal solutions to (P_k) piecewise linearly extended to the continuous-time interval $[0,T]$ converges to $\bar{x}(t)$ in the norm topology of $W^{1,2}([0,T];\mathbb{R}^n)$.*

Sketch of Proof. Observe first that the set of feasible solutions to each (P_k) with sufficiently large $k\in\mathbb{N}$ is nonempty. It follows from Theorem 3, Definition 1, and the construction of (P_k). Furthermore, we deduce from the constraints in (4.6) that the feasible set in (P_k) is bounded and hence compact due to its obvious closedness. Thus the existence assertion (i) holds by to the classical Weierstrass theorem.

To verify (ii), it is sufficient to show that

$$J_k[z_k]\to J[\bar{x}]\ \text{ as }\ k\to\infty, \tag{4.8}$$

where the approximating discrete trajectories $z_k(\cdot)$ are constructed in Theorem 3 for the given local minimizer $\bar{x}(\cdot)$ of (P). Assuming the contrary to (4.8) and using the compactness of the solution set to differential inclusions under the assumptions made (as follows from the result by Tolstonogov [16, Theorem 3.4.2]) allows us to find an absolutely continuous mapping $\tilde{x}\colon[0,T]\to\mathbb{R}^n$ such that $\bar{x}_k(\cdot)\to\tilde{x}(\cdot)$ uniformly on $[0,T]$ along a subsequence as $k\to\infty$, which corresponds to the weak convergence of the derivatives. Then the classical Mazur's theorem on weak closure and the imposed convexity of the sets $F(x)$ and the mapping $v\mapsto\ell(x,v)$ ensures that

$\tilde{x}(\cdot)$ is a feasible trajectory of (3.1). Using finally the assumed continuity of the functions φ and ℓ brings us to a contradiction.

We refer the reader to [10, Theorem 6.13] to a nonconvex and nonautonomous version of Theorem 3 for optimization problems of type (P) with infinite-dimensional state spaces.

5. Necessary Optimality Conditions

Now we are ready to present the major theorem on necessary optimality conditions for intermediate local minimizers (and hence for strong local minimizers as well) of the continuous-time problem (P) obtained by using the discrete approximation method together with the constructions and results of variational analysis and generalized differentiation.

Theorem 4. *Let $\bar{x}(\cdot)$ be an intermediate local minimizer for problem (P) under the assumptions in* (H1) *and* (H2). *Suppose in addition that the running cost ℓ and the terminal cost φ are locally Lipschitzian around $\bar{x}(\cdot)$ and that the constraint set Ω is locally closed around the endpoint $(\bar{x}(0), \bar{x}(T))$. Then there exist a number $\lambda \geq 0$ and an absolutely continuous adjoint arc $p\colon [0,T] \to \mathbb{R}^n$, not equal to zero simultaneously, satisfying for a.e. $t \in [0,T]$ the extended Euler-Lagrange inclusion*

$$\dot{p}(t) \in \operatorname{co}\Big\{u \in \mathbb{R}^n \Big|\ \big(u, p(t)\big) \in \lambda\partial\ell\big(\bar{x}(t), \dot{\bar{x}}(t)\big) + N\big((\bar{x}(t), \dot{\bar{x}}(t)); \operatorname{gph} F\big)\Big\},$$

the Weierstrass-Pontryagin maximum condition

$$\big\langle p(t), \dot{\bar{x}}(t)\big\rangle - \lambda\ell\big(\bar{x}(t), \dot{\bar{x}}(t)\big) = \max_{v \in F(\bar{x}(t))} \Big\{\big\langle p(t), v\big\rangle - \lambda\ell\big(\bar{x}(t), v\big)\Big\},$$

and also the transversality inclusion at both endpoints

$$\big(p(0), -p(T)\big) \in \lambda\partial\varphi\big(\bar{x}(0), \bar{x}(T)\big) + N\big((\bar{x}(0), \bar{x}(T)); \Omega\big).$$

Sketch of Proof. Having the well-posedness results of Theorem 3, we concentrate further on deriving necessary optimality conditions for problems (P_k), $k \in \mathbb{N}$. Note that each problem (P_k) with a fixed number $k \in \mathbb{N}$ is a discrete-time problem of dynamic optimization. Clearly it can be written in a nondynamic form of finite-dimensional optimization with specific types of constraints. In particular, the dynamic constraints in (3.2) are reduced to increasingly many *geometric constraints* expressed via the *graph* of the velocity mapping F from (1.2). Such graphical sets are intrinsically *nonconvex* and often have nonempty interiors. Nevertheless, the developed in [9] generalized differential calculus for the robust normal, subdifferential, and coderivative constructions reviewed in Section 2 allows us to derive necessary optimality conditions for the discrete-time problems (P_k).

To proceed, we first formulate for the reader's convenience the discrete-time problems under consideration in the simplified form: minimize

$$\begin{aligned} &\varphi(x_0,x_k)+h\sum_{j=0}^{k}\ell\Big(x_j,\frac{x_{j+1}-x_j}{h}\Big)\ \text{ subject to} \\ &x_{j+1}\in x_j+hF(x_j)\ \text{ for }\ j=0,\ldots,k-1,(x_0,x_k)\in\Omega. \end{aligned} \tag{5.1}$$

Then the aforementioned generalized differential calculus ensures that, without any convexity assumptions and Lipschitz continuity of F, for a given optimal solution $\{\bar{x}_j|\ j=0,\ldots,k\}$ to (5.1) there exist dual elements $\lambda\geq 0$ and $p_j\in\mathbb{R}^n|\ j=0,\ldots,k\}$, not equal to zero simultaneously, such that we have the *discrete Euler-Lagrange inclusions*

$$\Big(\frac{p_{j+1}-p_j}{h},p_{j+1}\Big)\in\lambda\partial\ell_j\Big(\bar{x}_j,\frac{\bar{x}_{j+1}-\bar{x}_j}{h}\Big)+N\Big(\Big(\frac{\bar{x}_{j+1}-\bar{x}_j}{h}\Big);\operatorname{gph}F_j\Big)$$

whenever $j=0,\ldots,k-1$ with the *transversality inclusion*

$$(p_0,-p_k)\in\lambda\partial\varphi(\bar{x}_0,\bar{x}_k)+N\big((\bar{x}_0,\bar{x}_k);\Omega\big).$$

These conditions can be clearly adjusted to the precise form of problem (P_k) for each $k\in\mathbb{N}$. The main issue now is to pass to the limit therein as $k\to\infty$. The usage of Theorem 3 on the *primal convergence* of optimal trajectories for problems (P_k) to the given local minimizer $\bar{x}(\cdot)$ of (P) is surely necessary, while not being sufficient to furnish this limiting procedure. Another important ingredient in the implementation of this approach is *robustness* of the generalized differential constructions from Section 2 with respect to perturbations of the initial data. However, the availability of both aforementioned results says nothing about an appropriate *dual convergence* of adjoint arcs that appear in the discrete Euler-Lagrange inclusions. Such a dual convergence in the strong topology of the space $\mathcal{C}([0,T];\mathbb{R}^n)$ is achieved due to the *coderivative criterion* for local Lipschitz continuity presented in Theorem 1. In this way we arrive at all the necessary optimality conditions of this theorem under the assumptions made.

We refer the reader to [10; 12; 13] for further developments in this direction and their various applications.

References

1. Clarke F.H. *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
2. Donchev T., Farkhi E., Mordukhovich B.S., Discrete approximations, relaxation, and optimization of one-sided Lipschitzian differential inclusions in Hilbert spaces. *J. Diff. Eqs*, 2007, vol. 243, pp. 301-328.

3. Euler L. *Methodus Inveniendi Curvas Lineas Maximi Minimive Proprietate Gaudentes Sive Solution Problematis Viso Isoperimetrici Latissimo Sensu Accepti*, Lausanne, 1774; reprinted in Opera Omnia, 1952, Ser. 1, vol. 24.
4. Gabasov R., Kirillova F.M. *The Qualitative Theory of Optimal Processes.* Marcel Dekker, New York, 1976.
5. Mordukhovich B.S. Maximum principle in problems of time optimal control with nonsmooth constraints. *J. Appl. Math. Mech.*, 1976, vol. 40, pp. 960-969.
6. Mordukhovich B.S. *Approximation Methods in Problems of Optimization and Control.* Moscow, Nauka Publ.,1988.
7. Mordukhovich B.S. Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions. *Trans. Amer. Math. Soc.*, 1993, vol. 340, pp. 1-35.
8. Mordukhovich B.S. Discrete approximations and refined Euler-Lagrange conditions for differential inclusions. *SIAM J. Control Optim.*, 1995, vol. 33, pp. 882-915.
9. Mordukhovich B.S. *Variational Analysis and Generalized Differentiation, I: Basic Theory.* Springer, Berlin, 2006.
10. Mordukhovich B.S. *Variational Analysis and Generalized Differentiation, II: Applications.* Springer, Berlin, 2006.
11. Mordukhovich B.S. *Variational Analysis and Applications.* Springer, Switzerland, 2018.
12. Mordukhovich B.S. Optimal control of Lipschitzian and discontinuous differential inclusions with various applications. *Proc. Inst. Math. Mech.*, 2019, vol. 45, pp. 52-74.
13. Mordukhovich B.S. Optimal control of differential inclusions, II: Sweeping. *The Bulletin of Irkutsk State University. Series Mathematics*, 2020, vol. 31.
14. Pontryagin L.S., Boltyanskii V.G., Gamkrelidze R. V., Mishchenko E.F. *The Mathematical Theory of Optimal Processes.* Wiley, New York, 1962.
15. Rockafellar R.T., Wets R.J-B. *Variational Analysis.* Springer, Berlin, 1998.
16. Tolstonogov A.A. *Differential Inclusions in a Banach Spaces.* Kluwer, Dordrecht, the Netherlands, 2000.
17. Vasiliev O.V. *Optimization Methods.* World Federation Publishers, Atlanta, GA, 1996.
18. Vinter R.B. *Optimal Control.* Birkhaüser, Boston, 2000.

Boris Mordukhovich, Ph. D. in Applied Mathematics, Distinguished University Professor, Department of Mathematics, Wayne State University, Detroit, Michigan, 48202, USA, e-mail: `boris@math.wayne.edu`,
ORCID iD https://orcid.org/0000-0002-6071-6049

Received 05.10.19

Оптимальное управление дифференциальными включениями, I: Липшицевы дифференциальные включения

Б. Ш. Мордухович

Государственный университет Уэйна, Детройт, США

Аннотация. Разработан метод дискретных аппроксимаций для изучения задач оптимального управления дифференциальными включениями с использованием современных инструментов вариационного анализа и обобщенного дифференцирования. Первая часть описывает метод, соответствующий механизм вариационного анализа, а затем представляет основной результат о необходимых условиях оптимальности в форме принципа максимума для липшицевых дифференциальных включений.

Ключевые слова: оптимальное управление, Липшицевы дифференциальные включения, вариационный анализ, дискретные аппроксимации, обобщенное дифференцирование.

Борис Шолимович Мордухович, доктор физико-математических наук, профессор, математический факультет, Государственный университет Уэйна, США, 48202, Мичиган, Детройт, e-mail: boris@math.wayne.edu,
ORCID iD https://orcid.org/0000-0002-6071-6049

Поступила в редакцию 05.10.19