Wayne State University

Mathematics Research Reports

Mathematics

9-1-2005

Discrete Approximations of Differential Inclusions in Infinite-Dimensional Spaces

Boris S. Mordukhovich
Wayne State University, boris@math.wayne.edu

DISCRETE APPROXIMATIONS OF DIFFERENTIAL INCLUSIONS IN INFINITE-DIMENSIONAL SPACES

B. S. MORDUKHOVICH

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2005 Series
#9

DISCRETE APPROXIMATIONS OF DIFFERENTIAL INCLUSIONS IN INFINITE-DIMENSIONAL SPACES

BORIS S. MORDUKHOVICH [1]

Department of Mathematics, Wayne State University, Detroit, Michigan 48202
boris@math.wayne.edu

Abstract: In this paper we study discrete approximations of continuous-time evolution systems governed by differential inclusions with nonconvex compact values in infinite-dimensional spaces. Our crucial result ensures the possibility of a strong Sobolev space approximation of every feasible solution to the continuous-time inclusion by its discrete-time counterparts extended as Euler's "broken lines." This result allows us to establish the value and strong solution convergences of discrete approximations of the Bolza problem for constrained infinite-dimensional differential/evolution inclusions under natural assumptions on the initial data.

Keywords: differential inclusions, infinite dimension, discrete approximations, optimal control, Bolza problem, relaxation stability, value and strong solution convergences.

Mathematics Subject Classifications (2000): 49J52, 49M25, 90C30.

1   Introduction

This paper is devoted to discrete/finite difference approximations of evolution systems with the continuous-time dynamics governed by differential inclusions in Banach spaces. It has been well recognized that discrete approximations of continuous-time systems (which go back to Leibniz and Euler in the classical calculus of variations) play a significant role in the study and applications of variational problems. Nowadays difference methods are mostly investigated and employed from the viewpoint of numerical analysis in order to approximate and compute solutions to continuous-time systems; see, e.g., the excellent survey by Dontchev and Lempio [5] devoted to these aspects of discrete approximations for differential inclusions in finite-dimensional spaces. On the other hand, discrete approximations can be considered as an efficient tool to derive qualitative results for continuous-time systems by reducing them, in a sense, to discrete-time systems and subsequently to the corresponding non-dynamic problems. The latter viewpoint was taken by Mordukhovich [9] in his study of necessary optimality conditions for optimal control problems governed by differential inclusions in finite dimensions. Then this approach was developed in many publications for various problems concerning optimal control of ordinary differential equations and inclusions, time-delay and functional differential systems, differential-algebraic systems, etc.; see the book [11] and the references therein. In the majority of publications in this direction, the method of discrete approximations was used for deriving new necessary optimality conditions for continuous-time systems based on their reduction to special problems of constrained mathematical programming with the usage of appropriate tools of variational analysis and generalized differentiation [10].

In this paper, we draw the main attention to the *well-posedness* of discrete approximations for

---

evolution systems governed by differential inclusions in *infinite-dimensional* spaces. By this we mean establishing results of the following three kinds:

(a) *strong approximation* (in the appropriate Sobolev space) of *any feasible* solution to the differential inclusion by solutions to its discrete-time counterparts extended as Euler's broken lines;

(b) constructing discrete approximations of the *generalized Bolza problem* for differential inclusions in such a way that *optimal values* of the cost functionals in the corresponding discrete problems converge to the optimal value in the original continuous-time problem;

(c) constructing discrete approximations of the generalized Bolza problem for differential inclusions in such a way that *optimal solutions* to corresponding discrete problems *strongly converge* to the given *local* optimal solution for the original problem.

In what follows we obtain general results of the types (a)–(c) under natural assumptions on the initial data. The notation used is standard in variational analysis; see, e.g., [10, 11, 14].

2 Differential Inclusions and Their Discrete Approximations

Let X be a Banach space (called the *state space* in what follows), and let $T := [a, b]$ be a *time interval* of the real line. Consider a set-valued mapping $F: X \times T \rightrightarrows X$ and define the *differential* or *evolution inclusion*

$$\dot{x}(t) \in F(x(t), t) \quad \text{a.e.} \quad t \in [a, b] \tag{2.1}$$

generated by F, where $\dot{x}(t)$ stands for the time derivative of $x(t)$, and where a.e. (almost everywhere) means as usual that the relation holds up to the Lebesgue measure zero on $I\!\!R$. Let us give the precise definition of solutions to the differential inclusion (2.1), which is used in this paper.

Definition 2.1 (solutions to differential inclusions). *By a* SOLUTION *to inclusion* (2.1) *we understand a mapping* $x: T \to X$, *which is Fréchet differentiable for a.e.* $t \in T$, *satisfies* (2.1) *and the* NEWTON-LEIBNIZ FORMULA

$$x(t) = x(a) + \int_a^t \dot{x}(s)\, ds \quad \text{for all} \quad t \in T,$$

where the integral in taken in the BOCHNER SENSE.

It is well known that for $X = I\!\!R^n$, $x(t)$ is a.e. differentiable on T and satisfies the Newton-Leibniz formula if and only if it is *absolutely continuous* on T in the standard sense, i.e., for any $\varepsilon > 0$ there is δ such that

$$\sum_{j=1}^{l} \|x(t_{j+1}) - x(t_j)\| \leq \varepsilon \quad \text{whenever} \quad \sum_{j=1}^{l} |t_{j+1} - t_j| \leq \delta$$

for the disjoint intervals $(t_j, t_{j+1}] \subset T$. However, for infinite-dimensional spaces X even the Lipschitz continuity may not imply the a.e. differentiability. On the other hand, there is a *complete characterization* of Banach spaces X, where the absolute continuity of every $x: T \to X$ is *equivalent* to its a.e. differentiability and the fulfillment of the Newton-Leibniz formula. This is the class of spaces with the so-called *Radon-Nikodým property*.

2

Definition 2.2 (Radon-Nikodým property). *A Banach space X has the* RADON-NIKODÝM PROPERTY *if for every finite measure space (Ξ, Σ, μ) and for each μ-continuous vector measure $m \colon \Sigma \to X$ of bounded variation there is $g \in L^1(\mu; \Xi)$ such that*

$$m(E) = \int_E g \, d\mu \quad for \ E \in \Sigma.$$

This fundamental property is well investigated in the general vector measure theory and in the geometric theory of Banach spaces; we refer the reader to the classical text by Diestel and Uhl [3] for the comprehensive study of the RNP and its applications. In particular, in [3, pp. 217–219] one can find the summary of equivalent formulations/charactetizations of the RNP and the list of specific Banach spaces for which the RNP automatically holds. It is important to observe that the latter list contains every *reflexive* space and every *weakly compactly generated dual* space, hence all *separable duals*. On the other hand, the classical spaces c_0, c, l^∞, $L^1[0,1]$, and $L^\infty[0,1]$ *don't* have the RNP. Let us mention a nice relationship between the RNP and Asplund spaces used in what follows: *given a Banach space X, the dual space X^* has the RNP if and only if X is Asplund.* Recall that a Banach space is *Asplund* if its every separable subspace has a separable dual; the reader can find more details, equivalent descriptions, and various implementations of the Asplund property in the book by Phelps [12] and the references therein.

Thus for Banach spaces with the RNP (and only for such spaces) the solution concept of Definition 2.1 agrees with the standard definition of *Carathéodory solutions* dealing with absolutely continuous mappings. In general Definition 2.1 postulates what we actually need for our purposes without appealing to Carathéodory solutions and the RNP. However, the RNP along with the Asplund property of X are essentially used for deriving some important results of this paper (but not all of them) from somewhat different prospectives not directly related to the adopted concept of solutions to differential inclusions in infinite-dimensional spaces.

It has been well recognized that differential inclusions, which are certainly of their own interest, provide a useful generalization of *control systems* governed by differential/evolution *equations* with control parameters:

$$\dot{x} = f(x, u, t), \quad u \in U(t), \tag{2.2}$$

where the control sets $U(\cdot)$ may also depend on the state variable x via $F(x,t) = f(x, U(x,t), t)$. In some cases, especially when the sets $F(x,t)$ are convex, the differential inclusions (2.1) admit parametric representations of type (2.2), but in general they cannot be reduced to parametric control systems and should be studied for their own sake. Note also that the *ODE form* (2.2) in Banach spaces is strongly related to various control problems for evolution *partial differential equations* of parabolic and hyperbolic types, where solutions may be understood in some other appropriate senses; see, e.g., the books by Fattorini [6] and by Li and Yong [8].

Our principal method to study differential inclusions involves *finite-difference* replacements of the Fréchet derivative

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

where the *uniform Euler scheme* is considered for simplicity. To formalize this process, we take any natural number $N \in I\!N$ and consider the *discrete grid/mesh* on T defined by

$$T_N := \{a, a + h_N, \ldots, b - h_N, b\}, \quad h_N := (b-a)/N,$$

with the *stepsize of discretization* h_N and the *mesh points* $t_j := a + jh_N$ as $j = 0, \ldots, N$, where $t_0 = a$ and $t_N = b$. Then the differential inclusion (2.1) is replaced by a sequence of its *finite-difference/discrete approximations*

$$x_N(t_{j+1}) \in x_N(t_j) + h_N F(x_N(t_j), t_j), \quad j = 0, \ldots, N-1. \tag{2.3}$$

Given a discrete trajectory $x_N(t_j)$ satisfying (2.3), we consider its *piecewise linear extension* $x_N(t)$ to the continuous-time interval T, i.e., the *Euler broken lines*. We also define the *piecewise constant extension* to T of the corresponding *discrete velocity* by

$$v_N(t) := \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}, \quad t \in [t_j, t_{j+1}), \quad j = 0, \ldots, N-1.$$

It follows from the very definition of the Bochner integral that

$$x_N(t) = x_N(a) + \int_a^t v_N(s) \, ds \text{ for } t \in T.$$

Our first goal is to show that *every* solution to the differential inclusion (2.1) can be *strongly approximated*, under reasonable assumptions, by extended trajectories to the discrete inclusions (2.3). By strong approximation we understand the convergence in the norm topology of the classical Sobolev space $W^{1,2}([a, b]; X)$ with the norm

$$\|x(\cdot)\|_{W^{1,2}} := \max_{t \in [a,b]} \|x(t)\| + \left(\int_a^b \|\dot{x}(t)\|^2 \, dt \right)^{1/2},$$

where the norm on the right-hand side is taken in the space X. Note that the convergence in $W^{1,2}([a, b]; X)$ implies the *uniform* convergence of the trajectories on $[a, b]$ and the *pointwise* (a.e. $t \in [a, b]$) convergence of (some subsequence of) their derivatives. The latter is crucial for our purposes, especially in the case of *nonconvex* values $F(x, t)$.

Let us formulate the *basic assumptions* for our study that apply not only to the next theorem but also to the subsequent results on differential inclusions via discrete approximations. Nevertheless these assumptions can be relaxed in some settings; see the remarks and discussions below. Roughly speaking, we assume that the set-valued mapping $F: X \times [a, b] \rightrightarrows X$ is compact-valued, locally Lipschitzian in x, and Hausdorff continuous in t a.e. on $[a, b]$. More precisely, the following hypotheses are imposed along a given trajectory $\bar{x}(\cdot)$ to (2.1), which is arbitrary in the next theorem but then will be a reference optimal solution to the variational problem under consideration.

(H1) There are an open set $U \subset X$ and positive numbers m_F and ℓ_F such that $\bar{x}(t) \in U$ for all $t \in [a, b]$, the sets $F(x, t)$ are nonempty and compact for all $(x, t) \in U \times [a, b]$, and one has

$$F(x, t) \subset m_F \mathbb{B} \text{ for all } (x, t) \in U \times [a, b]. \tag{2.4}$$

$$F(x_1, t) \subset F(x_2, t) + \ell_F \|x_1 - x_2\| \mathbb{B} \text{ for all } x_1, x_2 \in U, \quad t \in [a, b]. \tag{2.5}$$

(H2) $F(x, \cdot)$ is Hausdorff continuous for a.e. $t \in [a, b]$ uniformly in $x \in U$.

4

Note that inclusion (2.5) is equivalent to the *uniform Lipschitz continuity*

$$\text{haus}(F(x,t), F(u,t)) \le \ell_F \|x - u\|, \quad x, u \in U,$$

of $F(\cdot, t)$ with respect to the *Pompieu-Hausdorff metric* haus$(\cdot, \cdot)$ on the space of nonempty and compact subsets of X; see Rockafellar and Wets [14].

To handle efficiently the Hausdorff continuity of $F(x, \cdot)$ for a.e. $t \in [a, b]$, define the for F in $t \in [a, b]$ while $x \in U$ by

$$\tau(F; h) := \int_a^b \sigma(F; t, h) \, dt, \tag{2.6}$$

where $\sigma(F; t, h) := \sup \{\omega(F; x, t, h) | \ x \in U\}$ with

$$\omega(F; x, t, h) := \sup \left\{ \text{haus}(F(x, t_1), F(x, t_2)) \Big| \ t_1, t_2 \in [t - \tfrac{h}{2}, t + \tfrac{h}{2}] \cap [a, b] \right\}.$$

The following observation is due to Dontchev and Farkhi [4].

Proposition 2.3 (averaged modulus of continuity). *The Hausdorff continuity property* (H2) *holds if and only if* $\tau(F; h) \to 0$ *as* $h \to 0$.

Note that for single-valued mapping $f: [a, b] \to X$ the property $\tau(f; h) \to 0$ as $h \to 0$ is *equivalent to the Riemann integrability* of f on $[a, b]$; see Sendov and Popov [13]. The latter holds, as well known, if and only if f is continuous at almost all $t \in [a, b]$.

3 Strong Approximation of Solutions to Differential Inclusions

The *strong approximation* theorem established in this section plays a crucial role in the subsequent results based on discrete approximations of differential inclusions.

Theorem 3.1 (strong approximation by discrete trajectories). *Let* $\bar{x}(\cdot)$ *be a solution to the differential inclusion* (2.1) *under assumptions* (H1) *and* (H2), *where* X *is an arbitrary Banach space. Then there is a sequence of solutions* $\hat{x}_N(t_j)$ *to the discrete inclusions* (2.3) *such that*

$$\hat{x}_N(a) = \bar{x}(a) \quad \text{for all} \quad N \in \mathbb{N}$$

and the extensions $\hat{x}_N(t)$, $a \le t \le b$, *converge to* $\bar{x}(t)$ *strongly in* $W^{1,2}([a, b]; X)$ *as* $N \to \infty$.

Proof. By Definition 2.1 involving the Bochner integral, the derivative mapping $\dot{\bar{x}}(\cdot)$ is *strongly measurable* on $[a, b]$, and hence we can find (rearranging the mesh points t_j if necessary) a sequence of simple/step mappings $w_N(\cdot)$ on T such that $w_N(t)$ are constant on $[t_j, t_{j+1})$ for every $j = 0, \ldots, N-1$ and $w_N(\cdot)$ converge to $\dot{\bar{x}}(\cdot)$ in the norm topology of $L^1([a, b]; X)$ as $N \to \infty$. Combining this convergence with (2.1) and (2.4), we get

$$\int_a^b \|w_N(t)\| \, dt = \sum_{j=0}^{N-1} \|w_N(t_j)\| \, (t_{j+1} - t_j) \le (m_F + 1)(b - a) \tag{3.1}$$

5

for all large N. In the estimates below we use the numerical sequence

$$\xi_N := \int_a^b \|\dot{\bar{x}}(t) - w_N(t)\| \, dt \to 0 \quad \text{as} \quad N \to \infty.$$

Let us define the discrete functions $u_N(t_j)$ by

$$u_N(t_{j+1}) = u_N(t_j) + h_N w_N(t_j), \quad j = 0, \ldots, N-1, \quad u_N(t_0) := \bar{x}(a)$$

and observe that the functions

$$u_N(t) := \bar{x}(a) + \int_a^t w_N(s) \, ds, \quad a \le t \le b,$$

are piecewise linear extensions of $u_N(t_j)$ to the interval $[a, b]$ and that

$$\|u_N(t) - \bar{x}(t)\| \le \int_a^t \|w_N(s) - \dot{\bar{x}}(s)\| \, ds \le \xi_N \quad \text{for} \quad t \in [a, b]. \tag{3.2}$$

Therefore $u_N(t) \in U$ for all $t \in [a, b]$ whenever N is sufficiently large.

Taking the *distance function* $\text{dist}(\cdot; \Omega)$ to a set in X, one can directly check that the Lipschitz condition (2.5) is equivalent to

$$\text{dist}(w; F(x_1, t)) \le \text{dist}(w; F(x_2, t)) + \ell_F \|x_1 - x_2\|$$

whenever $w \in X$, $x_1, x_2 \in U$, and $t \in [a, b]$. By definition (2.6) of $\tau(F; h)$ and the obvious relation

$$\text{dist}(w; F(x, t_1)) \le \text{dist}(w; F(x, t_2)) + \text{haus}(F(x, t_1), F(x, t_2))$$

one has the estimate

$$\zeta_N : = \sum_{j=0}^{N-1} h_N \text{dist}(w_N(t_j); F(u_N(t_j), t_j)) = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \text{dist}(w_N(t_j); F(u_N(t_j), t)) \, dt$$

$$\le \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \text{dist}(w_N(t_j); F(u_N(t), t)) \, dt + \tau(F; h_N).$$

The Lipschitz property of F and the construction of $w_N(\cdot)$ imply

$$\text{dist}(w_N(t_j); F(u_N(t_j), t)) \le \text{dist}(w_N(t); F(u_N(t_j), t)) + \ell_F w_N(t_j)(t - t_j)$$

whenever $t \in [t_j, t_{j+1})$, and then

$$\text{dist}(w_N(t); F(u_N(t), t)) \le \text{dist}(w_N(t); F(\bar{x}(t), t)) + \ell_F \|u_N(t) - \bar{x}(t)\|$$

$$\le \|w_N(t) - \dot{\bar{x}}(t)\| + \ell_F \xi_N \quad \text{a.e.} \quad t \in [a, b].$$

Employing further (3.1) and (3.2), we arrive at the estimate

$$\zeta_N \le \gamma_N := (1 + \ell_F(b - a))\xi_N + \ell_F(b - a)(m_F + 1)/2 + \tau(F; h_N). \tag{3.3}$$

6

Observe that the functions $u_N(t_j)$ built above are *not* trajectories for the discrete inclusions (2.3), since one doesn't have $w_N(t_j) \in F(u_N(t_j), t_j)$. Now we use $w_N(t_j)$ to construct *actual trajectories* $\widehat{x}_N(t_j)$ for (2.3) that are close to $u_N(t_j)$ and enjoy the convergence property stated in the theorem.

Let us define $\widehat{x}_N(t_j)$ recurrently by the following *proximal algorithm*, which is realized due to the compactness assumption on the values of F:

$$\begin{cases} \widehat{x}_N(t_0) = \bar{x}(a), \quad \widehat{x}_N(t_{j+1}) = \widehat{x}_N(t_j) + h_N v_N(t_j), \quad j = 0, \ldots, N-1, \\[2mm] \text{where } v_N(t_j) \in F(\widehat{x}_N(t_j), t_j) \text{ with} \\[2mm] \|v_N(t_j) - w_N(t_j)\| = \text{dist}(w_N(t_j); F(\widehat{x}_N(t_j), t_j)). \end{cases} \qquad (3.4)$$

First we prove that algorithm (3.4) keeps $\widehat{x}_N(t_j)$ inside the neighborhood U from (H1) whenever N is sufficiently large. Indeed, let us consider any number $N \in \mathbb{N}$ satisfying $\bar{x}(t) + \eta_N \mathbb{B} \subset U$ for all $t \in [a, b]$, where

$$\eta_N := \gamma_N \exp\left(\ell_F(b - a)\right) + \xi_N$$

with ξ_N and γ_N defined above. We have $\eta_N \to 0$ as $N \to \infty$, since $\xi_N \to 0$ by the construction of ξ_N and since $\gamma_N \to 0$ due to assumption (H2) equivalent to $\tau(F; h_N) \to 0$ by Proposition 2.3. Arguing by induction, we suppose that $\widehat{x}_N(t_i) \in U$ for all $i = 0, \ldots, j$ and show that this also holds for $i = j + 1$. Using (2.5), (3.3), and (3.4), one gets

$$\|\widehat{x}_N(t_{j+1}) - u_N(t_{j+1})\| \leq \|\widehat{x}_N(t_j) - u_N(t_j)\| + h_N\|v_N(t_j) - w_N(t_j)\|$$

$$\leq \|\widehat{x}_N(t_j) - u_N(t_j)\| + h_N\text{dist}(w_N(t_j); F(u_N(t_j), t_j))$$

$$+\ell_F\|\widehat{x}_N(t_j) - u_N(t_j)\| \leq \ldots$$

$$\leq h_N \sum_{i=0}^{j}(1 + \ell_F h_N)^{j-i}\text{dist}(w_N(t_i); F(u_N(t_i), t_i))$$

$$\leq h_N \exp\left[\ell_F(b - a)\right] \sum_{i=0}^{j} \text{dist}(w_N(t_i); F(u_N(t_i), t_i))$$

$$\leq \gamma_N \exp\left(\ell_F(b - a)\right).$$

Due to (3.2) the latter implies that

$$\|\widehat{x}_N(t_{j+1}) - \bar{x}_N(t_{j+1})\| \leq \gamma_N \exp\left(\ell_F(b - a)\right) + \xi_N =: \eta_N, \qquad (3.5)$$

which proves that $\widehat{x}_N(t_j) \in U$ for all $j = 0, \ldots, N$. Taking this into account, we have by the previous arguments that

$$\sum_{j=0}^{N} \|\widehat{x}_N(t_j) - u_N(t_j)\| \leq (b - a) \exp\left(\ell_F(b - a)\right) \sum_{j=0}^{N-1} \text{dist}(w_N(t_j); F(u_N(t_j), t_j)).$$

7

Now let us estimate the quantity

$$\vartheta_N := \int_a^b \|\hat{\dot{x}}_N(t) - w_N(t)\| \, dt \quad \text{as} \ N \to \infty.$$

Using the last estimate above together with (3.3) and (3.5), we have

$$\vartheta_N = \sum_{j=0}^{N-1} h_N \|\hat{\dot{x}}_N(t_j) - w_N(t_j)\| = \sum_{j=0}^{N-1} h_N \text{dist}(w_N(t_j); F(\hat{x}_N(t_j), t_j))$$

$$\leq \sum_{j=0}^{N-1} h_N \text{dist}(w_N(t_j); F(u_N(t_j), t_j)) + \ell_F \sum_{j=0}^{N-1} h_N \|\hat{x}_N(t_j) - u_N(t_j)\|$$

$$\leq \gamma_N (1 + \ell_F(b-a) \exp{(\ell_F(b-a))}).$$

Thus one finally gets

$$\int_a^b \|\hat{\dot{x}}_N(t) - \dot{\bar{x}}(t)\| \, dt \ \leq \int_a^b \|\hat{\dot{x}}_N(t) - \dot{\bar{x}}(t)\| \, dt + \int_a^b \|w_N(t) - \dot{\bar{x}}(t)\| \, dt$$

$$\tag{3.6}$$

$$\leq \gamma_N (1 + \ell_F(b-a) \exp{(\ell_F(b-a))}) + \xi_N := \alpha_N.$$

Since $\alpha_N \to 0$ as $N \to \infty$, this obviously implies the desired convergence $\hat{x}_N(\cdot) \to \bar{x}(\cdot)$ in the norm of $W^{1,2}([a,b]; X)$ due to the Newton-Leibniz formula for $\hat{x}_N(t)$ and $\bar{x}(t)$ and due to the uniform boundedness assumption (2.4). $\triangle$

Remark 3.2 (numerical efficiency of discrete approximations). It follows from (3.6) by the Newton-Leibniz formula that

$$\|\hat{x}_N(t) - \bar{x}(t)\| \leq \int_a^b \|\hat{\dot{x}}_N(t) - \dot{\bar{x}}(t)\| \, dt \leq \alpha_N \ \text{for all} \ t \in [a,b].$$

Thus the error estimate and numerical efficiency of the discrete approximation in Theorem 3.1 depend on the evaluation of the averaged modulus of continuity $\tau(F; h)$ from (2.6) and the approximating quantity ξ_N defined in the proof of Theorem 3.1. Denoting

$$v(F) := \sup_k \Big\{ \sum_{i=1}^{k-1} \sup_x [\text{haus}(F(x, t_{i+1}), F(x, t_i)), \ x \in U], \ a \leq t_1 \leq \ldots \leq t_k \leq b \Big\},$$

it is not hard to check that

$$\tau(F; h) \leq v(F)h = O(h)$$

whenever $F(x, \cdot)$ has a *bounded variation* $v(F) < \infty$ uniformly in $x \in U$; see Dontchev and Farkhi [4]. Furthermore, one has the estimate

$$\xi_N \leq 2\tau(\dot{\bar{x}}; h_N)$$

by taking $w_N(t) = \hat{\dot{x}}_N(t) = \dot{\bar{x}}(t_j)$ for $t \in [t_j, t_j + h_N)$ if $\dot{\bar{x}}(\cdot)$ is *Riemann integrable* on $[a, b]$.

4 The Bolza Problem for Nonconvex Differential Inclusions and Relaxation Stability

Let us consider the following generalized *Bolza problem* (P) of dynamic optimization over solutions (in the sense of Definition 2.1) to differential inclusions in Banach spaces: minimize the functional

$$J[x] := \varphi(x(a), x(b)) + \int_a^b \vartheta(x(t), \dot{x}(t), t) \, dt \tag{4.1}$$

over trajectories $x: [a, b] \to X$ for the differential inclusion (2.1) such that $\vartheta(x(t), \dot{x}(t), t)$ is Bochner integrable on the fixed time interval $T := [a, b]$ subject to the endpoint constraints

$$(x(a), x(b)) \in \Omega \subset X^2. \tag{4.2}$$

We use the term *arc* for any solution $x = x(\cdot)$ to (2.1) with $J[x] < \infty$ and the term *feasible arc* for arcs $x(\cdot)$ satisfying the endpoint constraints (4.2). Since the dynamic (2.1) and endpoint (4.2) constraints are given explicitly, we may assume that both functions φ and ϑ in the cost functional (4.1) take finite values.

The formulated problem (P) covers a broad range of various problems of dynamic optimization in finite-dimensional and infinite-dimensional spaces. In particular, it contains both standard and nonstandard models in optimal control for parameterized control systems (2.2) with possibly *closed-loop* control sets $U(x, t)$. Note also that problems with free time (non-fixed time intervals), with integral constraints on $(x, \dot{x})$, and with some other types of state constraints can be reduced to the form of (P).

In what follows, we study *optimal solutions* to (P) in the sense of *intermediate local minimizers* introduced by the author [9] and then employed in many publications; see, e.g., [1, 7, 17, 16] and the references therein.

Definition 4.1 (intermediate local minima). *A feasible arc $\bar{x}$ is an* INTERMEDIATE LOCAL MINIMIZER *(i.l.m.) of rank $p \in [1, \infty)$ for (P) if there are numbers $\varepsilon > 0$ and $\alpha \geq 0$ such that $J[\bar{x}] \leq J[x]$ for any feasible arcs to (P) satisfying*

$$\|x(t) - \bar{x}(t)\| < \varepsilon \quad \text{for all } t \in [a, b] \quad \text{and} \tag{4.3}$$

$$\alpha \int_a^b \|\dot{x}(t) - \dot{\bar{x}}(t)\|^p \, dt < \varepsilon. \tag{4.4}$$

Relationships (4.3) and (4.4) actually mean that we consider a neighborhood of $\bar{x}$ in the Sobolev space $W^{1,p}([a, b]; X)$. If there is only requirement (4.3) in Definition 4.1, i.e., $\alpha = 0$ in (4.4), that one gets the classical *strong* local minimum corresponding to a neighborhood of $\bar{x}$ in the norm topology of $\mathcal{C}([a, b]; X)$. If instead of (4.4) one puts the more restrictive requirement

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

then we have the classical *weak* local minimum in the framework of Definition 4.1. This corresponds to considering a neighborhood of $\bar{x}$ in the topology of $W^{1,\infty}([a, b]; X)$. Thus the introduced notion of i.l.m. takes, for any $p \in [1, \infty)$, an *intermediate* position between the classical concepts of strong

($\alpha = 0$) and weak ($p = \infty$) local minima. Clearly all the results for intermediate local minimizers automatically hold for strong (and hence for global) minimizers. We refer the reader to [9, 11, 17] for various examples that illustrate relationships between weak, intermediate, and strong local minimizers in variational (particularly optimal control) problems.

In what follows, along with the original problem (P), we consider its *relaxed* counterpart that, roughly speaking, is obtained from (P) by the *convexification* procedure with respect to the velocity variable. Taking the integrand $\vartheta(x, v, t)$ in (4.1), we consider its restriction

$$\vartheta_F(x, v, t) := \vartheta(x, v, t) + \delta(v; F(x, t))$$

to the sets $F(x, t)$ in (2.1) and denote by $\widehat{\vartheta}_F(x, v, t)$ the *biconjugate* function to $\vartheta_F(x, \cdot, t)$, i.e.,

$$\widehat{\vartheta}_F(x, v, t) = (\vartheta_F)_v^{**}(x, v, t) \quad \text{for all} \quad (x, v, t) \in X \times X \times [a, b].$$

It is well known that $\widehat{\vartheta}_F(x, v, t)$ is the *greatest proper, convex, l.s.c.* function with respect to v, which is *majorized* by ϑ_F. Moreover, $\vartheta_F = \widehat{\vartheta}_F$ *if and only if* ϑ_F is proper, convex, and l.s.c. with respect to the velocity variable v.

Given the original variational problem (P), we define the *relaxed problem* (R), or the *relaxation* of (P), as follows:

$$\text{minimize} \quad \widehat{J}[x] := \varphi(x(a), x(b)) + \int_a^b \widehat{\vartheta}_F(x(t), \dot{x}(t), t) \, dt \tag{4.5}$$

over a.e. differentiable arcs $x : [a, b] \to X$ that are Bochner integrable on $[a, b]$ together with $\vartheta_F(x(t), (x)(t), t)$, satisfy the Newton-Leibniz formula on $[a, b]$ and the endpoint constraints (4.2). Note that, in contrast to (4.1), the integrand in (4.5) is extended-real-valued. Furthermore, the natural requirement $\widehat{J}[x] < \infty$ yields that $x(\cdot)$ is a solution (in the sense of Definition 2.1) to the *convexified differential inclusion*

$$\dot{x}(t) \in \text{clco} \, F(x(t), \dot{x}(t), t) \quad \text{a.e.} \quad t \in [a, b]. \tag{4.6}$$

Thus the relaxed problem (R) can be considered under explicit dynamic constrained given by the convexified differential inclusion (4.6). Any trajectory for (4.6) is called a *relaxed trajectory* for (2.1), in contrast to *original trajectories*/arcs for the latter inclusion.

There are deep relationships between relaxed and original trajectories for differential inclusion, which reflect *hidden convexity* inherent in continuous-time (nonatomic measure) dynamic systems defined by differential operators; see [11, Chapter 6] for various implementations of the hidden convexity phenomenon with more references and discussions. In particular, *any relaxed trajectory* of compact-valued and Lipschitz in x differential inclusion in Banach spaces may be *uniformly approximated* (in the space $\mathcal{C}([a, b]; X)$) by original trajectories starting with the same initial state $x(a) = x_0$. We need a version of this approximation/density property involving not only differential inclusions but also minimizing functionals. The following result, which holds when the underlying Banach space is *separable*, is proved in [2]; results of this type go back to the classical theorems of Bogolyubov and Young in the calculus of variations; see [2, 11] for more details.

Theorem 4.2 (approximation property for relaxed trajectories). *Let $x(\cdot)$ be a relaxed trajectory for the differential inclusion (2.1), where X is separable, and where $F : X \times [a, b] \rightrightarrows X$*

is compact-valued and uniformly bounded by a summable function, locally Lipschitzian in x, and measurable in t. Assume also that the integrand ϑ in (4.1) is continuous in (x,v), measurable in t, and uniformly bounded by a summable function near $x(\cdot)$. Then there is sequence of the original trajectories $x_k(\cdot)$ for (2.1) satisfying the relations

$$x_k(a) = x(a), \quad x_k(\cdot) \to x(\cdot) \ \text{in} \ C([a,b]; X), \quad \text{and}$$

$$\liminf_{k \to \infty} \int_a^b \vartheta(x_k(t), \dot{x}_k(t), t)\, dt \le \int_a^b \widehat{\vartheta}_F(x(t), \dot{x}(t), t)\, dt.$$

Note that Theorem 4.2 *does not* assert that the approximating trajectories $x_k(\cdot)$ satisfy the endpoint constraints (4.2). Indeed, there are examples showing that the latter may not be possible. If they do, then problem (P) has the property of *relaxation stability*:

$$\inf(P) = \inf(R), \tag{4.7}$$

where the infima of the cost functionals (4.1) and (4.5) are taken over all the feasible arcs in (P) and (R), respectively.

An obvious sufficient condition for the relaxation stability is the *convexity* of the sets $F(x,t)$ and of the integrand f in v. However, the relaxation stability goes far beyond the standard convexity due to the hidden convexity property of continuous-time differential systems. In particular, Theorem 4.2 ensures the relaxation stability of nonconvex problems (P) with no constraints on $x(b)$. There are other efficient conditions for the relaxation stability of nonconvex problems discussed, e.g., [11] and the references therein. Let us mention the classical *Bogolyubov theorem* ensuring the relaxation stability in variational problems of minimizing (4.1) with endpoint constraint (4.2) but with *no dynamic constraints* (2.1); relaxation stability of control systems *linear in state variables* via the fundamental *Lyapunov theorem* on the range convexity of nonatomic vector measures that largely justifies the hidden convexity; the *calmness* condition by Clarke for differential inclusion problems with endpoint constraints of the inequality type; the *normality* condition by Warga involving parameterized control systems (2.2), etc.

An essential part of our study relates to *local minima* that are *stable with respect to relaxation*. The corresponding counterpart of Definition 4.1 is formulated as follows.

Definition 4.3 (relaxed intermediate local minima). *The arc $\bar{x}$ is a* RELAXED INTERMEDIATE LOCAL MINIMIZER *(r.i.l.m.) of rank $p \in [1, \infty)$ for the original problem (P) if $\bar{x}$ is a feasible solution to (P) and provides an intermediate local minimum of this rank to the relaxed problem (R) with the same cost $J[\bar{x}] = \widehat{J}[\bar{x}]$.*

The notions of *relaxed weak* and *relaxed strong local minima* are defined similarly, with the same relationships between them as discussed above. Of course, there is no difference between the corresponding relaxed and usual (non-relaxed) notions of local minima for problems (P) with convex sets $F(x,t)$ and integrands f convex with respect to velocity. It is also clear that any relaxed intermediate (weak, strong) minimizer for (P) provides the corresponding non-relaxed local minimum to the original problem. The opposite requires a kind of *local relaxation stability*.

Next build *well-posed* discrete approximations of a *given* r.i.l.m. $\bar{x}(\cdot)$ in problem (P) such that optimal solutions to discrete-time problems *strongly converge* to $\bar{x}(\cdot)$ in the space $W^{1,2}([a,b]; X)$.

11

This issue is of undoubted interest for both qualitative and numerical aspects of variational analysis for differential inclusions. In particular, it is used in [11] (following the finite-dimensional development of [9]) for deriving *necessary conditions* for intermediate local minimizers.

5 Strong Convergence of Optimal Solutions

Considering differential inclusions and their finite-difference counterparts in Section 2, we established there that *every* trajectory for a differential inclusion in a general Banach space can be *strongly approximated* by extended trajectories for finite-difference inclusions under the natural assumptions made. This result doesn't directly relate to optimization problems involving differential inclusions, but we are going to employ it now in the optimization framework. The primary objective of this section is as follows.

Given a trajectory $\bar{x}(\cdot)$ for the differential inclusion (2.1), which provides a *relaxed intermediate local minimum* (r.i.l.m.) to the optimization problem (P) defined in the previous section, construct a *well-posed* family of approximating optimization problems (P_N) for finite-difference inclusions (2.3) such that (extended) *optimal solutions* $\bar{x}_N(\cdot)$ to (P_N) *strongly converge* to $\bar{x}(\cdot)$ in the norm topology of $W^{1,2}([a,b];X)$.

Imposing the standing hypotheses (H1) and (H2) formulated in Section 2, observe that the boundedness assumption (2.4) implies that the notion of r.i.l.m. from Definition 4.3 *doesn't depend on rank p* from the interval $[1,\infty)$. This means that $\bar{x}(\cdot)$ is an r.i.l.m. of some rank $p \in [1,\infty)$, then it is also an r.i.l.m. of any other rank $p \geq 1$. In what follows we take $p = 2$ and $\alpha = 1$ in (4.4) for simplicity.

To proceed, we need to impose proper assumptions on the other data ϑ, φ, and Ω of problem (P) in addition to those imposed on F. Dealing with the Bochner integral, we always identify measurability of mappings $f \colon [a,b] \to X$ with *strong measurability*. Recall that f is strongly measurable if it can be a.e. approximated by a sequence of step X-valued functions on measurable subsets of $[a,b]$. Given a neighborhood U of $\bar{x}(\cdot)$ and a constant m_F from (H1), we further assume that:

(H3) $\vartheta(\cdot,\cdot,t)$ is continuous on $U \times (m_F I\!B)$ uniformly in $t \in [a,b]$, while $\vartheta(x,v,\cdot)$ is measurable on $[a,b]$ and its norm is majorized by a summable function uniformly in $(x,v) \in U \times (m_F I\!B)$.

(H4) φ is continuous on $U \times U$; $\Omega \subset X \times X$ is locally closed around $(\bar{x}(a),\bar{x}(b))$ and such that the set $\mathrm{proj}_1 \Omega \cap (\bar{x}(a) + \varepsilon I\!B)$ is compact for some $\varepsilon > 0$, where $\mathrm{proj}_1 \Omega$ stands for the projection of Ω on the first space X in the product space $X \times X$.

Note that symmetrically we may assume the local compactness of the second projection of $\Omega \subset X \times X$; the first one is selected in (H4) just form definiteness.

Now taking the r.i.l.m. $\bar{x}(\cdot)$ under consideration, let us apply to this feasible arc Theorem 3.1 on the strong approximation by discrete trajectories. Thus we find a sequence of the extended discrete trajectories $\hat{x}_N(\cdot)$ approximating $\bar{x}(\cdot)$ and compute the numbers η_N in (3.5). Having $\varepsilon > 0$ from relations (4.3) and (4.4) of the intermediate minimizer $\bar{x}(\cdot)$ with $p = 1$ and $\alpha = 1$, we always suppose that $\bar{x}(t) + \varepsilon/2 \in U$ for all $t \in [a,b]$. Let us construct the sequence of discrete approximation

problems (P_N), $N \in I\!N$, as follows: minimize the discrete-time Bolza functional

$$J_N[x_N] := \varphi(x_N(t_0), x_N(t_N)) + \|x_N(t_0) - \bar{x}(a)\|^2$$

$$\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta\left(x_N(t_j), \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}, t\right) dt \tag{5.1}$$

$$+ \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\|\frac{x_N(t_{j+1}) - x_N(t_j)}{h_N} - \dot{\bar{x}}(t)\right\|^2 dt$$

over discrete trajectories $x_N = x_N(\cdot) = (x_N(t_0), \ldots, x_N(t_N))$ for the difference inclusions (2.3) subject to the constraints

$$(x(t_0), x_N(t_N)) \in \Omega + \eta_N I\!B, \tag{5.2}$$

$$\|x_N(t_j) - \bar{x}(t_j)\| \le \frac{\varepsilon}{2} \quad \text{for } j = 1, \ldots, N, \quad \text{and} \tag{5.3}$$

$$\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\|\frac{x_N(t_{j+1}) - x_N(t_j)}{h_N} - \dot{\bar{x}}(t)\right\|^2 dt \le \frac{\varepsilon}{2}. \tag{5.4}$$

As in Section 2, consider (without mentioning any more) piecewise linear extensions of $x_N(\cdot)$ to the whole interval $[a, b]$ with piecewise constant derivatives, for which one has

$$\begin{cases} x_N(t) = x_N(a) + \int_a^t \dot{x}_N(s)\, ds \quad \text{for all } t \in [a, b] \quad \text{and} \\ \\ \dot{x}_N(t) = \dot{x}_N(t_j) \in F(x_N(t_j), t_j), \quad t \in [t_j, t_{j+1}), \quad j = 0, \ldots, N-1. \end{cases} \tag{5.5}$$

The next theorem establishes that the given local minimizer $\bar{x}(\cdot)$ to (P) can be approximated by *optimal solutions* to (P_N) *strongly* in $W^{1,2}([a, b]; X)$, which implies a.e. *pointwise* convergence of the derivatives essential in what follows. To justify such an approximation, we need to impose both the Asplund structure and the Radon-Nikodým property (RNP) on the space X in question, which ensure the validity of the classical Dunford theorem on the weak compactness in $L^1([a, b]; X)$. It is worth noting that every *reflexive* space is Asplund and has the RNP simultaneously. Furthermore, the second dual space X^{**} enjoys the RNP (and hence so does $X \subset X^{**}$) if X^* is Asplund. Thus the class of Banach spaces X for which both X and X^* are Asplund satisfies the properties needed in the next theorem. As well known in the geometric theory of Banach spaces, there are *nonreflexive* (even separable) spaces that fall into this category.

Theorem 5.1 (strong convergence of discrete optimal solutions). *Let $\bar{x}(\cdot)$ be an r.i.l.m. for the Bolza problem (P) under assumptions (H1)–(H4), and let (P_N), $N \in I\!N$, be a sequence of discrete approximation problems built above. The following hold:*

(i) *Each (P_N) admits an optimal solution.*

(ii) *If in addition X is Asplund and has the RNP, then any sequence $\{\bar{x}_N(\cdot)\}$ of optimal solutions to (P_N) converges to $\bar{x}(\cdot)$ strongly in $W^{1,2}([a, b]; X)$.*

Proof. To justify (i), we observe that the set of feasible trajectories to each problem (P_N) is nonempty for all large N, since the extended functions $\widehat{x}_N(\cdot)$ from Theorem 3.1 satisfy (2.3) and the constraints (5.2)–(5.4) by construction. This follows immediately from (3.5) in the case of (5.2) and (5.3). In the case of (5.4) we get from (2.4) and (3.6) that

$$\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{\widehat{x}_N(t_{j+1}) - \widehat{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\|^2 dt = \int_a^b \|\dot{\widehat{x}}_N(t) - \dot{\bar{x}}(t)\|^2 dt$$

$$\leq 2m_F \alpha_N \leq \frac{\varepsilon}{2}$$

for large N by the formula for α_N in (3.6). The existence of optimal solutions to (P_N) follows now from the classical Weierstrass theorem due to the compactness and continuity assumptions made in (H1), (H3), and (H4).

It remains to prove the convergence assertion (ii). Check first that

$$J_N[\widehat{x}_N] \to J[\bar{x}] \quad \text{as} \quad N \to \infty \tag{5.6}$$

along some sequence of $N \in I\!N$. Considering the expression (5.1) for $J_N[\widehat{x}_N]$, we deduce from Theorem 3.1 that the second terms therein vanishes, the forth term tends to zero due to (2.4) and (3.6), and the first term tends to $\varphi(\bar{x}(a), \bar{x}(b))$ due to the continuity assumption on φ in (H4). It is thus sufficient to show that

$$\sigma_N := \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta\Big(\widehat{x}_N(t_j), \frac{\widehat{x}_N(t_{j+1}) - \widehat{x}_N(t_j)}{h_N}, t\Big) dt \to \int_a^b \vartheta(\bar{x}(t), \dot{\bar{x}}(t), t) dt$$

as $N \to \infty$. Using the sign "$\sim$" for expressions that are equivalent as $N \to \infty$, we get the following limiting relationships

$$\sigma_N = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta(\widehat{x}_N(t_j), \dot{\widehat{x}}_N(t), t) dt \sim \int_a^b \vartheta(\widehat{x}_N(t), \dot{\widehat{x}}_N(t), t) dt$$

$$\sim \int_a^b \vartheta(\bar{x}(t), \dot{\widehat{x}}_N(t), t) dt \sim \int_a^b \vartheta(\bar{x}(t), \dot{\bar{x}}(t), t) dt$$

by Theorem 3.1 ensuring the a.e. convergence $\dot{\widehat{x}}_N(t) \to \dot{\bar{x}}(t)$ along a subsequence of $N \to \infty$ and by the Lebesgue dominated convergence theorem for the Bochner integral that is valid under (H3).

Note that we have justified (5.6) for any intermediate (not relaxed) local minimizer $\bar{x}(\cdot)$ for the original problem (P) in an arbitrary Banach space X. Next let us show that (5.6) implies that

$$\lim_{N \to \infty} \Big[\beta_N := \|\bar{x}_N(a) - \bar{x}(a)\|^2 + \int_a^b \|\dot{\bar{x}}_N(t) - \dot{\bar{x}}(t)\|^2 dt \Big] = 0 \tag{5.7}$$

for every sequence of optimal solutions $\bar{x}_N(\cdot)$ to (P_N) provided that $\bar{x}(\cdot)$ is a *relaxed* intermediate local minimizer for the original problem, where the state space X is assumed to be Asplund and to satisfy the RNP.

Suppose that (5.7) is not true. Take a limiting point $\beta > 0$ of the sequence $\{\beta_N\}$ in (5.7) and let for simplicity that $\beta_N \to \beta$ for all $N \to \infty$. We are going to apply the Dunford theorem on the relative *weak compactness* in the space $L^1([a, b]; X)$ (see, e.g., [3, Theorem IV.1]) to the sequence

14

$\{\dot{\bar{x}}_N(\cdot)\}$, $N \in I\!N$. Due to (5.5) and (H1) this sequence satisfies the assumptions of the Dunford theorem. Furthermore, both spaces X and X^* have the RNP, since the latter property for X^* is equivalent to the Asplund structure on X, as mentioned above. Hence we suppose without loss of generality that there is $v \in L^1([a,b];X)$ such that

$$\dot{\bar{x}}_N(\cdot) \to v(\cdot) \quad \text{weakly in} \quad L^1([a,b];X) \quad \text{as} \quad N \to \infty.$$

Since the Bochner integral is a linear continuous operator from $L^1([a,b];X)$ to X, it remains continuous if the spaces $L^1([a,b];X)$ and X are endowed with the weak topologies. Due to (5.2) and the assumptions on Ω in (H4), the set $\{\bar{x}_N(a) | N \in I\!N\}$ is relatively compact in X. Using (5.5) and the *compactness* property of solution sets for differential inclusions under the assumptions made in (H1) and (H2), we conclude that the sequence $\{\bar{x}_N(\cdot)\}$ contains a subsequence that converges to some $\tilde{x}(\cdot)$ in the norm topology of the space $C([a,b];X)$. Now passing to the limit in the Newton-Leibniz formula for $\bar{x}_N(\cdot)$ in (5.5) and taking into account the above convergences, one has

$$\tilde{x}(t) = \tilde{x}(a) + \int_a^t v(s)\, ds \quad \text{for all} \quad t \in [a,b], \cdot$$

which implies the absolute continuity and a.e. differentiability of $\tilde{x}(\cdot)$ on $[a,b]$ with $v(t) = \dot{\tilde{x}}(t)$ for a.e. $t \in [a,b]$. Observe that $\tilde{x}(\cdot)$ is a solution to the convexified differential inclusions (4.6). Indeed, since a subsequence of $\{\bar{x}_N(\cdot)\}$ converges to $\tilde{x}(\cdot)$ weakly in $L^1([a,b];X)$, some *convex combinations* of $\bar{x}_N(\cdot)$ converge to $\dot{\tilde{x}}(\cdot)$ in the norm topology of $L^1([a,b];X)$, and hence *pointwisely* for a.e. $t \in [a,b]$. Passing to the limit in the differential inclusions for $\bar{x}_N(\cdot)$ in (5.5), we conclude that $\tilde{x}(\cdot)$ satisfies (4.6). By passing to the limit in (5.2) and (5.3), we also conclude that $\tilde{x}(\cdot)$ satisfies the endpoint constraints in (4.2) as well as

$$\|\tilde{x}(t) - \bar{x}(t)\| \le \varepsilon/2 \quad \text{for all} \quad t \in [a,b].$$

Furthermore, the integral functional

$$I[v] := \int_a^b \|v(t) - \dot{\bar{x}}(t)\|^2\, dt$$

is lower semicontinuous in the weak topology of $L^2([a,b];X)$ due to the convexity of the integrand in v. Since the weak convergence of $\dot{\bar{x}}_N(\cdot) \to \dot{\tilde{x}}(\cdot)$ in $L^1([a,b];X)$ implies the one in $L^2([a,b];X)$ by the boundedness assumption (2.4), and since

$$\int_a^b \|\dot{\bar{x}}_N(t) - \dot{\bar{x}}(t)\|^2\, dt = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\|^2 dt,$$

the above lower semicontinuity and relation (5.4) imply that

$$\int_a^b \|\dot{\tilde{x}}(t) - \dot{\bar{x}}(t)\|^2\, dt \le \liminf_{N\to\infty} \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\|^2 dt \le \frac{\varepsilon}{2}.$$

Thus the arc $\tilde{x}(\cdot)$ belongs to the ε neighborhood of $\bar{x}(\cdot)$ in the space $W^{1,2}([a,b];X)$.

Let us finally show that the arc $\tilde{x}(\cdot)$ gives a smaller value to cost functional (4.5) than $\bar{x}(\cdot)$. One always has

$$J_N[\bar{x}_N] \le J_N[\hat{x}_N] \quad \text{for all large} \quad N \in I\!N,$$

15

since each $\widehat{x}_N(\cdot)$ is feasible to (P_N). Now passing to the limit as $N \to \infty$ and taking into account the previous discussions as well as the construction of the convexified integrand $\widehat{\vartheta}_F$ in (4.5), we get from (5.6) that

$$\varphi(\widetilde{x}(a), \widetilde{x}(b)) + \int_a^b \widehat{\vartheta}_F(\widetilde{x}(t), \dot{\widetilde{x}}(t), t)\, dt + \beta \leq J[\widetilde{x}],$$

which yields by $\beta > 0$ that $\widehat{J}[\widetilde{x}] < J[\widetilde{x}] = \widehat{J}[\widetilde{x}]$. The latter is impossible, since $\bar{x}(\cdot)$ is an r.i.l.m. for (P). Thus (5.7) holds, which obviously implies the desired convergence $\widetilde{x}_N(\cdot) \to \bar{x}(\cdot)$ in the norm topology of the space $W^{1,2}([a, b]; X)$ and completes the proof of the theorem. $\triangle$

The arguments developed in the proof of Theorem 5.1 allow us to establish efficient conditions for the *value convergence* of discrete approximations, which means that the optimal/infimal values of the cost functionals in discrete approximation problems converge to the one in the original problem (P). Moreover, using the approximation property for relaxed trajectories from Theorem 4.2, we obtain in fact a *necessary and sufficient* condition for the value convergence in terms of an intrinsic property of the original problems.

6   Value Convergence of Discrete Approximations

Observe that the cost functional (5.1) as well as the constraints (5.3) and (5.4) in the discrete approximation problems (P_N) explicitly contain the given local minimizer $\bar{x}(\cdot)$ to (P). Considering below the value convergence of discrete approximations, we are *not* going to involve *any local minimizer* in the construction of discrete approximations and/or even to assume the *existence of optimal solutions* to the original problem. To furnish this, we consider a sequence of new discrete approximation problems $(\widetilde{P}_N)$ built as follows: minimize

$$\widetilde{J}_N[x_N] := \varphi(x_N(t_0), x_N(t_N)) + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta\left(x_N(t_j), \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}, t\right) dt$$

subject to the discrete inclusions (2.3) and the *perturbed* endpoint constraints (5.2), where the sequence η_N is not yet specified. Note that problems $(\widetilde{P}_N)$ are constructively built upon the initial data of the original continuous-time problem. In the next theorem the notation $\widetilde{J}_N^0 := \inf(\widetilde{P}_N)$, $\inf(P)$, and $\inf(R)$ stands for the optimal value of the cost functional in problems $(\widetilde{P}_N)$, (P), and (R), respectively. Observe that optimal solutions to the discrete-time problems $(\widetilde{P}_N)$ always *exist* due to the assumptions (H1)–(H4) made in Theorem 5.1 under proper perturbations η_N of the endpoint constraints; see its proof.

Theorem 6.1 (criterion for value convergence via relaxation stability). *Let $U \subset X$ be an open subset of a Banach space X such that $x_k(t) \in U$ as $t \in [a, b]$ and $k \in \mathbb{N}$ for a minimizing sequence of feasible solutions to (P). Assume that hypotheses (H1)–(H4) are fulfilled with this set U, where $\bar{x}(a) + \varepsilon \mathbb{B}$ is replaced by $\mathrm{cl}\, U$ in (H4). The following assertions hold:*

(i) There is a sequence of the endpoint constraint perturbations $\eta_N \downarrow 0$ in (5.2) such that

$$\inf(R) \leq \liminf_{N\to\infty} \widetilde{J}_N^0 \leq \limsup_{N\to\infty} \widetilde{J}_N^0 \leq \inf(P), \tag{6.1}$$

where the left-hand side inequality requires that X is Asplund and has the RNP. Therefore the relaxation stability (4.7) of (P) is sufficient for the value convergence of discrete approximations

$$\inf(\widetilde{P}_N) \to \inf(P) \quad as \quad N \to \infty$$

provided that X is Asplund and has the RNP.

   (ii) *Conversely, the relaxation stability of (P) is also a necessary condition for the value convergence $\inf(\widetilde{P}_N) \to \inf(P)$ of the discrete approximations with arbitrary perturbations $\eta_N \downarrow 0$ of the endpoint constraints provided that X is reflexive and separable.*

Proof. Let us first prove that the right-hand side inequality in (6.1) holds in any Banach space X. Taking the minimizing sequence of feasible arcs $x_k(\cdot)$ to (P) specified in the theorem, we apply to each $x_k(\cdot)$ Theorem 3.1 on the strong approximation by discrete trajectories. Involving the diagonal process, we build the extended discrete trajectories $\widehat{x}_N(\cdot)$ for (2.3) such that

$$\eta_N := \|(\widehat{x}_N(a), \widehat{x}_N(b)) - (x_{k_N}(a), x_{k_N}(b))\| \to 0 \quad as \, \ N \to \infty$$

and consider the sequence of discrete approximation problems $(\widetilde{P}_N)$ with these constraint perturbations η_N in (5.2). Similarly to the proof of the first part of Theorem 5.1, we show that each $(\widetilde{P}_N)$ admits an optimal solution and, arguing by contradiction, one has the right-hand side inequality in (6.1). To justify the left-hand side inequality in (6.1), we follow the proof of the second part of Theorem 5.1 assuming that X is Asplund and enjoys the RNP. This automatically implies the value convergence of $\inf(\widetilde{P}_N) \to \inf(P)$ under the relaxation stability of (P).

   To prove the conversed assertion (ii) in the theorem, we first observe that the relaxed problem (R) admits an optimal solution under the assumptions made; see Tolstonogov [15, Theorem A.1.3]. It follows from the arguments in the second part of Theorem 5.1 that actually justify, under the assumptions made, the compactness of feasible solutions to the relaxed problem and the lower semicontinuity of the minimizing functional (4.5) in the topology on the set of feasible solutions $x(\cdot)$ induced by the weak convergence of the derivatives $\dot{x}(\cdot) \in L^1([a,b];X)$ provided that X is Asplund and has the RNP. Assume now that X is reflexive and separable and, employing Theorem 4.2, approximate some relaxed optimal trajectory $\bar{x}(\cdot)$ by a sequence of the original trajectories $x_k(\cdot)$ converging to $\bar{x}(\cdot)$ as established in that theorem. In turn, each $x_k(\cdot)$ can be strongly approximated in $W^{1,2}([a,b];X)$ by discrete trajectories $\widehat{x}_{k_N}(\cdot)$ due to Theorem 3.1. Using the diagonal process, we get a sequence of the discrete trajectories $\widehat{x}_N(\cdot)$ approximating $\bar{x}(\cdot)$ and put

$$\eta_N := \|(\widehat{x}_N(a), \widehat{x}_N(b)) - (\bar{x}(a), \bar{x}(b))\| \to \infty \quad as \quad N \to \infty.$$

Now assume that problem (P) is not stable with respect to relaxation, i.e., $\inf(R) < \inf(P)$, and then show that

$$\liminf_{N \to \infty} \widetilde{J}_N^0 < \inf(P)$$

for a sequence of discrete approximation problems $(\widetilde{P}_N)$ with some perturbations η_N of the endpoint constraints (5.2). Indeed, having

$$\inf(R) = \varphi(\bar{x}(a), \bar{x}(b)) + \int_a^b \widehat{\vartheta}_F(\bar{x}(t), \dot{x}(t), t)\, dt < \inf(P)$$

17

for the relaxed optimal trajectory $\bar{x}(\cdot)$, we build η_N as above and consider problems $(\tilde{P}_N)$ with these perturbations of the endpoint constraints. Taking into account the approximation of $\bar{x}(\cdot)$ by $x_k(\cdot)$ due to Theorem 4.2, the strong approximation of $x_k(\cdot)$ by the discrete trajectories $\hat{x}_N(\cdot)$ in Theorem 3.1, and

$$
\tilde{J}_N^0 \leq \varphi(\hat{x}_N(t_0), \hat{x}_N(t_N)) + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta\left(\hat{x}_N(t_j), \frac{\hat{x}_N(t_{j+1}) - \hat{x}_N(t_j)}{h_N}, t\right) dt
$$

$$
= \varphi(\hat{x}_N(a), \hat{x}_N(b)) + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta(\hat{x}_N(t_j), \dot{\hat{x}}_N(t), t) \, dt,
$$

one gets by the absence of the relaxation stability that

$$
\liminf_{N \to \infty} \tilde{J}_N^0 \leq \liminf_{N \to \infty} \left[\varphi(\hat{x}_N(a), \hat{x}_N(b)) + \int_a^b \vartheta(\hat{x}_N(t), \dot{\hat{x}}_N(t), t) \, dt \right]
$$

$$
\leq \varphi(\bar{x}(a), \bar{x}(b)) + \int_a^b \hat{\vartheta}_F(\bar{x}(t), \dot{\bar{x}}(t), t) \, dt < \inf(P).
$$

Therefore we don't have the value convergence of discrete approximations for problems $(\tilde{P}_N)$ corresponding to the above perturbations of the endpoint constraints. This justifies (ii) and completes the proof of the theorem. $\triangle$

Thus the relaxation stability of (P), which is an intrinsic and natural property of continuous-time dynamic optimization problems, is actually a *criterion* for the value convergence of discrete approximations under appropriate perturbations of the endpoint constraints in (5.2). It follows from the proof of Theorem 6.1 that one can express the corresponding perturbations η_N via the averaged modulus of continuity (2.6) by

$$
\eta_N = \tau(\dot{\bar{x}}; h_N) \to \infty \quad \text{as} \quad N \to \infty
$$

provided that (P) admits an optimal solution $\bar{x}(\cdot)$ with the Riemann integrable derivative $\dot{\bar{x}}(\cdot)$ on $[a, b]$. Moreover, $\eta_N = O(h_N)$ if $\dot{\bar{x}}(t)$ is of bounded variation on this interval; see Section 2.

Remark 6.2 (simplified form of discrete approximations). Observe that if $\vartheta(x, v, \cdot)$ is *a.e. continuous* on $[a, b]$ uniformly in (x, v) in some neighborhood of the optimal solution $\bar{x}(\cdot)$, then the cost functional in (5.1) in problem (P_N) can be replaced in Theorem 5.1 by

$$
J_N[x_N] : = \varphi(x_N(t_0), x_N(t_N)) + \|x_N(t_0) - \bar{x}(a)\|^2
$$

$$
+ h_N \sum_{j=0}^{N-1} \vartheta\left(x_N(t_j), \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}, t_j\right)
$$

$$
+ \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\|^2 dt;
$$

(6.2)

similarly for the cost functional in problem $(\widetilde{P}_N)$ used in Theorem 6.1. Indeed, this is an easy consequence of the fact that $\tau(\vartheta; h_N) \to 0$ as $N \to \infty$ for the averaged modulus of continuity (2.6) when $\vartheta(x, v, \cdot)$ is a.e. continuous. Denote by $(\overline{P}_N)$ the discrete approximation problem that differs from (P_N) of that the cost functional (5.1) is replaced by the simplified one (6.2). In what follows we consider both problems (P_N) and $(\overline{P}_N)$ using them to derive necessary optimality conditions for the original problem. The results obtained in these ways are distinguished by the assumptions on the initial data that allow us to justify the desired necessary optimality conditions. Namely, while the use of the simplified problems $(\overline{P}_N)$ as $N \to \infty$ requires the a.e. continuity assumption on ϑ with respect to t (versus the measurability), it relaxes the requirements on the state space X needed in the case of (P_N); see [11, Chapter 6].

References

[1] F. H. Clarke (2005), Necessary conditions in dynamic optimization, *Mem. Amer. Math. Soc.* **173**, No. 816.

[2] F. S. De Blasi, G. Pianigiani and A. A. Tolstonogov (2004), A Bogolyubov type theorem with a nonconvex constraint in Banach spaces, *SIAM J. Control Optim.* **43**, 466-476.

[3] J. Diestel and J. J. Uhl, Jr. (1977), *Vector Measures,* American Mathematical Society, Providence, Rhode Island.

[4] A. L. Dontchev and E. M. Farkhi (1989), Error estimates for discretized differential inclusions, *Computing* **41**, 349-358.

[5] A. L. Dontchev and F. Lempio (1992), Difference methods for differential inclusions: A survey, *SIAM Rev.* **34**, 263-294.

[6] H. O. Fattorini (1999), *Infinite Dimensional Optimization and Control Theory*, Cambridge University Press, Cambridge, Massachusetts.

[7] A. D. Ioffe (1997), Euler-Lagrange and Hamiltonian formalisms in dynamic optimization, *Trans. Amer. Amer. Soc.* **349**, 2871-2900.

[8] X. Li and J. Yong (1995), *Optimal Control Theory for Infinite-Dimensional Systems*, Birkhäuser, Boston, Massachusetts.

[9] B. S. Mordukhovich (1995), Discrete approximations and refined Euler-Lagrange conditions for nonconvex differential inclusions, *SIAM J. Control Optim.* **33**, 882-915.

[10] B. S. Mordukhovich (2006), *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. **330**, Springer, Berlin.

[11] B. S. Mordukhovich (2006), *Variational Analysis and Generalized Differentiation, II: Applicationds*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. **331**, Springer, Berlin.

[12] R. R. Phelps (1993), *Convex Functions, Monotone Operators and Differentiability*, 2nd edition, Lecture Notes in Mathematics, Vol. **1364**, Springer, Berlin.

[13] B. Sendov and V. A. Popov (1988), *The Averaged Moduli of Smoothness*, Wiley, New York.

[14] R. T. Rockafellar and R. J-B. Wets (1998), *Variational Analysis*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. **317**, Springer, Berlin.

[15] A. A. Tolstonogov (2000), *Differential Inclusions in a Banach Spaces*, Kluwer, Dordrecht, The Netherlands.

[16] R. B. Vinter (2000), *Optimal Control*, Birkhäuser, Boston, Massachusetts.

[17] R. B. Vinter and P. D. Woodford (1997), On the occurrence of intermediate local minimizers that are not strong local minimizers, *Systems Cont. Lett.* **31**, 235–342.