

**OPTIMIZATION AND FINITE DIFFERENCE  
APPROXIMATIONS OF NONCONVEX DIFFERENTIAL  
INCLUSIONS WITH FREE TIME**

By

**Boris Mordukhovich**

**IMA Preprint Series # 1174**

September 1993

# OPTIMIZATION AND FINITE DIFFERENCE APPROXIMATIONS OF NONCONVEX DIFFERENTIAL INCLUSIONS WITH FREE TIME

BORIS MORDUKHOVICH<sup>1</sup>

Department of Mathematics

Wayne State University

Detroit, Michigan 48202

E-mail: boris@math.wayne.edu

July 1993

**Abstract.** This paper is concerned with a free-time optimal control problem for nonconvex-valued differential inclusions with a nonsmooth cost functional in the form of Bolza and general endpoint constraints involving free time. We develop a finite difference method for studying this problem and focus on two major topics: 1) constructions of well-posed discrete approximations ensuring a strong convergence of optimal solutions, and 2) necessary optimality conditions for free-time differential inclusions obtaining by the limiting process from discrete approximations. As a result, we construct a sequence of discrete approximations with the strong convergence of optimal solutions in the  $W^{1,2}$ -norm. Then using the convergence result and appropriate tools of nonsmooth analysis, we prove necessary optimality conditions for differential inclusions in the refined Euler-Lagrange form with a new relation for an optimal free time.

---

<sup>1</sup>This research was supported in part by grant DMS-9206989 from the National Science Foundation and the Career Development Chairs Award at Wayne State University and also by the Institute for Mathematics and its Applications with funds provided by the National Science Foundation

# 1. Introduction

In this paper we study problem  $(P)$  of minimizing the real-valued Bolza functional

$$J[x, T] := \varphi(x(0), x(T), T) + \int_0^T f(t, x(t), \dot{x}(t))dt \quad (1.1)$$

on trajectories for the compact-valued and Lipschitz continuous differential inclusion

$$\dot{x}(t) \in F(t, x(t)) \quad (1.2)$$

over a varying time interval subject to general endpoint constraints on  $(x(0), x(T), T)$ . For brevity, we shall refer to this problem as to a problem of Bolza with *free time*. If  $f = 0$ , then  $(P)$  is said to be a Mayer free-time problem.

Such differential inclusion problems are natural generalizations of free-time problems in both the calculus of variations and optimal control. The latter case corresponds to the representation

$$Q(t, x) = \{g(t, x, U) \mid u \in U\} \quad (1.3)$$

with some vector function  $g$  and set  $U$  which may depend on the time variable. Moreover, the differential inclusion model (1.2) allows us to consider closed-loop control systems with time-dependent control regions  $U = U(x)$  in (1.3).

It is known even in classical settings that problems with free time have certain qualitative and technical distinctions from corresponding fixed-time problems. Of course, fixed-time problems can be considered as special cases of problems with free time. But usually results for fixed-time problems can be easier obtained under more general assumptions about the dependence on time variables.

On the other hand, in many situations necessary optimality conditions for free-time problems may be derived from corresponding results for fixed-time problems under additional regularity assumptions on the time-dependence of the data. Such transformation techniques are widely used in the classical variational and control problems; see, e.g., [26].

To our knowledge, the first set of necessary optimality conditions for differential inclusions with free time was published in Clarke's book [11, Section 3.6]. Considering the autonomous case for *convex* (this means convex-valued) differential inclusions, he obtained necessary

conditions for a free-time Mayer problem by reducing it a fixed-time one and using the so-called (true) *Hamiltonian*

$$\mathcal{H}(x, p) := \max\{\langle p, v \rangle \mid v \in F(x)\} \quad (1.4)$$

which is a Lipschitz continuous function in the state variable  $x$  and the adjoint variable  $p$ .

Clarke established the Hamiltonian necessary conditions for an optimal solution  $\{\bar{x}(t), 0 \leq t \leq \bar{T}\}$  in terms of his generalized gradients  $\partial_C$  of Lipschitz continuous functions. These conditions involve the *Hamiltonian inclusion*

$$(-\dot{p}(t), \dot{x}(t)) \in \partial_C \mathcal{H}(\bar{x}(t), p(t)) \quad \text{a.e. } t \in [0, \bar{T}] \quad (1.5)$$

with additional relations for  $(p(0), p(\bar{T}), \bar{T})$  expressed in terms of the generalized gradients of  $\varphi$  and endpoint constraint functions. If, in particular,  $\varphi$  and endpoint constraints do not depend on free time  $T$ , then Clarke's conditions imply that

$$\mathcal{H}(\bar{x}(t), p(t)) \equiv 0 \quad \text{on } [0, \bar{T}]. \quad (1.6)$$

The latter constancy relation is well known for the classical problems of optimal control where it actually follows from the maximum condition in the Pontryagin maximum principle; see [26, 46, 57].

The nonautonomous case can be easily reduced to the autonomous one if  $F$  is Lipschitzian with respect to both variables  $(t, x)$ . The case of merely continuous dynamics is more complicated even if a standard differential equation formulation is adopted; see Berkovitz [3] for smooth problems and Mordukhovich [34, 35, 37] for nonsmooth problems of optimal control with free time. We refer the reader to the papers of Clarke and Vinter [13, 14], Clarke, Loewen, and Vinter [15], and Rowland and Vinter [54] for studying as well as applications of free-time and related optimization problems for convex differential inclusions with discontinuous (in general measurable) time dependence. The mentioned papers contain necessary optimality conditions of the Hamiltonian type generalized the results of [11].

Another version of necessary optimality conditions for convex differential inclusions was developed by Mordukhovich first for fixed-time [33, 35] and then for free-time [39] Mayer problems. In this version, the main differential relation is obtained in the following form (all

the results in the rest of this section are formulated only for autonomous systems):

$$\begin{aligned} (\dot{p}(t), \dot{\bar{x}}(t)) \in \text{co}\{(u, v) \mid (u, p(t)) \in N((\bar{x}(t), v); \text{gph } F), \\ v \in M(\bar{x}(t), p(t))\} \quad \text{a.e. } t \in [0, \bar{T}] \end{aligned} \quad (1.7)$$

where "co" stands for the convex hull and

$$M(x, p) := \{v \in F(x) \mid \langle p, x \rangle = \mathcal{H}(x, p)\}. \quad (1.8)$$

In (1.7),  $N$  is not Clarke's normal cone but its *nonconvex* counterpart which was first used in Mordukhovich [31] for obtaining transversality conditions in nonsmooth optimal control problems. Now it is clear that this normal cone and the corresponding nonconvex subdifferential  $\partial$  are, probably, the most convenient tools for describing transversality and related conditions for dynamic optimization problems as well as necessary conditions in finite dimensional nonsmooth optimization; see, e.g., [12, 22, 28, 29, 33–39, 52–54] and references therein. We shall consider some properties of these objects in Section 4.

It is rather surprising that the dynamic relationships (1.5) and (1.7) appear to be *equivalent* in the general framework of necessary conditions for convex differential inclusions. This has been recently proved by Rockafellar [51] in the direction (1.5)  $\implies$  (1.7) and Ioffe (personal communication; see also [24, Section 3.5]) in the opposite direction.

Let us consider condition (1.7) and its further improvements in more details. First observe that (1.7) implies for a.e.  $t \in [0, \bar{T}]$  both the *maximum condition*

$$\langle p(t), \dot{\bar{x}}(t) \rangle = \mathcal{H}(\bar{x}(t), p(t)) \quad (1.9)$$

and an analogue of the *Euler-Lagrange inclusion* in the form

$$\dot{p}(t) \in \text{co}\{u \mid (u, p(t)) \in N((\bar{x}(t), v); \text{gph } F), v \in M(\bar{x}(t), p(t))\}. \quad (1.10)$$

Note that version (1.10) is, in general, independent of Clarke's version of the Euler-Lagrange inclusion [10]

$$(\dot{p}(t), p(t)) \in \text{cl co } N((\bar{x}(t), \dot{\bar{x}}(t)); \text{gph } F); \quad (1.11)$$

see examples in [29]. On the other hand, if the maximum set  $M(\bar{x}(t), p(t))$  in (1.8) is a *singleton* for a.e.  $t \in [0, \bar{T}]$  (in particular, when the sets  $F(x)$  are *strictly convex* along  $\bar{x}(\cdot)$ ), then (1.10) is reduced to the following *refined* form

$$\dot{p}(t) \in \text{co}\{u \mid (u, p(t)) \in N((\bar{x}(t), \dot{\bar{x}}(t)); \text{gph } F)\} \quad (1.12)$$

which obviously implies (1.11). Observe that the refined condition (1.12) requires less convexification: only to the components involving derivatives of the adjoint function instead of to all components at once. This makes (1.12) essentially stronger than (1.11) in certain situations; see Remark 4.5 in Section 4.

It has been recently proved by Loewen and Rockafellar [29] that the general case of fixed-time Mayer problems for convex differential inclusions can be actually reduced to the case when the sets  $F(x)$  are strictly convex along the optimal trajectory under consideration. In this way, using the Hamiltonian analysis, they establish the refined condition (1.12) with no single-valuedness assumption on the maximum set (1.8).

Note that in the convex-valued setting for  $F$ , the refined Euler-Lagrange inclusion (1.12) automatically implies the maximum condition (1.9); see Proposition 4.6 stated below. Therefore, it also implies the "fuzzy" inclusion (1.7) as well as the Hamiltonian inclusion (1.5), and thus it appears to be the strongest result in this direction for convex problems.

The principal necessary conditions for differential inclusions considered above are obtained under the convexity assumption on  $F(x)$  which is essentially used in their proofs. What happens if the sets  $F(x)$  are no longer assumed to be convex? First note that in this setting neither Clarke's form (1.11) of the Euler-Lagrange inclusion nor the refined form (1.12) implies the maximum condition (1.9). Could one ensure that these inclusions themselves are necessary for optimality in *nonconvex* problems?

The positive answer for the case of (1.11) can be found in Clarke's paper [10] under the *calmness* hypothesis imposed on a Mayer fixed-time problem. The latter hypothesis is a kind of regularity (constraint qualification) assumption for problems with endpoint constraints which ensures *normality* in transversality conditions; see Section 2 for more information. Recently Kaskosz and Lojasiewicz [27] have released the calmness assumption proving that Clarke's form of the Euler-Lagrange inclusion holds for any boundary trajectory in nonconvex

differential inclusions.

We have considered fixed-time optimization problems for nonconvex differential inclusions in the recent paper [43]. In that paper, we prove that the refined form (1.12) of the Euler-Lagrange inclusion is a necessary condition for optimality in Mayer problems as well as for boundary trajectories. Therein we also prove that the following analogue

$$\dot{p}(t) \in \text{co}\{u \mid (u, p(t)) \in \lambda \partial f(\bar{x}(t), \dot{\bar{x}}(t)) + N((\bar{x}(t), \dot{\bar{x}}(t)); \text{gph } F)\} \quad (1.13)$$

of the refined Euler-Lagrange inclusion (1.12) with a multiplier  $\lambda \geq 0$  appears to be a necessary optimality condition for fixed-time Bolza problems involving nonconvex differential inclusions. The latter result (in contrast to those for Mayer problems and boundary trajectories) is proved under a certain *relaxation stability* assumption; cf. Section 2. The core of our approach in [43] consists in using a method of *discrete (finite difference) approximations* together with appropriate tools of *nonsmooth analysis*.

The primal objective of the present paper is to extend the main constructions and results of [43] to *free-time Bolza problems for nonconvex differential inclusions*. To the best of our knowledge, such problems have never been considered in the literature and results for them cannot be directly derived from those for fixed-time problems.

Here we pursue a twofold goal. First, to develop a discrete approximation approach for free-time Bolza problems involving nonconvex differential inclusions. And second, to obtain necessary optimality conditions in the refined Euler-Lagrange form (1.13) with an additional characterization of the optimal time interval  $[0, \bar{T}]$ .

Discrete approximation techniques have been long time recognized as a powerful tool for studying and solving infinite dimensional variational problems. This approach goes back to Euler (1744) who used finite differences (broken lines) to prove the classical Euler-Lagrange equation in the calculus of variations as well as for proving the existence of solutions to differential equations.

In further applications, Euler's finite difference method and its modifications are mostly employed for computational purposes. There is a number of works devoted to numerical aspects of discrete approximations for optimal control and differential inclusion problems. We refer the reader to the recent papers of Dontchev and Lempio [19] and Polak [45] containing

surveys and new developments in this area; see also Dontchev's contribution to this volume [17].

Besides a numerical analysis of variational problems, consistent discrete approximations provide a possibility to obtain qualitative (theoretical) results for infinite dimensional problems by passing to the limit in corresponding results for their finite dimensional approximations. For instance, in this way one can derive necessary optimality conditions for variational problems using those in finite dimensional optimization. Such an approach may be successful if it is possible

- 1) to construct "right" discrete approximations with appropriate convergence properties;
- 2) to derive "robust" necessary conditions in the finite dimensional optimization problems obtained; and then
- 3) to justify the convergence procedure in those optimality conditions.

Some implementations of this approach applied to optimal control and differential inclusion problems can be found in Mordukhovich [33, 35, 36, 39, 43], Pshenichnyi [47], and Smirnov [56]. Observe that finite dimensional approximations for such problems always contain many equality and/or geometric constraints arising from finite difference replacements of differential relations.

Due to the natural presence of many geometric constraints, the finite dimensional problems obtained in this way appear to be objects of nonsmooth analysis and optimization even for the case of smooth functional data in the original models. Moreover, to achieve the purposes 2) and 3) stated above, they require to use only generalized differential constructions with special properties (or postulate such properties as in [47]). One can see that our constructions in Section 4 just fit all the requirements. At the same time, other widely spread constructions in nonsmooth analysis cannot be employed without some restrictive assumptions; see, in particular, discussions in Remark 4.9.

Previous efforts in using discrete approximations to obtain necessary optimality conditions for differential inclusions were mostly concerned with fixed-time Mayer problems under the convexity assumption on  $F(x)$ . In [39], we consider a free-time problem of Mayer with convexity and formulate necessary optimality conditions grouped around the "fuzzy" Euler-Lagrange inclusion (1.7). The recent paper [43] employs discrete approximations to prove



the refined Euler-Lagrange inclusion (1.13) for fixed-time Bolza problems in the nonconvex setting.

In the present paper we develop a discrete approximation procedure for free-time Bolza problems involving nonconvex differential inclusions. This procedure and the results obtained have some essential distinctions from the previous considerations.

To perform finite difference approximations for differential inclusion problems with free time, we use the simplest uniform Euler scheme but on a *varying time interval*. This implies that for each step of approximation, a *discrete grid* (stepsize) is variable and becomes a *subject to optimization*. The latter requires *more regularity* with respect to the time variable in order to prove necessary conditions involving optimal time.

In this way, we construct a sequence of discrete approximation problems  $(P_K)$  with a varying grid whose optimal solutions *strongly* converge in  $W^{1,2}$ -norm to a reference optimal solution  $\{\bar{x}(\cdot), \bar{T}\}$  for the original Bolza problem  $(P)$  with general endpoint constraints. Problems  $(P_K)$  with discrete-time dynamics can be reduced to special static problems of non-smooth optimization in finite dimensions with many equality, inequality, and geometric type constraints. Necessary optimality conditions for  $(P_K)$  are directly derived from a *generalized Lagrange multiplier rule* in nondifferentiable programming. Passing to the limit as  $K \rightarrow \infty$  and using convergence results together with appropriate tools of nonsmooth analysis, we get a set of necessary optimality conditions for the original nonconvex problem  $(P)$  with free time.

In these necessary conditions, the differential relation is expressed in the form of the *refined Euler-Lagrange inclusion* (1.13) for a.e.  $t \in [0, \bar{T}]$  with additional relations for  $(p(0), p(\bar{T}), \bar{T})$  in terms of (nonconvex) subdifferentials of the cost function  $\varphi$  in (1.1) and endpoint constraints. In the case when  $\varphi$  and endpoint constraints do not depend on  $T$ , the relations obtained imply that

$$\int_0^{\bar{T}} H(\bar{x}(t), \dot{\bar{x}}(t), p(t), \lambda) dt = 0 \quad (1.14)$$

in terms of the so-called *pseudo-Hamiltonian*

$$H(x, v, p, \lambda) := \langle p, v \rangle - \lambda f(x, v) \quad (1.15)$$

of the Bolza problem  $(P)$  calculated on the optimal solution  $\{\bar{x}(\cdot), \bar{T}\}$  and the corresponding adjoint pair  $\{p(\cdot), \lambda\}$ .

Note that in general nonconvex setting under consideration, the refined Euler-Lagrange inclusion (1.13) does not imply the maximum condition

$$H(\bar{x}(t), \dot{\bar{x}}(t), p(t), \lambda) = \mathcal{H}(\bar{x}(t), p(t), \lambda) := \max\{\langle p, v \rangle - \lambda f(x, v) \mid v \in F(x)\}$$

for a.e.  $t \in [0, \bar{T}]$  as well as the constancy condition

$$\mathcal{H}(\bar{x}(t), p(t), \lambda) \equiv \text{const on } [0, \bar{T}].$$

If the latter conditions hold, then (1.14) is reduced to (1.6) for the Hamiltonian  $\mathcal{H}$  of the Bolza problem. In general, (1.14) appears to be an independent integral condition in terms of the pseudo-Hamiltonian (1.15) over the optimal time interval  $[0, \bar{T}]$ .

The remainder of the paper is organized as follows. In Section 2 we formulate the problem and consider the property of relaxation stability employed for obtaining the principal results. Section 3 is devoted to constructing correct discrete approximations of the original problem and proving the strong convergence of optimal solutions. In Section 4 we review some concepts and results in nonsmooth analysis used in the paper. Section 5 is concerned with necessary optimality conditions for discrete approximation problems. In Section 6 we prove the main theorem about necessary optimality conditions for the differential inclusion problem under consideration. In the concluding Section 7 we discuss some open questions and further generalizations of the results obtained.

In this paper we basically use standard notation. Some special symbols are introduced and explained in Section 4. Throughout the paper, the set  $B$  stands for the unit closed ball of the space in question; the adjoint (transposed) matrix to  $A$  is denoted by  $A^*$ . Note that we consider all finite dimensional vectors to be vector-columns although they may be written as vector-rows for the purpose of convenience.

## 2. Problem Formulation and Relaxation

Let us consider the following problem  $(P)$  of dynamic optimization with a varying time

interval:

$$\text{minimize } J[x, T] := \varphi_0(x(0), x(T), T) + \int_0^T f(t, x(t), \dot{x}(t)) dt \quad (2.1)$$

over all arcs  $x(\cdot) \in W^{1,\infty}[0, T]$  satisfying the differential inclusion

$$\dot{x}(t) \in F(t, x(t)) \text{ a.e. } t \in [0, T] \quad (2.2)$$

and the general endpoint constraints

$$\varphi_i(x(0), x(T), T) \leq 0 \text{ for } i = 1, 2, \dots, q; \quad (2.3)$$

$$\varphi_i(x(0), x(T), T) = 0 \text{ for } i = q + 1, q + 2, \dots, q + r; \quad (2.4)$$

$$(x(0), x(T), T) \in \Omega \subset \mathbf{R}^{2n+1}. \quad (2.5)$$

Here  $F$  is a set-valued mapping (multifunction) from  $\mathbf{R}^{n+1}$  into  $\mathbf{R}^n$  and  $f, \varphi_i$  are real-valued functions defined on  $\mathbf{R}^{2n+1}$ . We call problem  $(P)$  the *Bolza problem for differential inclusion with free time*. If  $f \equiv 0$  in (2.1), then  $(P)$  is reduced to the corresponding *Mayer problem* for differential inclusions.

Any solution  $x(\cdot) \in W^{1,\infty}$  to (2.2) is called an *(original) trajectory* for the differential inclusion, and any trajectory for (2.2) satisfying constraints (2.3)–(2.5) is called a *feasible solution* to problem  $(P)$ . A feasible solution  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , minimizing the cost functional (2.1) over all other feasible solutions  $x(t)$ ,  $0 \leq t \leq T$ , (with different  $T$ ) is called an *(original) optimal solution* to  $(P)$ .

Along with problem  $(P)$ , we consider its *relaxation* (variational extension) which is defined by the following way going back to the classical works of Bogoljubov and Young in the 30s; cf. [4, 7, 21, 25, 57, 59]. Let

$$f_F(x, v, t) := f(t, x, v) + \delta(v, F(t, x)) \quad (2.6)$$

where  $\delta(v, \Lambda) = 0$  if  $v \in \Lambda$  and  $\delta(v, \Lambda) = \infty$  if  $v \notin \Lambda$  (the indicator function). Denote by  $\hat{f}_F(t, x, v)$  the *converification* (the biconjugate function) for  $f_F$  in the  $v$  variable, i.e., the

largest convex function majorized by  $f_F(t, x, \cdot)$  for each  $x$  and  $t$ . The *relaxed problem* (R) is defined as follows:

$$\text{minimize } \hat{J}[x, T] := \varphi_0(x(0), x(T), T) + \int_0^T \hat{f}_F(t, x(t), \dot{x}(t)) dt \quad (2.7)$$

over all arcs  $x(\cdot) \in W^{1,\infty}[0, T]$  subject to constraints (2.3)–(2.5). Observe that if  $\hat{J}[x, T] < \infty$ , then  $x(\cdot)$  satisfies the convexified differential inclusion

$$\dot{x}(t) \in \text{co } F(t, x(t)) \text{ a.e. } t \in [0, T]. \quad (2.8)$$

Any trajectory for (2.8) is called an *relaxed trajectory* for (2.2) in contrast to original trajectories satisfying (2.2). It is well known (cf., e.g., [1, 7, 21, 25]) that under natural assumptions involving the Lipschitz continuity of  $F$  in  $x$ , the following *approximation property* holds:

*Every relaxed trajectory  $x(t)$ ,  $0 \leq t \leq T$ , can be uniformly in  $[0, T]$  approximated by original trajectories  $x_k(t)$ ,  $0 \leq t \leq T$ , starting from the same initial point (but may be not satisfied endpoint constraints) such that*

$$\liminf_{k \rightarrow \infty} \int_0^T f(t, x_k(t), \dot{x}_k(t)) dt \leq \int_0^T \hat{f}_F(t, x(t), \dot{x}(t)) dt. \quad (2.9)$$

Probably the first result in this vein has been obtained by Bogoljubov [4] who proved such a property for the classical problem of the calculus of variations (no differential inclusions). The approximation property is related to the so-called "hidden convexity" of continuous-time systems under consideration which is reflected (from an abstract viewpoint) by the celebrated Lyapounov theorem about the range convexity of nonatomic vector measures [30] (the same as the convexity of Aumann's multivalued integral; see, e.g., [11, 26]).

The discussions above make natural the following definition.

2.1. DEFINITION. Denote by  $\inf(P)$  and  $\inf(R)$  the infima of the cost functionals in problems (P) and (R) respectively. Then one says that (P) possesses the property of *relaxation stability* (or (P) is *stable with respect to relaxation*) if

$$\inf(P) = \inf(R). \quad (2.10)$$

Note that by virtue of the convexity with respect to velocities in the relaxed problem, the infimum in (R) is attained under well known growth conditions of the Tonelli type; see,

e.g., [12, 26]. This means the existence of optimal solutions to the relaxed problem which is not the case for the original problem  $(P)$ . On the other hand, the approximation property allows us to obtain a minimizing sequence of original trajectories which, however, may not exactly satisfy the boundary conditions.

Obviously, we always have  $\inf(R) \leq \inf(P)$ . To establish the relaxation stability, one should prove the opposite inequality which is somehow connected with the approximation property. Let us present a result in this direction going back to the classical Bogoljubov theorem.

**2.2. PROPOSITION.** *Let  $F \equiv 0$  in  $(P)$ , and let the integrand  $f$  be continuous. Then  $(P)$  is stable with respect to relaxation and*

$$f(t, \bar{x}(t), \dot{\bar{x}}(t)) = \hat{f}(t, \bar{x}(t), \dot{\bar{x}}(t)) \text{ a.e. } t \in [0, \bar{T}] \quad (2.11)$$

for any optimal solution  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , to  $(P)$ .

*Proof.* Assume that  $\inf(R) < \inf(P)$ . Then there exist  $T > 0$  and a function  $x(\cdot) \in W^{1,\infty}[0, T]$  which satisfy all the constraints (2.3)–(2.5) and

$$\varphi_0(x(0), x(T), T) + \int_0^T \hat{f}(t, x(t), \dot{x}(t)) dt < \inf(P). \quad (2.12)$$

According to the version of Bogoljubov's theorem in [26, Section 9.2.4], for this  $x(\cdot)$  we can find a sequence of  $x_k(\cdot) \in W^{1,\infty}[0, T]$  such that  $x_k(0) = x(0)$ ,  $x_k(T) = x(T)$ ,  $x_k(\cdot)$  converges to  $x(\cdot)$  uniformly in  $[0, T]$ , and (2.9) holds with  $f_F = f$ . So, all  $x_k(t)$ ,  $0 \leq t \leq T$ , appear to be feasible solutions to problem  $(P)$  in (2.1), (2.3)–(2.5) with

$$\liminf_{k \rightarrow \infty} I[x_k, T] < \inf(P)$$

due to (2.9) and (2.12). This contradiction proves the relaxation stability of  $(P)$ .

If  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , is an optimal solution to  $(P)$ , then by virtue of (2.10) it solves the relaxed problem  $(R)$  as well. This ensures

$$\int_0^{\bar{T}} [f(t, \bar{x}(t), \dot{\bar{x}}(t)) - \hat{f}(t, \bar{x}(t), \dot{\bar{x}}(t))] dt = 0. \quad (2.13)$$

From the definition one always has

$$f(t, \bar{x}(t), \dot{\bar{x}}(t)) - \hat{f}(t, \bar{x}(t), \dot{\bar{x}}(t)) \geq 0 \text{ a.e. } t \in [0, \bar{T}].$$

Therefore, (2.13) is equivalent to (2.11).  $\square$

2.3. *Remark.* Proposition 2.2 implies the corresponding result of Clarke [7, Theorem 1] where  $f$  is assumed to be Lipschitz continuous in  $(x, v)$ . As it was observed in [7], (2.11) is the essence of the necessary condition of Weierstrass in the calculus of variations.

It follows directly from the approximation property stated above that the relaxation stability is inherent in any problem  $(P)$  involving (Lipschitz) differential inclusions with endpoint constraints at either  $t = 0$  or  $t = T$  (one of the ends is *free*). In general, the relaxation stability is clearly related to a kind of the value stability (regularity) of  $(P)$  with respect to perturbations of endpoint constraints. Such a regularity condition has been developed by Clarke and Rockafellar under the name of *calmness*; see [6–11, 50] and references therein. As it has been proved in Clarke [7, 9], the calmness property implies the relaxation stability and actually allows us to reduce problems involving differential inclusions to the classical ones as in Proposition 2.2.

Moreover, the calmness property is fulfilled for *most* endpoint constraints (at least of inequality type; cf. [6, 8]) and shows that the relaxation stability may fail only for ill-posed problems where small perturbations of endpoint constraints produce proportionally unbounded variations of the minimum (value function).

Note also that according to Clarke [9, 10], the calmness hypothesis implies that corresponding necessary optimality conditions can be taken *normal*. A general result that *normality implies relaxation stability* for optimal control systems has been obtained by Warga [57, 58].

For special classes of problems  $(P)$  with arbitrary endpoint constraints, the relaxation stability holds with *no calmness or normality* assumptions. In particular, let differential inclusion (2.2) be represented in the *linear form*:

$$\dot{x}(t) \in F_1(t)x(t) + F_2(t)$$

where the multifunctions  $F_1$  is convex-valued while  $F_2$  is not. If, in addition, the function  $f$  in (2.1) is convex in  $v$ , then any of such problems possesses the property of relaxation stability. This can be proved by using the Lyapounov-Aumann theorem about the convexity of set-valued integrals; cf. the arguments in Mordukhovich [35, Theorem 19.7]. Similarly,

the relaxation stability holds for general problems ( $P$ ) involving *one-dimensional* differential inclusions; see Remark 19.2 in [35].

In the subsequent sections of the paper, we shall use the property of relaxation stability to establish convergence results for discrete approximations and to obtain necessary optimality conditions for differential inclusions.

### 3. Finite Difference Approximations

This section is devoted to constructing finite difference (discrete) approximations of the original problem of Bolza for differential inclusions with free time and to establishing the principal theorem about the strong convergence of optimal solutions to discrete approximations in the  $W^{1,2}$ -norm.

In what follows we use the simplest uniform *Euler scheme*

$$\dot{x}(t) \approx \frac{x(t+h) - x(t)}{h}$$

for the replacement of the derivative in (2.2). Note that most of the qualitative results obtained below hold for many other first-order and higher-order finite difference approximations.

For any positive integer  $K = 1, 2, \dots$ , we consider a real number  $T_K$  approximating  $T$  and the uniform grid

$$t_0 = 0, \quad t_{j+1} = t_j + h_K \quad \text{for } j = 0, 1, \dots, K-1$$

with the stepsize  $h_K = T_K/K$  (so  $t_K = T_K$ ). We define a *discrete approximation inclusion* as follows

$$x_K(t_{j+1}) \in x_K(t_j) + h_K F(t_j, x_K(t_j)) \quad \text{for } j = 0, 1, \dots, K-1 \quad (3.1)$$

with some initial state  $x_K(0) = x_{0K}$ . A collection of vectors  $\{x_{jK} := x_K(t_j) \mid j = 0, \dots, K\}$  satisfying (3.1) is called a *discrete trajectory* for (3.1). The corresponding collection  $\{v_{jK} := (x_K(t_{j+1}) - x_K(t_j))/h_K \mid j = 0, \dots, K-1\}$  is called a *discrete velocity*.

We also consider piecewise-constant extensions of discrete velocities

$$v_K(t) := \frac{x_K(t_{j+1}) - x_K(t_j)}{h_K} \text{ for } t \in [t_j, t_{j+1}) \text{ \& } j = 0, \dots, K-1 \quad (3.2)$$

and the corresponding piecewise-linear extensions of discrete trajectories

$$x_K(t) := x_{0K} + \int_0^t v_K(s) ds \text{ for } t \in [0, T_K] \quad (3.3)$$

to the continuous time interval  $[0, T_K]$ . We shall call (3.2) and (3.3), respectively, *extended discrete velocities* and *extended discrete trajectories* for (3.1). It follows from (3.3) that

$$\dot{x}_K(t) = v_K(t) \text{ a.e. } t \in [0, T_K]$$

for extended discrete trajectories and velocities. So, given an extended discrete trajectory  $x_K(\cdot)$ , one can always identify the corresponding discrete velocity with  $\dot{x}_K(\cdot)$ .

Now let us fix arbitrary original trajectory  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , for the differential inclusion (2.2) and formulate an important result about its strong approximation by discrete trajectories. For this purposes, we assume that the multifunction  $F$  is bounded and locally Lipschitzian in  $x$  around  $\bar{x}(\cdot)$  and it is Hausdorff continuous in  $t$  a.e. on  $[0, \bar{T}]$ . In the sequel we actually need these assumptions and the result only for the case when  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , is an optimal solution to the initial problem (P).

More precisely, we impose the following hypotheses:

(H1) There are an open set  $U \subset \mathbf{R}^n$  and positive numbers  $m_F, l_F$  such that  $\bar{x}(t) \in U$  for any  $t \in [0, \bar{T}]$ , the sets  $F(t, x)$  are closed for all  $(t, x) \in [0, \bar{T}] \times U$ , and

$$F(t, x) \subset m_F B \quad \forall (t, x) \in [0, \bar{T}] \times U, \quad (3.4)$$

$$F(t, x_1) \subset F(t, x_2) + l_F |x_1 - x_2| B \quad \forall x_1, x_2 \in U, \quad t \in [0, \bar{T}]. \quad (3.5)$$

(H2) The multifunction  $F(\cdot, x)$  is Hausdorff continuous for a.e.  $t \in [0, \bar{T}]$  uniformly in  $x \in U$ .

Following Dontchev and Farkhi [18], we consider the so-called *averaged modulus of continuity* for the multifunction  $F(t, x)$  in  $t \in [0, \bar{T}]$  when  $x \in U$ . This modulus  $\tau(F; h)$  depending on the parameter  $h > 0$  is defined as

$$\tau(F; h) := \int_0^{\bar{T}} \sigma(F; t, h) dt \quad (3.6)$$



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

$$\omega(F; t, x, h) := \sup\{\text{haus}(F(t', x), F(t'', x)) \mid t', t'' \in [t - h/2, t + h/2] \cap [0, \bar{T}]\},$$

$\text{haus}(\cdot, \cdot)$  is the Hausdorff distance between compact sets.

It is proved in [18] that *if  $F(\cdot, x)$  is Hausdorff continuous for a.e.  $t \in [0, \bar{T}]$  uniformly in  $x \in U$ , then  $\tau(F; h) \rightarrow 0$  as  $h \rightarrow 0$ . Moreover,  $\tau(F; h) = O(h)$  if  $F(\cdot, x)$  has a bounded variation [18] uniformly in  $x \in U$  (in particular, if  $F$  is Lipschitz continuous in  $t$  with a uniform Lipschitz constant).*

Note that in the case of single-valued bounded functions  $f(t)$  not depending on  $x$ , the construction (3.6) has been originally developed in Sendov and Popov [55] under the name of "averaged modulus of smoothness". It has been proved in [55] that  $\tau(f; h) \rightarrow 0$  as  $h \rightarrow 0$  if and only if  $f$  is Riemann integrable on  $[0, \bar{T}]$ , i.e.,  $f$  is continuous for a.e.  $t \in [0, \bar{T}]$ . If  $f$  is of bounded variation on  $[0, \bar{T}]$ , then  $\tau(f; h) = O(h)$ . In this paper we shall use the name "averaged modulus of continuity" for both single-valued and multi-valued cases.

Now we formulate an auxiliary approximation result which is of independent interest for qualitative and numerical aspects of discrete approximations.

**3.1. LEMMA.** *Let  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , be an original trajectory for the differential inclusion (2.2) under hypotheses (H1) and (H2). Then there exists a sequence of solutions  $\{z_N(t_j) \mid j = 0, \dots, K\}$  to discrete inclusions (3.1) with  $T_K \equiv \bar{T}$  such that  $z_K(0) = \bar{x}(0)$  for any  $K = 1, 2, \dots$ , and the extended discrete trajectories  $z_K(t)$ ,  $0 \leq t \leq \bar{T}$ , converge to  $\bar{x}(\cdot)$  as  $K \rightarrow \infty$  in the norm topology of  $W^{1,2}[0, \bar{T}]$ .*

*Proof.* The complete proof of this result can be found in Mordukhovich [43]. The main idea is related to the so-called *proximal algorithm* to construct discrete trajectories for (3.1) by using projections of the derivative  $\dot{\bar{x}}(t)$  on the admissible velocity sets  $F(t_j, z_K(t_j))$ ; cf. [35, 36, 56]. In this way, we establish the *strong*  $L^2[0, \bar{T}]$ -convergence of the extended discrete velocities  $\dot{z}_K(\cdot)$  to  $\dot{\bar{x}}(\cdot)$  with effective *error estimates*. The latter estimates involve the boundedness and Lipschitz constants  $m_F$  and  $l_F$  in (3.4), (3.5) as well as the averaged modulus of continuity (3.6).  $\square$

Now let  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , be a given optimal solution to the original problem ( $P$ ) for the differential inclusion (3.1) satisfying (H1) and (H2) around  $\bar{x}(\cdot)$ . Because of  $U$  in (H1) is an

open neighborhood of  $\bar{x}(t)$  for all  $t \in [0, \bar{T}]$ , one can find a number  $\epsilon > 0$  such that

$$B_\epsilon(\bar{x}(t)) \subset U \quad \forall t \in [0, \bar{T}]. \quad (3.7)$$

Using Lemma 3.1, we construct a sequence of discrete approximation problems for finite difference inclusions whose optimal solutions strongly  $W^{1,2}$ -converge to the given to the given trajectory  $\bar{x}(\cdot)$ .

For any  $K = 1, 2, \dots$ , we consider the numbers

$$\alpha_{iK} := \varphi_i(\bar{x}(0), z_K(\bar{T}), \bar{T}) - \varphi_i(\bar{x}(0), \bar{x}(\bar{T}), \bar{T}) \quad \text{for } i = 1, \dots, q; \quad (3.8)$$

$$\beta_{iK} := |\varphi_i(\bar{x}(0), z_K(\bar{T}), \bar{T})| \quad \text{for } i = q + 1, \dots, q + r; \quad (3.9)$$

$$\gamma_K := |\bar{x}(\bar{T}) - z_K(\bar{T})| \quad (3.10)$$

where  $z_K(t)$ ,  $0 \leq t \leq \bar{T}$ , is the extended discrete trajectory from Lemma 3.1.

According to this lemma,  $\gamma_K \rightarrow 0$  as  $K \rightarrow \infty$ , and also  $\alpha_{iK}, \beta_{iK} \rightarrow 0$  as  $K \rightarrow \infty$  when the functions  $\varphi_i$  are continuous in the second variable at the point  $(\bar{x}(0), \bar{x}(\bar{T}), \bar{T})$ . It follows directly from (3.8)–(3.10) that

$$|\alpha_{iK}| \leq l_i \gamma_K \quad \text{for } i = 1, \dots, q \quad \& \quad \beta_{iK} \leq l_i \gamma_K \quad \text{for } i = q + 1, \dots, q + r$$

if the corresponding functions  $\varphi_i(\bar{x}(0), \cdot, \bar{T})$  are Lipschitz continuous around  $\bar{x}(\bar{T})$  with constants  $l_i$ . Note that the number  $\gamma_K$  in (3.10) can be effectively estimated in terms of the initial data of the problem; see [43].

Now for each  $K = 1, 2, \dots$ , we define the *discrete approximation problem* ( $P_K$ ) as follows:

$$\text{minimize } J_K[x_K, T_K] := \varphi_0(x_K(0), x_K(T_K), T_K) + |x_K(0) - \bar{x}(0)|^2 + (T_K - \bar{T})^2 + \quad (3.11)$$

$$h_K \sum_{j=0}^{K-1} f(t_j, x_K(t_j), \frac{x_K(t_{j+1}) - x_K(t_j)}{h_K}) + \sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} |\frac{x_K(t_{j+1}) - x_K(t_j)}{h_K} - \dot{\bar{x}}(t)|^2 dt$$

over trajectories  $\{x_K(t_j) \mid j = 0, \dots, K\}$  for the finite difference inclusion (3.1) subject to the constraints

$$\varphi_i(x_K(0), x_K(T_K), T_K) \leq \alpha_{iK} \quad \text{for } i = 1, \dots, q; \quad (3.12)$$

$$-\beta_{iK} \leq \varphi_i(x_K(0), x_K(T_K), T_K) \leq \beta_{iK} \quad \text{for } i = q+1, \dots, q+r; \quad (3.13)$$

$$(x_K(0), x_K(T_K), T_K) \in \Omega_K := \Omega + \gamma_K B; \quad (3.14)$$

$$x_K(t_j) \in B_\epsilon(\bar{x}(t_j)) \quad \forall j = 0, \dots, K; \quad T_K \leq \bar{T} + \epsilon \quad (3.15)$$

where  $\epsilon$  satisfies (3.7) and  $\epsilon$  is a given (arbitrarily small) positive number.

Let us emphasize that in each problem  $(P_K)$  the final time  $T_K$  and the discretization step  $h_K$  are *variable* for any fixed  $K = 1, 2, \dots$

We assume that the multifunction  $F$  satisfies hypotheses (H1) and (H2) along the given optimal solution  $\bar{x}(\cdot)$  for  $t \in [0, \bar{T} + \epsilon]$ . In addition to this, we impose the following hypotheses on  $f$ ,  $\varphi_i$ , and  $\Omega$ :

(H3)  $f(\cdot, x, v)$  is continuous for a.e.  $t \in [0, \bar{T} + \epsilon]$  and bounded uniformly in  $(x, v) \in U \times (m_F B)$ .

(H4) There exists  $\nu > 0$  such that the function  $f(t, \cdot, \cdot)$  is continuous on the set

$$A_\nu(t) := \{(x, v) \in U \times (m_F + \nu)B \mid v \in F(t', x) \text{ for some } t' \in (t - \nu, t)\}$$

uniformly in  $t \in [0, \bar{T} + \epsilon]$ .

(H5) The functions  $\varphi_0, \varphi_1, \dots, \varphi_q$  are lower semicontinuous on  $U \times U \times [0, \bar{T} + \epsilon]$  being continuous in the second variable at  $(\bar{x}(0), \bar{x}(\bar{T}), \bar{T})$ .

(H6) The functions  $\varphi_{q+1}, \dots, \varphi_{q+r}$  are continuous on  $U \times U \times [0, \bar{T} + \epsilon]$ .

(H7) The set  $\Omega$  is closed around  $(\bar{x}(0), \bar{x}(\bar{T}), \bar{T})$ .

**3.2. THEOREM.** *Let  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , be an original optimal solution to problem  $(P)$  which possesses the property of relaxation stability. Assume that hypotheses (H1)–(H7) hold.*

*Then:*

(i) *The discrete approximation problems  $(P_K)$  admit optimal solutions  $\{\bar{x}_K(t_j) \mid j = 0, \dots, K\}$  for all  $K$  large enough.*

(ii) *For any sequence of extended optimal trajectories  $\{\bar{x}_K(t), 0 \leq t \leq \bar{T}_K\}$  in  $(P_K)$  one has*

$$\bar{T}_K \rightarrow \bar{T} \text{ as } K \rightarrow \infty, \quad (3.16)$$

$$\max\{|\bar{x}_K(t) - \bar{x}(t)| : t \in \bar{T}_K\} \rightarrow 0 \text{ as } K \rightarrow \infty, \quad (3.17)$$

$$\int_0^{\bar{T}_K} |\dot{\bar{x}}_K(t) - \dot{\bar{x}}(t)|^2 dt \rightarrow 0 \text{ as } K \rightarrow \infty. \quad (3.18)$$

*Proof.* First let us show that the discrete approximation problem  $(P_K)$  has a feasible solution for any  $K$  large enough. In fact, we check that the trajectory  $\{z_K(t_j) | j = 0, \dots, K\}$  constructed in Lemma 3.1 satisfies all constraints (3.12)–(3.15). For the case of (3.12)–(3.14), it immediately follows from the definitions of the perturbations  $\alpha_{iK}$ ,  $\beta_{iK}$ , and  $\gamma_K$  in (3.8)–(3.10).

According to Lemma 3.1, the extended trajectories  $z_K(t)$ ,  $0 \leq t \leq \bar{T}$ , converge to  $\bar{x}(t)$  uniformly in  $[0, \bar{T}]$ . Now taking any  $\epsilon$  satisfying (3.7), one can find a natural number  $\bar{K}$  such that

$$|z_K(t) - \bar{x}(t)| \leq \epsilon \quad \forall K \geq \bar{K}.$$

Therefore, we ensure (3.15) for  $z_K(\cdot)$  when  $K \geq \bar{K}$ . Now the existence of optimal solutions to  $(P_K)$  for  $K \geq \bar{K}$  follows directly from the classical Weierstrass theorem due to the compactness and continuity assumptions made. This proves assertion (i).

To prove (ii), let us start with establishing the inequality

$$\limsup_{K \rightarrow \infty} J_K[\bar{x}_K, \bar{T}_K] \leq J[\bar{x}, \bar{T}] \quad (3.19)$$

for any sequence of optimal solutions to  $(P_K)$ . If (3.19) is not true, then there is a sequence  $\mathcal{N}$  of natural numbers  $K \rightarrow \infty$  such that

$$\varphi_0(\bar{x}(0), \bar{x}(\bar{T}), \bar{T}) + \int_0^{\bar{T}} f(t, \bar{x}(t), \dot{\bar{x}}(t)) dt < J_K[\bar{x}_K, \bar{T}_K] \quad \forall K \in \mathcal{N}. \quad (3.20)$$

To get a contradiction from (3.20), let us again use Lemma 3.1 and consider an approximating sequence of the discrete trajectories  $\{z_K(t_j) | j = 0, \dots, K; T_K = \bar{T}\}$  which are proved to be feasible for  $(P_K)$ . By virtue of  $z_K(0) = \bar{x}(0)$  and the continuity of  $\varphi_0$  in the second variable, one has

$$\varphi_0(z_K(0), z_K(T_K), T_K) \rightarrow \varphi_0(\bar{x}(0), \bar{x}(\bar{T}), \bar{T}) \text{ as } K \rightarrow \infty.$$

Observe that in the cost functional expression (3.11) for  $z_K(\cdot)$ , the second and third terms vanish. For the last term therein we get

$$\sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} \left| \frac{z_K(t_{j+1}) - z_K(t_j)}{h_K} - \bar{x}(t) \right|^2 dt = \int_0^{\bar{T}} |\dot{z}_K(t) - \bar{x}(t)|^2 dt \rightarrow 0 \text{ as } K \rightarrow \infty$$

by virtue of the extension rules (3.2), (3.3) and the convergence result in Lemma 3.1.

It remains to estimate the fourth term in (3.11). Let us prove that

$$\eta_K := h_K \sum_{j=0}^{K-1} f(t_j, z_K(t_j), \frac{z_K(t_{j+1}) - z_K(t_j)}{h_K}) \rightarrow \int_0^{\bar{T}} f(t, \bar{x}(t), \dot{\bar{x}}(t)) dt \text{ as } K \rightarrow \infty$$

under assumptions (H1)–(H4). Note that (H3) implies that  $\tau(f; h_K) \rightarrow 0$  as  $K \rightarrow \infty$  for the averaged modulus of continuity (3.6).

In what follows we use the sign " $\sim$ " for expressions which are equivalent as  $K \rightarrow \infty$ . Due to (3.2), (3.3), (3.6), and Lemma 3.1, one gets

$$\begin{aligned} \eta_K &= \sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} f(t, z_K(t_j), \dot{z}_K(t)) dt \sim \sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} f(t, z_K(t_j), \dot{z}_K(t)) dt + \tau(f; h_K) \sim \\ &\sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} f(t, \bar{x}(t), \dot{z}_K(t)) dt = \int_0^{\bar{T}} f(t, \bar{x}(t), \dot{z}_K(t)) dt \sim \int_0^{\bar{T}} f(t, \bar{x}(t), \dot{\bar{x}}(t)) dt. \end{aligned}$$

To justify the last conclusion, we observe that the strong  $L^2[0, \bar{T}]$ -convergence of  $z_K(\cdot) \rightarrow \dot{\bar{x}}(\cdot)$  in Lemma 3.1 implies the existence of a subsequence  $\{\dot{z}_K(t)\}$  converging to  $\dot{\bar{x}}(t)$  for a.e.  $t \in [0, \bar{T}]$ . Now one can employ the classical Lebesgue theorem about the limit under the integral sign.

Thus we have proved that

$$J_K[z_K, \bar{T}] \rightarrow J[\bar{x}, \bar{T}] \text{ as } K \rightarrow \infty.$$

By virtue of (3.20) and the feasibility of  $z_K(\cdot)$  for  $(P_K)$ , this contradicts the optimality of  $\bar{x}_K(\cdot)$  for the discrete approximations under consideration. Therefore, we ensure (3.19).

In the discussions above we have heavily used Lemma 3.1 but have not yet used the property of relaxation stability for the original problem  $(P)$ . Now let us prove that the relaxation stability property together with result (3.19) imply the desirable convergences (3.16)–(3.18). We are going to show that

$$\lim_{K \rightarrow \infty} [\delta_K := |\bar{x}_K(0) - \bar{x}(0)|^2 + |\bar{T}_K - \bar{T}|^2 + \int_0^{\bar{T}_K} |\dot{\bar{x}}_K(t) - \dot{\bar{x}}(t)|^2 dt] = 0. \quad (3.21)$$

Suppose that (3.21) doesn't hold and consider any limiting point  $\delta > 0$  of the sequence  $\{\delta_K\}$  in (3.21). For the purpose of simplicity, we assume that  $\delta = \lim \delta_K$  for all  $K \rightarrow \infty$ .

Using the boundedness of  $\bar{T}_K$  by virtue of (3.7), one can find a real number  $\tilde{T} \leq \bar{T} + \varepsilon$  such that  $\bar{T}_K \rightarrow \tilde{T}$  as  $K \rightarrow \infty$  (here and on we take all natural  $K$  without loss of generality). Let us consider the extended discrete trajectories  $\bar{x}_K(t)$  on the time interval  $[0, \tilde{T}]$  defining  $\bar{x}_K(t) \equiv \bar{x}_K(T_K)$  for  $t \in (T_K, \tilde{T}]$  when  $T_K < \tilde{T}$ .

Taking into account the boundedness conditions (3.7) and (3.4) for  $t \in [0, \tilde{T}]$  and employing the classical compactness results, we claim the existence of an absolutely continuous function  $\tilde{x}(t)$  in  $[0, \tilde{T}]$  such that  $\bar{x}_K(\cdot) \rightarrow \tilde{x}(t)$  uniformly in  $[0, \tilde{T}]$  and  $\dot{\bar{x}}_K(\cdot) \rightarrow \dot{\tilde{x}}(\cdot)$  weakly in  $L^2[0, \tilde{T}]$  as  $K \rightarrow \infty$ . Due to assumptions (H5)–(H7), the limiting function  $\tilde{x}(t)$ ,  $0 \leq t \leq \tilde{T}$ , satisfies endpoint constraints (2.3)–(2.5). Now we study the limit of the cost functional  $J_K[\bar{x}_K, \bar{T}_K]$  in (3.11) as  $K \rightarrow \infty$ .

According to the well-known Masur theorem, there is a sequence of convex combinations of  $\dot{\bar{x}}_K(\cdot)$  which converges to  $\dot{\tilde{x}}(\cdot)$  in the norm topology of  $L^2[0, \tilde{T}]$ . Hence it contains a subsequence converging to  $\dot{\tilde{x}}(\cdot)$  for a.e.  $t \in [0, \tilde{T}]$ .

From here one can easily conclude that  $\tilde{x}(\cdot)$  satisfies the convexified differential inclusion (2.8). Moreover, taking into account that

$$h_K \sum_{j=0}^{K-1} f(t_j, \bar{x}_K(t_j), \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K}) \sim \int_0^{\tilde{T}} f(t, \bar{x}_K(t), \dot{\bar{x}}_K(t)) dt \text{ as } K \rightarrow \infty$$

and also the definition of the convexified function  $\hat{f}_F$  for (2.6), we get

$$\int_0^{\tilde{T}} \hat{f}_F(t, \tilde{x}(t), \dot{\tilde{x}}(t)) dt \leq \liminf_{K \rightarrow \infty} h_K \sum_{j=0}^{K-1} f(t_j, \bar{x}_K(t_j), \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K}). \quad (3.22)$$

Let us observe that

$$\sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} \left| \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K} - \dot{\tilde{x}}(t) \right|^2 dt = \int_0^{\tilde{T}_K} |\dot{\bar{x}}_K(t) - \dot{\tilde{x}}(t)|^2 dt \sim \int_0^{\tilde{T}} |\dot{\bar{x}}_K(t) - \dot{\tilde{x}}(t)|^2 dt \text{ as } K \rightarrow \infty.$$

By virtue of the convexity in  $v$  of the function  $g(v, t) := |v - \dot{\tilde{x}}(t)|^2$ , the integral functional

$$I[v] := \int_0^{\tilde{T}} |v(t) - \dot{\tilde{x}}(t)|^2 dt$$

is lower semicontinuous in the weak topology of  $L^2[0, \tilde{T}]$ . Therefore.

$$\int_0^{\tilde{T}} |\dot{\tilde{x}}(t) - \dot{\tilde{x}}(t)|^2 dt \leq \liminf_{K \rightarrow \infty} \sum_{j=0}^{K-1} \left| \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K} - \dot{\tilde{x}}(t) \right|^2 dt. \quad (3.23)$$

Now passing to the limit in (3.11) for  $J_K[\bar{x}_K, \bar{T}_K]$  and taking into account (3.22), (3.23), and the lower semicontinuity of  $\varphi_0$ , one has

$$\varphi_0(\tilde{x}(0), \tilde{x}(\tilde{T}), \tilde{T}) + \int_0^{\tilde{T}} \hat{f}_F(t, \tilde{x}(t), \dot{\tilde{x}}(t)) dt + \delta \leq \liminf_{K \rightarrow \infty} J_K[\bar{x}_K, \bar{T}_K]. \quad (3.24)$$

By virtue of (3.19) and our assumption about  $\delta > 0$ , (3.24) implies that

$$\hat{J}[\tilde{x}, \tilde{T}] < J[\bar{x}, \bar{T}].$$

But the latter is impossible because

$$J[\bar{x}, \bar{T}] = \inf(P) = \inf(R)$$

due to the relaxation stability property for  $(P)$ . Therefore,  $\delta = 0$  and we get relationship (3.21).

The relationship obtained directly implies (3.16), (3.18), and the convergence  $\bar{x}_K(0) \rightarrow \bar{x}(0)$  as  $K \rightarrow \infty$ . The latter and (3.18) provide (3.17). This ends the proof of the theorem.

□

3.3. *Remark.* Denote by  $\bar{J}_K$  the optimal value of the cost functional in the discrete approximation problem  $(P_K)$  for each  $K = 1, 2, \dots$ . We have actually proved in Theorem 3.2 that

$$\inf(R) \leq \liminf_{K \rightarrow \infty} \bar{J}_K \leq \limsup_{K \rightarrow \infty} \bar{J}_K \leq \inf(P) \quad (3.25)$$

with no assumption about the relaxation stability. Moreover, (3.25) implies that *the relaxation stability property for  $(P)$  is in fact equivalent to the value convergence  $\bar{J}_K \rightarrow \inf(P)$  of the discrete approximations under appropriate perturbations of constraints.* We refer to Mordukhovich [32, 35, 36] and the recent book of Dontchev and Zolezzi [20] for related properties and discussions in other problems of optimal control.

3.4. *Remark.* In the convergence results stated above, one may avoid the continuity hypothesis (H3) on  $f$  in  $t$  by changing the approximation

$$h_K \sum_{j=0}^{K-1} f(t_j, x_K(t_j), \frac{x_K(t_{j+1}) - x_K(t_j)}{h_K}) \quad \text{for} \quad \sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} f(t, x_K(t_j), \frac{x_K(t_{j+1}) - x_K(t_j)}{h_K}).$$

Indeed, we can handle the latter approximation in the same way as the last term in (3.11) under the measurability (summability) assumption  $f$  in  $t$ . On the other hand, one can change

$$\sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} \left| \frac{x_K(t_{j+1}) - x_K(t_j)}{h_K} - \dot{\bar{x}}(t) \right|^2 dt \quad \text{for} \quad h_K \sum_{j=0}^{K-1} \left| \frac{x_K(t_{j+1}) - x_K(t_j)}{h_K} - \dot{\bar{x}}(t_j) \right|^2$$

assuming that  $\dot{\bar{x}}(\cdot)$  is continuous for a.e.  $t \in [0, \bar{T}]$ . These two kinds of approximations are treated by different techniques from the viewpoint of necessary optimality conditions; cf. [43] and Section 5 below.

**3.5. Remark.** The convergence results obtained in this section allow us to make a bridge between variational problems for differential inclusions with free time and dynamic optimization problems in finite dimensions. The latter can be reduced to special finite dimensional problems of mathematical programming; see Section 5. In this paper we develop this procedure to obtain necessary optimality conditions for the original variational problem ( $P$ ) by passing to the limit in corresponding necessary conditions for finite dimensional optimization problems ( $P_K$ ).

Note that the convergence results in Theorem 3.2 ensure that optimal solutions to problems ( $P_K$ ) belong to the *interiority* of the constraints in (3.15). Therefore, from the viewpoint of necessary conditions, constraints (3.15) can be omitted without any loss of generality. *In what follows we shall always consider problems ( $P_K$ ) with no constraints (3.15).*

**3.6. Remark.** Problems ( $P_K$ ) and equivalent problems of mathematical programming always have *many geometric constraints* arising from the approximation of differential inclusions. Such problems appear to be objects of *nonsmooth analysis and optimization*.

For the variational analysis of these problems and then for passing to the limit in necessary optimality conditions as  $K \rightarrow \infty$ , we need to use generalized differential constructions with special properties which are the subject of the next section.

## 4. Generalized Differentiation

In this section we briefly review some constructions and results on the generalized differentiation of nonsmooth and set-valued mappings which are widely employed in the paper.



Most of these results with detailed proofs and discussions can be found in Mordukhovich [35, 40, 41]. We also refer the reader to Clarke [12], Ioffe [22–24], Loewen [28], Rockafellar [48, 52], and Rockafellar and Wets [53] for related and additional material.

Developing a geometric approach to the generalized differentiation, we begin with the definition of a normal cone to arbitrary sets in finite dimensions.

Let  $\Omega$  be a nonempty set in  $\mathbf{R}^n$ , and let

$$\Pi(x, \Omega) := \{\omega \in \text{cl } \Omega \mid |x - \omega| = \text{dist}(x, \Omega)\}$$

be the (multi-valued) *Euclidean projector* of  $x$  on the set  $\text{cl } \Omega$ . In the following definition, "cone" stands for the conic hull of a set and "Limsup" denotes the well-known *Kuratowski-Painlevé upper limit* for multifunctions, i.e., the collection of all limiting points in their values (see, e.g., [2, p. 41]).

4.1. DEFINITION. Given  $\bar{x} \in \text{cl } \Omega$ , the closed cone

$$N(\bar{x}; \Omega) := \text{Limsup}_{x \rightarrow \bar{x}}[\text{cone}(x - \Pi(x, \Omega))] \quad (4.1)$$

is called the *normal cone* to the set  $\Omega$  at the point  $\bar{x}$ . If  $\bar{x} \notin \text{cl } \Omega$ , we put  $N(\bar{x}; \Omega) = \emptyset$ .

This definition of the normal cone first appeared in Mordukhovich [31] motivated by applications to optimal control problems. The normal cone defined (or equivalent constructions) are widely used in the literature, sometimes under different names: the "approximate normal cone" [22], the "cone of the limiting proximal normals" [48], the "prenormal cone" [12], the "limiting normal cone" [28, 29, 54], etc.

If  $\Omega$  is convex, then (4.1) is reduced to the normal cone of convex analysis. In general, the normal cone (4.1) is frequently *nonconvex* and its convex closure coincides with the *Clarke normal cone*

$$N_C(\bar{x}; \Omega) = \text{cl co } N(\bar{x}; \Omega) \quad (4.2)$$

which is the dual (polar) construction to Clarke's tangent cone [11]. Furthermore, (4.1) always admits the representation

$$N(\bar{x}; \Omega) = \text{Limsup}_{x \rightarrow \bar{x}} \hat{N}(x; \Omega)$$

where the cone

$$\hat{N}(x; \Omega) := \{x^* \in \mathbf{R}^n \mid \limsup_{x' \in \Omega \rightarrow x} \frac{\langle x^*, x' - x \rangle}{|x' - x|} \leq 0\} \text{ for } x \in \text{cl } \Omega \quad (4.3)$$

appears to be *dual* to the well-known Bouligand *contingent cone*; see, e.g., [2, Chapter 2].

Observe that our basic normal cone in (4.1) is *not dual* to any tangent cone because it is not convex, except situations where all three normal cones (4.1)–(4.3) coincide at  $\bar{x}$ . (The latter fails in many important settings, e.g., for sets  $\Omega$  which can be locally represented as graphs of nonsmooth Lipschitz continuous functions; see Remark 4.5.)

Despite its nonconvexity, the normal cone (4.1) possesses many nice properties essential for applications. First, it is always *robust* with respect to perturbations of  $\bar{x}$ , i.e., the multifunction  $N(\cdot, \Omega)$  has closed graph. What is really surprising a priori, that this nonconvex normal cone and related differential objects enjoy *rich calculi* which is even better than ones for convex-valued counterparts; see below. The progress in this direction has been achieved by using a *variational approach* instead of convex analysis.

Now we consider generalized differentiation constructions for multifunctions and nonsmooth mappings induced by the normal cone (4.1) to their *graphs*. The following notion was first introduced in [33].

4.2. DEFINITION. Let  $F$  be a multifunction from  $\mathbf{R}^n$  into  $\mathbf{R}^m$  which graph

$$\text{gph } F := \{(x, y) \in \mathbf{R}^n \times \mathbf{R}^m \mid y \in F(x)\}$$

is nonempty, and let  $(\bar{x}, \bar{y}) \in \text{cl}(\text{gph } F)$ . The multifunction  $D^*F(\bar{x}, \bar{y})$  from  $\mathbf{R}^m$  into  $\mathbf{R}^n$  defined by

$$D^*F(\bar{x}, \bar{y})(y^*) := \{x^* \in \mathbf{R}^n \mid (x^*, -y^*) \in N((\bar{x}, \bar{y}); \text{gph } F)\} \quad (4.4)$$

is called the *coderivative* of  $F$  at  $(\bar{x}, \bar{y})$ . We put  $D^*F(\bar{x}, \bar{y})(y^*) = \emptyset$  if  $(\bar{x}, \bar{y}) \notin \text{cl}(\text{gph } F)$ . The symbol  $D^*F(\bar{x})$  is used in (4.4) when  $F$  is single-valued at  $\bar{x}$  and  $\bar{y} = F(\bar{x})$ .

One can see that the coderivative  $D^*F(\bar{x}, \bar{y})(\cdot)$  is a positive homogeneous multifunction with closed values. These values may be not convex by virtue of the nonconvexity of (4.1). Therefore, the coderivative (4.4) is *not dual* to any tangentially generated *derivative* of multifunctions (see, e.g., [2, Chapter 5]).

Let us present two useful representations of the coderivative (4.4) in the following classical settings.

4.3. PROPOSITION. *Let  $F$  be single-valued around  $\bar{x}$  and strictly differentiable at  $\bar{x}$  with the Jacobian  $\nabla F(\bar{x}) \in \mathbf{R}^{m \times n}$ , i.e.,*

$$\lim_{x, x' \rightarrow \bar{x}} \frac{F(x) - F(x') - \nabla F(\bar{x})(x - x')}{|x - x'|} = 0.$$

*Then one has*

$$D^*F(\bar{x})(y^*) = \{(\nabla F(\bar{x}))^*y^*\} \quad \forall y^* \in \mathbf{R}^m.$$

4.4. PROPOSITION. *Let  $F$  be a multifunction of convex graph. Then for any point  $(\bar{x}, \bar{y}) \in \text{cl}(\text{gph } F)$  and for any  $y^* \in \mathbf{R}^m$  one has*

$$D^*F(\bar{x}, \bar{y})(y^*) = \{x^* \in \mathbf{R}^n \mid \langle x^*, \bar{x} \rangle - \langle y^*, \bar{y} \rangle = \sup_{(x, y) \in \text{gph } F} [\langle x^*, x \rangle - \langle y^*, y \rangle]\}.$$

4.5. *Remark.* Note that in both cases considered above the coderivative (4.4) coincides with the *Clarke coderivative*  $D_C^*F(\bar{x}, \bar{y})(\cdot)$  which is obtained in the scheme (4.4) by replacing the normal cone (4.1) with its Clarke's counterpart (4.2). However, this is not longer true for a broad class of multifunctions whose graphs are nonsmooth *Lipschitzian manifolds* in the sense of Rockafellar [50], i.e., they are locally homeomorphic around  $(\bar{x}, \bar{y})$  to graphs of nonsmooth Lipschitz continuous functions. Besides locally Lipschitzian vector functions, this class includes maximal monotone relations, in particular, subdifferential operators for convex, concave, and saddle functions; see [50].

Indeed, for such multifunctions the coderivative (4.4) is always *strictly smaller* than its Clarke's counterpart; moreover, they may be distinguished *in dimensions*. (We refer the reader to [42, Section 3] for more details and related discussions.) That is why the refined form of the Euler-Lagrange inclusion in (1.12) involving the coderivative (4.4) is essentially stronger than Clarke's one (1.11) involving  $D_C^*F$ .

Let us consider some general properties of the coderivative (4.4) which are useful in the framework of this paper. First we note that it is a *robust* construction with respect

to perturbations of the data  $(\bar{x}, \bar{y}, y^*)$ . The next result [35, Theorem 3.1] shows that the considered Euler-Lagrange conditions for differential and discrete inclusions automatically imply the maximum (minimum) conditions in problems with convex velocities. In what follows, we use a conventional concept of lower (inner) semicontinuity for multifunctions; see, e.g., [2, p. 39].

4.6. PROPOSITION. *Let  $F$  be convex-valued around  $\bar{x}$  and lower semicontinuous at  $\bar{x}$ . Then one has*

$$[D^*F(\bar{x}, \bar{y})(y^*) \neq \emptyset] \implies [\langle y^*, \bar{y} \rangle = \min\{\langle y^*, y \rangle \mid y \in F(\bar{x})\}].$$

One of the most important advantages of the coderivative (4.4) in the general setting consists in effectively using this construction for *complete dual characterizations of Lipschitzian properties* of multifunctions and nonsmooth mappings. These results play a *crucial role* to justify the convergence of adjoint functions in necessary optimality conditions for discrete approximations; see the proof of Theorem 6.1 in Section 6.

Recall that the multifunction  $F$  is said to be *pseudo-Lipschitzian* around  $(\bar{x}, \bar{y}) \in \text{gph } F$  if there is a neighborhood  $U$  of  $\bar{x}$ , a neighborhood  $V$  of  $\bar{y}$ , and a constant  $l \geq 0$  such that

$$F(x') \cap V \subset F(x) + l|x' - x|B \quad \forall x, x' \in U. \quad (4.5)$$

This definition goes back to Aubin who imposed the additional condition  $F(x) \cap V \neq \emptyset$  for all  $x \in U$ ; see [2, Definition 1.4.5]. Rockafellar [49] obtained some characterizations of the pseudo-Lipschitzian property and established its interrelations with other Lipschitzian properties of multifunctions. Note that the pseudo-Lipschitzian property of  $F$  appears to be equivalent to such fundamental properties as *metric regularity* and *openness at linear rate* for the *inverse mapping*  $F^{-1}$ ; see Borwein and Zhuang [5], Penot [44], and Mordukhovich [40] for various modifications.

The next dual characterizations were obtained in Mordukhovich [38, Theorem 5.1] and [40, Theorem 5.7]. Moreover, therein one can find formulae for calculating the *exact bound* of Lipschitz moduli  $l$  in (4.5).

4.7. PROPOSITION. *Let  $F$  be a multifunction from  $\mathbf{R}^n$  into  $\mathbf{R}^m$  whose graph is closed*

around  $(\bar{x}, \bar{y})$ . Then the following are equivalent:

- (a)  $F$  is pseudo-Lipschitzian around  $(\bar{x}, \bar{y})$ ;
- (b) there exist a neighborhood  $U$  of  $\bar{x}$ , a neighborhood  $V$  of  $\bar{y}$ , and a constant  $l \geq 0$  such that

$$\sup\{|x^*| : x^* \in D^*F(x, y)(y^*)\} \leq l|y^*| \quad (4.6)$$

for any  $x \in U$ ,  $y \in F(x) \cap V$ , and  $y^* \in \mathbf{R}^m$ ;

- (c)  $D^*F(\bar{x}, \bar{y})(0) = \{0\}$ .

If  $F$  is locally bounded around  $\bar{x}$ , then its classical (Hausdorff) Lipschitzian behavior ( $V = \mathbf{R}^m$  in (4.5)) is equivalent to  $F$  being pseudo-Lipschitzian around  $(\bar{x}, \bar{y})$  for every  $\bar{y} \in F(\bar{x})$ ; see [49]. Therefore, we get dual criteria for the classical local Lipschitz continuity of multifunctions.

4.8. COROLLARY. *Let a closed-graph multifunction  $F$  be locally bounded around the point  $\bar{x}$  where  $F(\bar{x}) \neq \emptyset$ . Then the following are equivalent:*

- (a)  $F$  is locally Lipschitzian around  $\bar{x}$ ;
- (b) there are a neighborhood  $U$  of  $\bar{x}$  and a number  $l \geq 0$  such that estimate (4.6) holds for any  $x \in U$ ,  $y \in F(x)$ , and  $y^* \in \mathbf{R}^m$ ;
- (c)  $D^*F(\bar{x}, \bar{y})(0) = \{0\} \quad \forall \bar{y} \in F(\bar{x})$ .

4.9. *Remark.* If one replaces the coderivative (4.4) by its Clarke's counterpart in the "null-criteria" (c) of Propositions 4.7 and 4.8, then the conditions obtained are far removed from the necessity for  $F$  to possess the Lipschitzian properties. Indeed, they are *never fulfilled* even in the case of single-valued Lipschitz continuous functions which do not happen to be strictly differentiable at  $\bar{x}$  (and also in more general settings with nonsmooth Lipschitzian manifolds in Remark 4.5). In fact, the condition  $D_C^*F(\bar{x}, \bar{y})(0) = \{0\}$  implies a "strictly smooth" property of multifunctions whose graphs are Lipschitzian manifolds; see [42, 50]. So, one cannot ensure an analogue of the basic estimate (4.6) in terms of  $D_C^*$  for such nonsmooth multifunctions.

Now let us consider some *calculus rules* for the coderivative (4.4) and related differential constructions which are valid under natural assumptions and appear to be of great impor-

tance for applications. The proofs of the following and other calculus results based on an *extremal principle* for systems of sets can be found in Mordukhovich [41].

4.10. PROPOSITION. *Let  $F_1$  and  $F_2$  be closed-graph multifunctions from  $\mathbf{R}^n$  into  $\mathbf{R}^m$ , and let  $\bar{y} \in F_1(\bar{x}) + F_2(\bar{x})$ . Assume that the multifunction  $S$  from  $\mathbf{R}^{n+m}$  into  $\mathbf{R}^{2m}$  defined by*

$$S(x, y) := \{(y_1, y_2) \in \mathbf{R}^{2m} \mid y_1 \in F_1(x), y_2 \in F_2(x), y_1 + y_2 = y\}$$

*is locally bounded around  $(\bar{x}, \bar{y})$  and the qualification condition*

$$D^*F_1(\bar{x}, y_1)(0) \cap (-D^*F_2(\bar{x}, y_2)(0)) = \{0\} \quad \forall (y_1, y_2) \in S(\bar{x}, \bar{y}) \quad (4.7)$$

*is fulfilled. Then for any  $y^* \in \mathbf{R}^m$  one has*

$$D^*(F_1 + F_2)(\bar{x}, \bar{y})(y^*) \subset \bigcup_{(y_1, y_2) \in S(\bar{x}, \bar{y})} [D^*F_1(\bar{x}, y_1)(y^*) + D^*F_2(\bar{x}, y_2)(y^*)]$$

*where equality holds if either  $F_1$  or  $F_2$  is strictly differentiable at  $\bar{x}$ .*

Observe that according to Proposition 4.7, the qualification condition (4.7) is automatically fulfilled if for each  $(y_1, y_2) \in S(\bar{x}, \bar{y})$  either  $F_1$  is *pseudo-Lipschitzian* around  $(\bar{x}, y_1)$  or  $F_2$  is *pseudo-Lipschitzian* around  $(\bar{x}, y_2)$ .

The next calculus result is concerned with general *chain rules* for the composition

$$(G \circ F)(x) = G(F(x)) := \bigcup_{y \in F(x)} G(y)$$

of multifunctions  $F$  from  $\mathbf{R}^n$  into  $\mathbf{R}^m$  and  $G$  from  $\mathbf{R}^m$  into  $\mathbf{R}^q$ .

4.11. PROPOSITION. *Let  $F$  and  $G$  be closed-graph, and let  $\bar{z} \in (G \circ F)(\bar{x})$ . Assume that the multifunction*

$$(x, z) \rightarrow F(x) \cap G^{-1}(z) = \{y \in F(x) \mid z \in G(y)\}$$

*is locally bounded around  $(\bar{x}, \bar{z})$  and the qualification condition*

$$D^*G(y, \bar{z})(0) \cap \ker D^*F(\bar{x}, y) = \{0\} \quad \forall y \in F(\bar{x}) \cap G^{-1}(\bar{z}) \quad (4.8)$$

*is fulfilled. Then one has*

$$D^*(G \circ F)(\bar{x}, \bar{z}) \subset \bigcup_{y \in F(\bar{x}) \cap G^{-1}(\bar{z})} [D^*F(\bar{x}, y) \circ D^*G(y, \bar{z})]$$

where equality holds in each of the following cases:

(i)  $F$  is strictly differentiable at  $\bar{x}$  with the quadratic ( $n = m$ ) and nonsingular Jacobian matrix  $\nabla F(\bar{x})$ ;

(ii)  $F$  is single-valued and Lipschitz continuous around  $\bar{x}$  while  $G$  is strictly differentiable at  $\bar{y} = F(\bar{x})$ .

Note that according to criterion (c) in Proposition 4.7 and the inverse criterion for the (local) metric regularity in [40, Corollary 4.3], the qualification condition (4.8) is automatically fulfilled if for each  $y \in F(\bar{x}) \cap G^{-1}(\bar{z})$  either  $G$  is pseudo-Lipschitzian around  $(y, \bar{z})$  or  $F$  is metrically regular around  $(\bar{x}, y)$ .

These major calculus results implies some other calculus rules for the coderivatives of multifunctions and associated constructions of the first and second order subdifferentials for *extended-real-valued functions*  $f : \mathbf{R}^n \rightarrow \bar{\mathbf{R}} = [\infty, \infty]$ ; see [35, 41]. Now we consider some results for the (first order) subdifferentials of such functions important in what follows.

4.12. DEFINITION. Let  $|f(\bar{x})| < \infty$ , and let

$$E_f(x) := \{\mu \in \mathbf{R} \mid \mu \geq f(x)\}$$

be the epigraphical multifunction ( $\text{gph } E_f = \text{epi } f$ ) associated with  $f$ . The set

$$\partial f(\bar{x}) := D^* E_f(\bar{x}, f(\bar{x}))(1) = \{x^* \in \mathbf{R}^n \mid (x^*, -1) \in N((\bar{x}, f(\bar{x})); \text{epi } f)\} \quad (4.9)$$

is called the *subdifferential* of  $f$  at  $\bar{x}$ . We put  $\partial f(\bar{x}) = \emptyset$  if  $|f(\bar{x})| = \infty$ .

The subdifferential (4.9) has been studied in details in the book of Mordukhovich [35] and previous publications starting from [31]. There are various analytic representations of (4.9) in terms of other subdifferential mappings named Dini, Frechét, viscosity, proximal subdifferentials, etc. We refer the reader to [12, 16, 22, 28, 35, 52, 53] for more information about these and related questions. Therein construction (4.9) is often used in some equivalent forms under different names (e.g., the approximate subdifferential, the presubdifferential, the set of limiting proximal subgradients or basic subgradients).

Note that for functions  $f : \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$  *lower semicontinuous* around  $\bar{x}$ , the subdifferential (4.9) can be expressed directly in terms of the coderivative (4.4) for  $f$ :

$$\partial f(\bar{x}) = D^* f(\bar{x})(1),$$

i.e., (epi  $f$ ) in (4.9) can be replaced by (gph  $f$ ); see [35, Proposition 2.1].

On the other hand, if  $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$  is a *vector function Lipschitz continuous* around  $\bar{x}$ , then the coderivative of  $f$  can be expressed in terms of the subdifferential (4.9) for its *Lagrange scalarization*:

$$D^*f(\bar{x})(y^*) = \partial\langle y^*, f \rangle(\bar{x}) \quad \forall y^* \in \mathbf{R}^m$$

where  $\langle y^*, f \rangle(x) := \langle y^*, f(x) \rangle$ ; see proofs and comments in [22, Section 5] and [35, Section 3].

For the case of real-valued Lipschitz continuous functions we have the following well-known results (see, e.g., [12, Propositions 1.1 and 1.2] and [35, Theorem 2.1]).

4.13. PROPOSITION. *Let a real-valued function  $f$  be locally Lipschitzian around  $\bar{x}$  with modulus  $l$ . Then:*

- (i)  $\partial f(\bar{x}) \neq \emptyset$  &  $|x^*| \leq l \quad \forall x^* \in \partial f(\bar{x})$ ;
- (ii) *one has*

$$\partial_C f(\bar{x}) = \text{co } \partial f(\bar{x}) = \text{co} \left\{ \lim_{\nu \rightarrow \infty} \nabla f(x_\nu) \mid f \text{ is differentiable at } x_\nu \rightarrow \bar{x}, x_\nu \notin S \right\}$$

for the *generalized gradient of Clarke*, where  $S$  is arbitrary set of measure zero.

It is easy to see that

$$N(\bar{x}; \Omega) = \partial\delta(\bar{x}, \Omega) \quad \text{if } \bar{x} \in \Omega$$

for the *indicator function* of the set  $\Omega$ . We shall also use another representation of the normal cone (4.1) in terms of the subdifferential (4.9) for the Lipschitz continuous *distance function*  $\text{dist}(\cdot, \Omega)$ . The following result is proved in [35, Proposition 2.7].

4.14. PROPOSITION. *For any nonempty set  $\Omega$ , one has*

$$N(\bar{x}; \Omega) = \text{cone}[\partial\text{dist}(\bar{x}, \Omega)] \quad \text{if } \bar{x} \in \text{cl } \Omega.$$

Various calculus results for the subdifferentials (4.9) and the normal cones (4.1) can be easily deduced from the coderivative results stated above; cf. [41]. A number of calculus rules for (4.1) and (4.9) have been proved in [12, 22, 28, 35, 38, 53] (see also references therein) by using different but somewhat close variational approximations. In this paper we employ the following simple corollary of Proposition 4.10.



4.15. PROPOSITION. Let  $f_1$  and  $f_2$  be extended-real-valued and lower semicontinuous functions one of which is Lipschitz continuous around  $\bar{x}$ . Then one has

$$\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x}). \quad (4.10)$$

Moreover, (4.10) becomes an equality if either  $f_1$  or  $f_2$  is strictly differentiable at  $\bar{x}$ .

## 5. Nonsmooth Optimization in Finite Dimensions

In this section we study nonsmooth optimization problems in finite dimensions and obtain necessary optimality conditions by using generalized differentiation constructions in Section 4. First we consider a broad class of mathematical programming problems with many inequality, equality, and geometric type constraints. We provide necessary conditions for such problems in the form of a *generalized Lagrange multiplier rule*. Then we employ this result for studying the finite difference problems  $(P_K)$  with a *variable stepsize* (varying discrete time interval) approximating the differential inclusion problem  $(P)$  with free time. We derive necessary conditions for the discrete approximation problems in a *refined Euler-Lagrange form* with *no convexity* in the original and/or adjoint inclusions.

For our purposes of applications to dynamic optimization problems like  $(P_K)$ , we need to consider mathematical programming problems with *many geometric* constraints, *nonsmooth inequality* constraints, and *smooth equality* constraints.

Given real-valued functions  $\phi_i$ ,  $n$ -vector-valued functions  $g_j$ , and sets  $\Delta_j$  in the space  $\mathbf{R}^d$ , we formulate the *mathematical programming* problem  $(MP)$  as follows:

$$\text{minimize } \phi_0(z) \text{ for } z \in \mathbf{R}^d \text{ subject to} \quad (5.1)$$

$$\phi_i(z) \leq 0 \text{ for } i = 1, \dots, s; \quad (5.2)$$

$$g_j(z) = 0 \in \mathbf{R}^n \text{ for } j = 0, \dots, m; \quad (5.3)$$

$$z \in \Delta_j \text{ for } j = 0, \dots, l. \quad (5.4)$$

5.1. PROPOSITION. Let  $\bar{z}$  be an optimal solution to problem (MP). Assume that the functions  $\phi_i$  are Lipschitz continuous, the functions  $g_j$  are smooth, and the sets  $\Delta_j$  are closed around  $\bar{z}$ . Then there exist real numbers  $\{\mu_i \mid i = 0, \dots, s\}$  as well as vectors  $\{\psi_j \in \mathbf{R}^n \mid j = 0, \dots, m\}$  and  $\{z_j^* \in \mathbf{R}^d \mid j = 0, \dots, l\}$ , not all zero, such that

$$z_j^* \in N(\bar{z}; \Delta_j) \text{ for } j = 0, \dots, l; \quad (5.5)$$

$$\mu_i \geq 0 \text{ for } i = 0, \dots, s; \quad (5.6)$$

$$\mu_i \phi_i(\bar{z}) = 0 \text{ for } i = 1, \dots, s; \quad (5.7)$$

$$-z_0^* - \dots - z_l^* \in \partial\left(\sum_{i=0}^s \mu_i \phi_i\right)(\bar{z}) + \sum_{j=0}^m (\nabla g_j(\bar{z}))^* \psi_j. \quad (5.8)$$

*Proof.* This proposition follows from general necessary conditions in nondifferentiable programming proved in Mordukhovich [33, Theorem 1] and [35, Corollary 7.5.1] on the basis of the so-called *metric approximation method*. This method allows us to approximate the nonsmooth constrained problem (MP) by a special parametric family of smooth unconstrained minimization problems and then to get necessary conditions (5.5)–(5.8) by passing to the limit in the classical Fermat stationary rule. In such a way, we obtain more general results in the presence of nonsmooth equality constraints, vector objective functions, non-Lipschitzian data, etc.; see [35, Chapter 2].  $\square$

Now let us reduce the discrete approximation problems ( $P_K$ ) in (3.11)–(3.14) to the (MP) form (5.1)–(5.4). For any fixed  $K = 1, 2, \dots$ , we consider the following minimization problem with respect to variables  $\theta > 0$  and  $(x_0, x_1, \dots, x_K) \in \mathbf{R}^{(K+1)n}$ :

$$\text{minimize } \varphi_0(x_0, x_K, \theta) + |x_0 - \bar{x}(0)|^2 + (\theta - \bar{T})^2 + \quad (5.9)$$

$$(\theta/K) \sum_{j=0}^{K-1} f(j\theta/K, x_j, (K/\theta)(x_{j+1} - x_j)) + \sum_{j=0}^{K-1} \int_{j\theta/K}^{(j+1)\theta/K} |(K/\theta)(x_{j+1} - x_j) - \dot{x}(t)|^2 dt$$

subject to the constraints

$$x_{j+1} \in x_j + (\theta/K)F(j\theta/K, x_j) \text{ for } j = 0, \dots, K-1; \quad (5.10)$$

$$\varphi_i(x_0, x_K, \theta) \leq \alpha_{iK} \text{ for } i = 1, \dots, q; \quad (5.11)$$

$$-\beta_{iK} \leq \varphi_i(x_0, x_K, \theta) \leq \beta_{iK} \text{ for } i = q + 1, \dots, q + r; \quad (5.12)$$

$$(x_0, x_K, \theta) \in \Omega_K := \Omega + \gamma_K B \quad (5.13)$$

where  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , is the considered optimal solution to the original problem ( $P$ ) and  $(\alpha_K, \beta_K, \gamma_K)$  are the defined constraint perturbations in ( $P_K$ ).

One can easily see that for each  $K = 1, 2, \dots$  problems ( $P_K$ ) and (5.9)–(5.13) are equivalent by virtue of the following correspondence between decision variables:

$$\{x_K(t_j), T_K\} \longleftrightarrow \{x_j, \theta\} \text{ for } j = 0, \dots, K. \quad (5.14)$$

For convenience, let us introduce the auxiliary variables

$$y_j := (K/\theta)(x_{j+1} - x_j) \text{ for } j = 0, \dots, K - 1.$$

Now problem (5.9)–(5.13) (and, therefore, ( $P_K$ )) can be rewritten in the equivalent form (5.1)–(5.4) with respect to variables  $z = (x_0, \dots, x_K, y_0, \dots, y_{K-1}, \theta) \in \mathbf{R}^{(2K+1)n+1}$  as follows:

$$\text{minimize } \phi_0(z) := \varphi_0(x_0, x_K, \theta) + |x_0 - \bar{x}(0)|^2 + (\theta - \bar{T})^2 + \quad (5.15)$$

$$(\theta/K) \left( \sum_{j=0}^{K-1} f(j\theta/K, x_j, y_j) + \sum_{j=0}^{K-1} \int_{j\theta/K}^{(j+1)\theta/K} |y_j - \dot{\bar{x}}(t)|^2 dt \right)$$

subject to

$$\phi_i(z) := \varphi_i(x_0, x_K, \theta) - \alpha_{iK} \leq 0 \text{ for } i = 1, \dots, q; \quad (5.16)$$

$$\phi_i(z) := \varphi_i(x_0, x_K, \theta) - \beta_{iK} \leq 0 \text{ for } i = q + 1, \dots, q + r; \quad (5.17)$$

$$\phi_{r+i}(z) := -\varphi_i(x_0, x_K, \theta) - \beta_{iK} \leq 0 \text{ for } i = q + 1, \dots, q + r; \quad (5.18)$$

$$g_j(z) := x_{j+1} - x_j - (\theta/K)y_j = 0 \text{ for } j = 0, \dots, K - 1; \quad (5.19)$$

$$z \in \Delta_j := \{(x_0, \dots, y_{K-1}, \theta) \in \mathbf{R}^{(2K+1)n+1} \mid y_j \in F(j\theta/K, x_j)\} \text{ for } j = 0, \dots, K - 1; \quad (5.20)$$

$$z \in \Delta_K := \{(x_0, \dots, y_{K-1}, \theta) \in \mathbf{R}^{(2K+1)n+1} \mid (x_0, x_K, \theta) \in \Omega_K\}. \quad (5.21)$$

Obviously, (5.15)–(5.21) is a problem of mathematical programming in the form (5.1)–(5.4) where  $d = (2K + 1)n + 1$ ,  $s = q + 2r$ ,  $m = K - 1$ , and  $l = K$ . Moreover, the real-valued functions  $\phi_i$ , the smooth vector functions  $g_j$ , and the sets  $\Delta_j$  have special structures. Now one can employ Proposition 5.1 to obtain necessary optimality conditions for the dynamic optimization problem  $(P_K)$  taking into account the special nature of (5.15)–(5.21) and calculus rules for the generalized differentiation constructions in Section 4.

For simplicity, in the following theorem we deal with autonomous systems where  $F$  and  $f$  do not depend on the time variable. Moreover, to avoid technical complications and taking into account applications in Section 6, we focus on the case when the derivative  $\dot{\bar{x}}(\cdot)$  of the reference optimal solution to  $(P)$  is Riemann integrable on  $[0, \bar{T}]$ , i.e., continuous for a.e.  $t \in [0, \bar{T}]$ . The general case of  $\bar{x}(\cdot) \in W^{1,\infty}[0, \bar{T}]$  will be discussed in Remark 5.5.

**5.2. THEOREM.** *Let  $\{\bar{x}_K(\cdot), \bar{T}_K\}$  be an optimal solution to problem  $(P_K)$  of autonomous dynamics where  $\dot{\bar{x}}(\cdot)$  is Riemann integrable on  $[0, \bar{T}_K]$ . Assume that the functions  $\varphi_i$  and  $f$  are Lipschitz continuous and the sets  $\Omega$  and  $\text{gph } F$  are closed around  $\{\bar{x}_K(\cdot), \bar{T}_K\}$ . Then there exist real numbers  $\{\lambda_{iK} \mid i = 0, \dots, q+r\}$  and a discrete  $n$ -vector function  $\{p_K(t_j) \mid j = 0, \dots, K\}$ , not all zero, such that*

$$\lambda_{iK} \geq 0 \text{ for } i = 0, \dots, q; \quad (5.22)$$

$$\lambda_{iK}(\varphi_i(\bar{x}_K(0), \bar{x}_K(\bar{T}_K), \bar{T}_K) - \alpha_{iK}) = 0 \text{ for } i = 1, \dots, q; \quad (5.23)$$

$$\left(\frac{p_K(t_{j+1}) - p_K(t_j)}{h_K}, p_K(t_{j+1}) + \frac{\lambda_{0K}}{h_K} \rho_{jK}\right) \in \lambda_{0K} \partial f(\bar{x}_K(t_j), \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K}) + \quad (5.24)$$

$$N((\bar{x}_K(t_j), \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K}); \text{gph } F) \text{ for } j = 0, \dots, K - 1;$$

$$(p_K(0) + 2\lambda_{0K}(\bar{x}(0) - \bar{x}_K(0)), -p_K(\bar{T}_K), \bar{H}_K + 2\lambda_{0K}(\bar{T} - \bar{T}_K) + \lambda_{0K} \varrho_K) \in \quad (5.25)$$

$$\partial \left( \sum_{i=0}^{q+r} \lambda_{iK} \varphi_i(\bar{x}_K(0), \bar{x}_K(\bar{T}_K), \bar{T}_K) + N((\bar{x}_K(0), \bar{x}_K(\bar{T}_K), \bar{T}_K); \Omega_K) \right)$$

where

$$h_K := \bar{T}_K / K; \quad t_j := j h_K \text{ for } j = 0, \dots, K \quad (t_K = \bar{T}_K); \quad (5.26)$$

$$\bar{H}_K := \frac{1}{K} \sum_{j=0}^{K-1} \left[ \langle p_K(t_{j+1}), \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K} \rangle - \lambda_{0K} f(\bar{x}_K(t_j), \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K}) \right]; \quad (5.27)$$

$$\rho_{jK} := 2 \int_{t_j}^{t_{j+1}} \left( \dot{\bar{x}}(t) - \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K} \right) dt \quad \text{for } j = 0, \dots, K-1; \quad (5.28)$$

$$\varrho_K := \sum_{j=0}^{K-1} \left[ \frac{j}{K} \left| \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K} - \dot{\bar{x}}(t_j) \right|^2 - \frac{(j+1)}{K} \left| \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K} - \dot{\bar{x}}(t_{j+1}) \right|^2 \right]. \quad (5.29)$$

*Proof.* Let  $\bar{z} = (\bar{x}_0, \dots, \bar{x}_K, \bar{y}_0, \dots, \bar{y}_{K-1}, \bar{\theta})$  be an optimal solution to the mathematical programming problem (5.15)–(5.21). According to Proposition 5.1, there exist real numbers  $\{\mu_i \mid i = 0, \dots, q+2r\}$  as well as vectors  $\psi_j \in \mathbf{R}^n$  ( $j = 0, \dots, K-1$ ) and  $z_j^* := (x_{j0}^*, \dots, x_{jK}^*, \bar{y}_{j0}, \dots, \bar{y}_{jK-1}, \theta_j^*) \in \mathbf{R}^{(2K+1)n+1}$  ( $j = 0, \dots, K$ ), not all zero, such that conditions (5.5)–(5.8) are fulfilled for the initial data in (5.15)–(5.21). Note that these  $\mu_i$ ,  $\psi_j$ , and  $z_j^*$  as well as the solutions  $\bar{z}$  depend on  $K$  but here we omit the index "K" for simplicity.

Without loss of generality one can always suppose that  $\beta_{iK} > 0$  for  $i = q+1, \dots, q+r$  and all  $K = 1, 2, \dots$ . Then the complementary slackness conditions (5.17) and (5.18) obviously imply that

$$\mu_i \cdot \mu_{r+i} = 0 \quad \forall i = q+1, \dots, q+r.$$

Denoting

$$\lambda_i := \mu_i \quad \text{for } i = 0, \dots, q \quad \& \quad \lambda_i := \mu_i - \mu_{r+i} \quad \text{for } i = q+1, \dots, q+r,$$

we get conditions

$$\lambda_i \geq 0 \quad \text{for } i = 0, \dots, q \quad \text{and} \quad (5.30)$$

$$\lambda_i (\varphi_i(\bar{x}_0, \bar{x}_K, \bar{\theta}) - \alpha_{iK}) = 0 \quad \text{for } i = 1, \dots, q \quad (5.31)$$

ensuring that  $\{\lambda_i \mid i = 0, \dots, q+r\}$  are not equal to zero simultaneously with  $\{\psi_j \mid j = 0, \dots, K-1\}$  and  $\{z_j^* \mid j = 0, \dots, K\}$ .

Now let us express condition (5.8) in terms of the initial data in problem (5.15)–(5.21). From the structure of (5.15)–(5.18) and the calculus rule in Proposition 4.15 one has

$$\partial \left( \sum_{i=0}^{q+2r} \mu_i \phi_i \right) (\bar{z}) \subset \partial \left( \sum_{i=0}^{q+r} \lambda_i \varphi_i \right) (\bar{z}) + \lambda_0 \partial \zeta (\bar{z}) + \lambda_0 \partial \xi (\bar{z}) + \quad (5.32)$$

$$2(\lambda_0(\bar{x}_0 - \bar{x}(0)), 0, \dots, 0, \lambda_0(\bar{\theta} - \bar{T}))$$

where

$$\zeta(z) = \zeta(x_0, \dots, x_K, y_0, \dots, y_{K-1}, \theta) := \frac{\theta}{K} \sum_{j=0}^{K-1} f(x_j, y_j), \quad (5.33)$$

$$\xi(z) = \xi(x_0, \dots, x_K, y_0, \dots, y_{K-1}, \theta) := \sum_{j=0}^{K-1} \int_{j\theta/K}^{(j+1)\theta/K} |y_j - \dot{\bar{x}}(t)|^2 dt. \quad (5.34)$$

It is easy to see that

$$\partial \left( \sum_{i=0}^{q+r} \lambda_i \varphi_i \right) (\bar{z}) = \{u_0, 0, \dots, u_K, 0, \vartheta\} \mid (u_0, u_K, \vartheta) \in \partial \left( \sum_{i=0}^{q+r} \lambda_i \varphi_i \right) (\bar{x}_0, \bar{x}_K, \bar{\theta}), \quad (5.35)$$

$$\partial \xi(\bar{z}) = \{(v_0, \dots, v_{K-1}, w_0, \dots, w_{K-1}, \frac{1}{K} \sum_{j=0}^{K-1} f(\bar{x}_j, \bar{y}_j)) \mid (v_j, w_j) \in \partial f(\bar{x}_j, \bar{y}_j)\}. \quad (5.36)$$

Next let us consider the function  $\xi$  in (5.34). This function is obviously differentiable in the  $(x, y)$  variables. Moreover, under the Riemann integrability assumption on  $\dot{\bar{x}}(\cdot)$ , this function is also differentiable in  $\theta$ . The latter follows from a variant of the fundamental theorem of the classical calculus where an integrand may not be continuous but admits a primitive and takes all the intermediate values of its range.

Now using the classical formulae for the differentiation under the integral sign and with respect to varying integral limits, one gets

$$\nabla_{x_j} \xi(\bar{z}) = 0 \quad \text{for } j = 0, \dots, K; \quad (5.37)$$

$$\nabla_{y_j} \xi(\bar{z}) = \xi_j := 2 \int_{j\bar{\theta}/K}^{(j+1)\bar{\theta}/K} (\bar{y}_j - \dot{\bar{x}}(t)) dt \quad \text{for } j = 0, \dots, K-1; \quad (5.38)$$

$$\nabla_{\theta} \xi(\bar{z}) = \xi_{\theta} := \sum_{j=0}^{K-1} \left[ \frac{j+1}{K} |\bar{y}_j - \dot{\bar{x}}((j+1)\bar{\theta}/K)|^2 - \frac{j}{K} |\bar{y}_j - \dot{\bar{x}}(j\bar{\theta}/K)|^2 \right]. \quad (5.39)$$

Note that to justify (5.39) one needs to assume the one-sided continuity of  $\dot{\bar{x}}(\cdot)$  at the points  $j\bar{\theta}/K$  for  $j = 0, \dots, K-1$  (otherwise, we should take  $\dot{\bar{x}}(\cdot)$  at points nearby). But this does not restrict generality as  $K \rightarrow \infty$  due to the procedure in (6.18).

Next let us compute the vectors  $(\nabla g_j(\bar{z}))^* \psi_j$  for  $g_j$  in (5.19). For each  $j = 0, \dots, K-1$ , we have the following expressions

$$(\nabla_{x_j} g_j(\bar{z}))^* \psi_j = -\psi_j, \quad (\nabla_{x_{j+1}} g_j(\bar{z}))^* \psi_j = \psi_j, \quad (\nabla_{x_k} g_j(\bar{z}))^* \psi_j = 0 \text{ if } k \neq j; \quad (5.40)$$

$$(\nabla_{y_j} g_j(\bar{z}))^* \psi_j = -(\bar{\theta}/K) \psi_j, \quad (\nabla_{y_k} g_j(\bar{z}))^* \psi_j = 0 \text{ if } k \neq j; \quad (5.41)$$

$$(\nabla_{\theta} g_j(\bar{z}))^* \psi_j = -\frac{1}{K} \sum_{j=0}^{K-1} \langle \psi_j, \bar{y}_j \rangle. \quad (5.42)$$

Thus we have calculated the left-hand side of (5.8) for the data in (5.15)–(5.19). To represent the right-hand side of (5.8), we observe that conditions (5.5) for the sets  $\Delta_j$  in (5.20) and (5.21) are equivalent to

$$(x_{jj}^*, y_{jj}^*) \in N((\bar{x}_j, \bar{y}_j); \text{gph } F) \quad \& \quad (5.43)$$

$$x_{jk}^* = y_{jk}^* = 0 \text{ if } k \neq j, \quad \forall j = 0, \dots, K-1;$$

$$(x_{K0}^*, x_{KK}^*) \in N((\bar{x}_0, \bar{x}_K); \Omega_K) \quad \& \quad x_{Kj}^* = y_{Kj}^* = 0 \text{ otherwise.} \quad (5.44)$$

Now putting all the calculations (5.32)–(5.44) together in Proposition 5.1, one gets the equalities:

$$\begin{aligned} -x_{00}^* - x_{K0}^* &= u_0 + 2\lambda_0(\bar{x}_K - \bar{x}(0)) + \frac{\lambda_0 \bar{\theta}}{K} v_0 - \psi_0; \\ -x_{jj}^* &= \frac{\lambda_0 \bar{\theta}}{K} v_j + \psi_{j-1} - \psi_j \text{ for } j = 1, \dots, K-1; \\ -x_{KK}^* &= u_K + \psi_{K-1}; \\ -y_{jj}^* &= \frac{\lambda_0 \bar{\theta}}{K} w_j - \frac{\theta}{K} \psi_j + \lambda_0 \xi_j \text{ for } j = 0, \dots, K-1; \\ 0 &= \vartheta + \frac{\lambda_0}{K} \sum_{j=0}^{K-1} f(\bar{x}_j, \bar{y}_j) - \frac{1}{K} \sum_{j=0}^{K-1} \langle \psi_j, \bar{y}_j \rangle + 2\lambda_0(\bar{\theta} - \bar{T}) + \lambda_0 \xi_\theta \end{aligned}$$

where  $\xi_j$  and  $\xi_\theta$  are defined in (5.38) and (5.39),

$$(u_0, u_K, \vartheta) \in \partial \left( \sum_{i=0}^{q+r} \lambda_0 \varphi_i \right) (\bar{x}_0, \bar{x}_K, \bar{\theta}), \quad \text{and}$$

$$(v_j, w_j) \in \partial f(\bar{x}_j, \bar{y}_j) \text{ for } j = 0, \dots, K-1.$$

Thus denoting

$$p_0 := x_{K0}^* + u_0 + 2\lambda_0(\bar{x}_0 - \bar{x}(0)) \ \& \ p_j := \psi_{j-1} \ \text{for } j = 1, \dots, K,$$

we obtain the following necessary conditions for the solution  $(\bar{x}_0, \dots, \bar{x}_K, \bar{y}_0, \dots, \bar{y}_{K-1}, \bar{\theta})$  to problem (5.15)–(5.21): there are real numbers  $\{\lambda_i \mid i = 0, \dots, q+r\}$  and  $n$ -vectors  $\{p_j \mid j = 0, \dots, K\}$ , not all zero, such that one has (5.30), (5.31), and

$$((K/\bar{\theta})(p_{j+1} - p_j), p_{j+1} - (\lambda_0 K/\bar{\theta})\xi_j) \in \lambda_0 \partial f(\bar{x}_j, (K/\bar{\theta})(\bar{x}_{j+1} - \bar{x}_j)) + \quad (5.45)$$

$$N((\bar{x}_j, (K/\bar{\theta})(\bar{x}_{j+1} - \bar{x}_j)); \text{gph } F) \ \text{for } j = 0, \dots, K-1;$$

$$(p_0 + 2\lambda_0(\bar{x}(0) - \bar{x}_0), -p_K, 2\lambda_0(\bar{T} - \bar{\theta}) - \xi_\theta) + \quad (5.46)$$

$$\frac{1}{K} \sum_{j=0}^{K-1} [\langle p_{j+1}, (K/\bar{\theta})(\bar{x}_{j+1} - \bar{x}_j) \rangle - \lambda_0 f(\bar{x}_j, (K/\bar{\theta})(\bar{x}_{j+1} - \bar{x}_j))] \in \partial(\sum_{j=0}^{q+r} \lambda_j \varphi_j)(\bar{x}_0, \bar{x}_K, \bar{\theta}).$$

Now let  $\{\bar{x}_K(\cdot), \bar{T}_K\}$  be an optimal solution to the discrete approximation problem  $(P_K)$ . Using correspondence (5.14), one gets an optimal solution  $(\bar{x}_0, \dots, \bar{x}_K, \bar{\theta})$  to problem (5.15)–(5.21) for which the necessary optimality conditions (5.30), (5.31), (5.45), and (5.46) hold. For any fixed  $K$  in these optimality conditions, we mark the dependence of the multipliers  $\lambda_i$  on  $K$  and set

$$p_K(t_j) := p_j \ \text{for } j = 0, \dots, K$$

taking (5.26) into account. Then conditions (5.30), (5.31), (5.45), and (5.46) turn into the necessary optimality conditions (5.22)–(5.25) for the problem  $(P_K)$  under consideration. This ends the proof of the theorem.  $\square$

**5.3. COROLLARY.** *In addition to the assumptions of Theorem 5.2, let us suppose that the multifunction  $F$  is pseudo-Lipschitzian around  $(\bar{x}_K(t_j), (\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j))/h_K)$  for each  $j = 0, \dots, K-1$ . Then conditions (5.22)–(5.25) hold with  $(\lambda_{0K}, \dots, \lambda_{q+rK}, p_K(\bar{T}_K)) \neq 0$ . Therefore, one can set*

$$|p_K(\bar{T}_K)| + \sum_{i=0}^q \lambda_{iK} + \sum_{i=q+1}^{q+r} |\lambda_{iK}| = 1 \ \forall K = 1, 2, \dots \quad (5.47)$$

*Proof.* If  $\lambda_{0K} = 0$ , then (5.24) is represented as

$$\frac{p_K(t_{j+1}) - p_K(t_j)}{h_K} \in D^* F(\bar{x}_K(t_j), \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K})(-p_K(t_{j+1})) \ \text{for } j = 0, \dots, K-1$$



in terms of the coderivative (4.4). Now using criterion (c) in Proposition 4.7, one gets that  $p_K(\bar{T}_K) = 0$  implies  $p_K(t_j) = 0$  for all  $j = 0, \dots, K - 1$ . This proves the corollary.  $\square$

5.4. *Remark.* From Theorem 5.2, one can easily obtain necessary optimality conditions for nonautonomous problems  $(P_K)$  using a standard reduction to the autonomous case with respect to a new state variable  $z = (t, x)$ . The extended finite difference inclusion for  $z$  is written in the form

$$z_K(t_{j+1}) \in z_K(t_j) + h_K \tilde{F}(z_K(t_j)) \text{ for } j = 0, \dots, K - 1$$

where  $\tilde{F}(z) := (1, F(z))$ .

5.5. *Remark.* If the time  $T_K$  is fixed in problem  $(P_K)$ , then  $\theta$  is fixed in the equivalent problem (5.9)–(5.13). In this case we do not need the assumption about the Riemann integrability of  $\dot{\bar{x}}(\cdot)$  in the proof of Theorem 5.2 and its applications in Section 6 which work for any  $\bar{x}(\cdot) \in W^{1,\infty}$ ; cf. [43]. The only place where we use the Riemann integrability of  $\dot{\bar{x}}(\cdot)$  in the latter proof is the differentiability of function (5.34) in the  $\theta$  variable.

In the general setting when  $\bar{x}(\cdot) \in W^{1,\infty}$ , the function  $\xi$  in (5.34) is Lipschitz continuous in its variables jointly being differentiable at almost all points. Moreover, one can compute the gradients of  $\xi$  at the points of differentiability according to the classical results. This allows us to use effectively the construction of Clarke's generalized gradient and its connection with the subdifferential (4.9) for estimating the impact of the term (5.34) to the conditions of Theorem 5.2.

It follows from (5.34) and Proposition 4.15 that

$$\partial\xi(\bar{z}) \subset \partial\xi_0(\bar{z}) + \dots + \partial\xi_{K-1}(\bar{z}) \tag{5.48}$$

where

$$\xi_j(z) = \xi_j(x_0, \dots, x_K, y_0, \dots, y_{K-1}, \theta) := \int_{j\theta/K}^{(j+1)\theta/K} |y_j - \dot{\bar{x}}(t)|^2 dt \tag{5.49}$$

for  $j = 0, \dots, K - 1$ . By virtue of Proposition 4.13(ii) one has

$$\partial\xi_j(\bar{z}) \subset \partial_C \xi_j(\bar{z}) = \text{co}\{\lim_{\nu \rightarrow \infty} \nabla \xi_j(z_\nu) \mid \xi_j \text{ is differentiable at } z_\nu \rightarrow \bar{z}\}. \tag{5.50}$$

where the last expression is invariant to "excluding sets of measure zero".

Now let us calculate the partial derivatives of  $\xi_j$  at  $z_\nu = (x_{0\nu}, \dots, x_{K\nu}, y_{0\nu}, \dots, y_{K-1\nu}, \theta_\nu)$  for any  $j = 0, \dots, K-1$ . From (5.49) one obviously gets

$$\nabla_{x_s} \xi_j(z_\nu) = 0 \quad (s = 0, \dots, K) \quad \& \quad \nabla_{y_s} \xi_j(z_\nu) = 0 \quad (s = 0, \dots, K-1; s \neq j); \quad (5.51)$$

$$\nabla_{y_j} \xi_j(z_\nu) = 2 \int_{j\theta_\nu/K}^{(j+1)\theta_\nu/K} (y_{j\nu} - \dot{\bar{x}}(t)) dt. \quad (5.52)$$

Moreover, we can compute

$$\nabla_{\theta} \xi_j(z_\nu) = \frac{j+1}{K} |y_{j\nu} - \dot{\bar{x}}((j+1)\theta_\nu/K)|^2 - \frac{j}{K} |y_{j\nu} - \dot{\bar{x}}(j\theta_\nu/K)|^2. \quad (5.53)$$

at almost all  $z_\nu$  where  $\xi_j$  is differentiable. Thus using (5.48)–(5.53), we obtain

$$\partial \xi(\bar{z}) \subset (\nabla_{x_0} \xi(\bar{z}), \dots, \nabla_{x_K} \xi(\bar{z}), \nabla_{y_0} \xi(\bar{z}), \dots, \nabla_{y_{K-1}} \xi(\bar{z}), C_\theta(\bar{z})) \quad (5.54)$$

where  $\nabla_{x_j} \xi(\bar{z})$  and  $\nabla_{y_j} \xi(\bar{z})$  are expressed in (5.37) and (5.38) while

$$C_\theta(\bar{z}) := \frac{1}{K} \sum_{j=0}^{K-1} \text{co} \left\{ \lim_{\theta_\nu \rightarrow \bar{\theta}} [(j+1)|\bar{y}_j - \dot{\bar{x}}((j+1)\theta_\nu/K)|^2 - j|\bar{y}_j - \dot{\bar{x}}(j\theta_\nu/K)|^2] \right\}. \quad (5.55)$$

(Note that similarly to Clarke [11, Example 2.2.5], one can express the convex hulls of the limiting points in (5.55) in terms of the essential supremum and essential infimum of the functions

$$\eta_{jK}(\theta) := (j+1)|\bar{y}_j - \dot{\bar{x}}((j+1)\theta/K)|^2 - j|\bar{y}_j - \dot{\bar{x}}(j\theta/K)|^2 \quad \text{for } j = 0, \dots, K-1$$

at  $\theta = \bar{\theta}$ ).

Now coming back to the notation of problem  $(P_K)$ , we provide necessary optimality conditions in form (5.22)–(5.28) where instead of equality (5.29) for  $\varrho_K$  one has the inclusion

$$\varrho_K \in \frac{1}{K} \sum_{j=0}^{K-1} \text{co} \left\{ \lim_{\theta \rightarrow \bar{T}_K} [j \left| \frac{\bar{x}_K(t_{j+1}^\theta) - \bar{x}_K(t_j^\theta)}{h_K} - \dot{\bar{x}}_K(t_j^\theta) \right|^2 - (j+1) \left| \frac{\bar{x}_K(t_{j+1}^\theta) - \bar{x}_K(t_j^\theta)}{h_K} - \dot{\bar{x}}_K(t_{j+1}^\theta) \right|^2] \right\} \quad (5.56)$$

with  $t_j^\theta := j\theta/K$  for  $j = 0, \dots, K-1$ .

From (5.56) one can conclude that  $\varrho_K \rightarrow 0$  as  $K \rightarrow \infty$  when  $\dot{\bar{x}}(\cdot)$  is Riemann integrable on  $[0, \bar{T}]$ ; cf. the arguments in (6.18). The question is about such a conclusion in more general settings for problems with non-fixed time.

## 6. Necessary Conditions for Differential Inclusions

In this section we prove necessary optimality conditions for the original Bolza problem  $(P)$  with free time by passing to the limit in the necessary conditions obtained above for its discrete approximations. The following two circumstances are most important in this procedure:

- 1) the strong  $W^{1,2}$ -convergence of discrete optimal solutions proved in Section 2; and
- 2) robustness of the generalized differential constructions and the dual differential characterization of the Lipschitzian behavior of multifunctions in Section 4 ensuring the convergence of the adjoint arcs in optimality conditions.

Note that the collection of generalized differentiation properties we need is inherent in objects of Section 4 but not in other known generalized differential constructions.

Now we obtain necessary optimality conditions for problem  $(P)$  in the refined Euler-Lagrange form. Using Theorem 5.1, we assume in what follows that the derivative  $\dot{\bar{x}}(\cdot)$  of the optimal trajectory under consideration is *Riemann integrable* on  $[0, \bar{T}]$ . First we deal with the autonomous case when  $F$  and  $f$  do not depend on  $t$ .

6.1. THEOREM. *Let  $\bar{x}(t)$ ,  $0 \leq t \leq \bar{T}$ , be an optimal solution to the autonomous problem  $(P)$  which possesses the property of relaxation stability. In addition to (H1) and (H7), we assume that the functions  $\varphi_i$  are Lipschitz continuous around  $(\bar{x}(0), \bar{x}(\bar{T}), \bar{T})$  and the function  $f$  is Lipschitz continuous with modulus  $l_f$  around any point of the set*

$$A := \{(x, v) \in \mathbf{R}^{2n} \mid x \in U, v \in F(x)\}. \quad (6.1)$$

*Then there exist real numbers  $\lambda_0, \dots, \lambda_{q+r}$  and an absolutely continuous function  $p : [0, \bar{T}] \rightarrow \mathbf{R}^n$ , not all zero, such that*

$$\lambda_i \geq 0 \text{ for } i = 0, \dots, q; \quad (6.2)$$

$$\lambda_i \varphi_i(\bar{x}(0), \bar{x}(\bar{T}), \bar{T}) = 0 \text{ for } i = 1, \dots, q; \quad (6.3)$$

$$\dot{p}(t) \in \text{co}\{u \mid (u, p(t)) \in \lambda_0 \partial f(\bar{x}(t), \dot{\bar{x}}(t)) + N((\bar{x}(t), \dot{\bar{x}}(t)); \text{gph } F) \text{ a.e. } t \in [0, \bar{T}]; \quad (6.4)$$

$$(p(0), -p(\bar{T}), \bar{H}) \in \partial\left(\sum_{i=0}^{q+r} \lambda_i \varphi_i\right)(\bar{x}(0), \bar{x}(\bar{T}), \bar{T}) + N((\bar{x}(0), \bar{x}(\bar{T}), \bar{T}); \Omega) \quad (6.5)$$

where

$$\bar{H} := \frac{1}{\bar{T}} \int_0^{\bar{T}} [\langle p(t), \dot{\bar{x}}(t) \rangle - \lambda_0 f(\bar{x}(t), \dot{\bar{x}}(t))] dt. \quad (6.6)$$

*Proof.* Let us consider a sequence of discrete approximations  $(P_K)$  in (3.1), (3.11)–(3.14) and a sequence of optimal solutions  $\{\bar{x}_K(\cdot), \bar{T}_K\}$  to  $(P_K)$  which converges to  $\{\bar{x}(\cdot), \bar{T}\}$  in the sense of Theorem 3.2. Now using the necessary optimality conditions for  $\{\bar{x}_K(\cdot), \bar{T}_K\}$  in Theorem 5.2, we find sequences of real numbers  $\{\lambda_{iK} \mid i = 0, \dots, q+r\}$  and discrete  $n$ -vector functions  $\{p_K(t_j) \mid j = 0, \dots, K\}$ , satisfying (5.22)–(5.29). Due to Corollary 5.3, one can always impose the normalization condition (5.47).

We shall consider the piecewise-linear extensions  $\bar{x}_K(t)$  and  $p_K(t)$  of the corresponding discrete functions on the continuous-time intervals  $[0, \bar{T}_K]$  according to (3.2) and (3.3). We also denote

$$\rho_K(t) := \rho_{jK}/h_K \text{ for } t \in [t_j, t_{j+1}), \quad j = 0, \dots, K-1$$

where the numbers  $\rho_j$  are defined in (5.28). One can see that

$$\begin{aligned} \int_0^{\bar{T}_K} |\rho_K(t)| dt &= \sum_{j=0}^{K-1} |\rho_{jK}| \leq 2 \sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} \left| \dot{\bar{x}}(t) - \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K} \right| dt = \\ &2 \int_0^{\bar{T}_K} |\dot{\bar{x}}(t) - \dot{\bar{x}}_K(t)| dt \rightarrow 0 \text{ as } K \rightarrow \infty \end{aligned} \quad (6.7)$$

by virtue of Theorem 3.2. Considering (if necessary) the piecewise-constant extensions of  $\dot{\bar{x}}_K(t)$  and  $\rho_K(t)$  on the interval  $[0, \bar{T}]$  and using the classical results, we can suppose without any loss of generality that

$$\dot{\bar{x}}_K(t) \rightarrow \dot{\bar{x}}(t) \text{ and } \rho_K(t) \rightarrow 0 \text{ a.e. } t \in [0, \bar{T}] \text{ as } K \rightarrow \infty. \quad (6.8)$$

Now let us obtain an estimate of the adjoint arcs  $p_K(\cdot)$  for large  $K$  which appears to be a crucial point of the proof. According to the adjoint discrete inclusion (5.24) and Definition 4.2 of the coderivative  $D^*F$ , one can find vectors  $(v_{jK}, w_{jK}) \in \partial f(\bar{x}_K(t_j), (\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j))/h_K)$  such that

$$\frac{p_K(t_{j+1}) - p_K(t_j)}{h_K} - \lambda_{0K} v_{jK} \in D^*F(\bar{x}_K(t_j), \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K})(\lambda_{0K} w_{jK} - \quad (6.9)$$

$$\lambda_{0K}\rho_{jK}/h_K - p_K(t_{j+1})) \text{ for } j = 0, \dots, K-1.$$

By virtue of the Lipschitz continuity of  $F$  around  $\bar{x}(t)$  with modulus  $l_F$  in (3.5) and due to the uniform convergence  $\bar{x}_K(\cdot) \rightarrow \bar{x}(\cdot)$ , we employ Corollary 4.8 in (6.9) and provide the estimate

$$\left| \frac{p_K(t_{j+1}) - p_K(t_j)}{h_K} - \lambda_{0K}v_{jK} \right| \leq l_F |\lambda_{0K}w_{jK} - \lambda_{0K}\rho_{jK}/h_K - p_K(t_{j+1})| \quad (6.10)$$

for  $j = 0, \dots, K-1$  when  $K$  is large enough. Taking into account that  $\lambda_{0K} \in [0, 1]$ , one gets from (6.10) the following recurrent sequence for  $p_K(t_j)$ :

$$|p_K(t_j)| \leq (1 + l_F h_K) |p_K(t_{j+1})| + h_K (|v_{jK}| + l_F |w_{jK}|) + l_F |\rho_{jK}| \quad (6.11)$$

for  $j = 0, \dots, K-1$ . Now observe that

$$(\bar{x}_K(t_j), \frac{\bar{x}_K(t_{j+1}) - \bar{x}_K(t_j)}{h_K}) \in A \text{ for } j = 0, \dots, K-1$$

and all large  $K$ , where the set  $A$  is defined in (6.1). So we can employ Proposition 4.13(i) and obtain the estimates

$$|v_{jK}| \leq l_f \ \& \ |w_{jK}| \leq l_f \ \forall j = 0, \dots, K-1 \text{ and large } K \quad (6.12)$$

in terms of the Lipschitz modulus  $l_f$  for the function  $f$  in our assumptions. From (6.11), (6.12), and  $|p_K(\bar{T}_K)| \leq 1$  due to (5.47), one gets

$$|p_K(t_j)| \leq (1 + h_K l_F) |p_K(t_{j+1})| + h_K l_f (1 + l_F) + l_F |\rho_{jK}| \leq \dots \leq \quad (6.13)$$

$$(1 + l_F \bar{T}_K / K)^K + \bar{T}_K l_f (1 + l_F) + l_F \sum_{j=0}^{K-1} |\rho_{jK}| \quad \forall j = 0, \dots, K-1$$

when  $K$  is large enough. Taking into account (6.7) and

$$\bar{T}_K \rightarrow \bar{T} \ \& \ (1 + l_F \bar{T}_K / K)^K \rightarrow \exp[l_F \bar{T}] \text{ as } K \rightarrow \infty,$$

we establish the uniform boundedness of  $\{p_K(t)\}$  on  $[0, \bar{T}_K]$  as  $K \rightarrow \infty$ .

Then employing (6.10) and (6.12), one gets the estimate of the extended adjoint velocities  $\dot{p}_K(t)$  as follows:

$$|\dot{p}_K(t)| = \left| \frac{p_K(t_{j+1}) - p_K(t_j)}{h_K} \right| \leq l_f + l_F (l_f + |\rho_K(t)| + |p_K(t_{j+1})|) \text{ for } t_j \leq t < t_{j+1}. \quad (6.14)$$

Without loss of generality, we can consider  $\dot{p}_K(t)$  on the interval  $[0, \bar{T}]$  using, if necessary, the piecewise-constant extension on  $[\bar{T}_K, \bar{T}]$ . By virtue of (6.7), (6.13), and the classical compactness criterion, estimate (6.14) implies that the sequence  $\{\dot{p}_K(t)\}$  is weakly compact in  $L^1[0, \bar{T}]$ . Therefore, one can find an absolutely continuous function  $p : [0, \bar{T}] \rightarrow \mathbf{R}^n$  such that  $p_K(\cdot) \rightarrow p(\cdot)$  uniformly in  $[0, \bar{T}]$  and  $\dot{p}_K(\cdot) \rightarrow \dot{p}(\cdot)$  weakly in  $L^1[0, \bar{T}]$  (as usual we take all  $K = 1, 2, \dots$ ).

Now we carry out the limiting process in the necessary optimality conditions of Theorem 5.2. Passing to the limit in (5.22) and (5.23) and taking into account that  $\alpha_{iK} \rightarrow 0$  as  $K \rightarrow \infty$ , one easily gets the sign and complementary slackness conditions (6.2) and (6.3) respectively. Taking the limit in (5.47), we provide the normalization condition

$$|p(\bar{T})| + \sum_{i=0}^q \lambda_i + \sum_{i=q+1}^{q+r} |\lambda_i| = 1 \quad (6.15)$$

which ensures that  $\lambda_0, \dots, \lambda_{q+r}$  and  $p(\cdot)$  are not equal to zero simultaneously. Employing the limiting procedure in (6.13), we can actually conclude that if  $p(t_0) = 0$  at some point  $t_0 \in [0, \bar{T}]$ , then  $p(t) \equiv 0$  in the whole interval  $[0, \bar{T}]$ .

Let us pass to the limit in the discrete Euler-Lagrange inclusion (5.24). Using our continuous-time notation, we represent (5.24) in the following equivalent form:

$$\dot{p}_K(t) \in \{u \in \mathbf{R}^n \mid (u, p_K(t_{j+1}) + \lambda_{0K} \rho_K(t)) \in \lambda_{0K} \partial f(\bar{x}_K(t_j), \dot{\bar{x}}(t)) + \quad (6.16)$$

$$N((x_K(t), \dot{\bar{x}}(t)); \text{gph } F)\} \text{ for } t \in [t_j, t_{j+1}), \quad j = 0, \dots, K-1.$$

Employing the classical Masur theorem, we find a sequence of convex combinations of functions  $\dot{p}_k(t)$  on  $[0, \bar{T}]$  which converges to  $\dot{p}(t)$  for a.e.  $t \in [0, \bar{T}]$ . Now passing to the limit in (6.15) as  $K \rightarrow \infty$  and using (6.8) as well as the robustness of the subdifferential  $\partial f(\cdot)$  and the normal cone  $N(\cdot; \text{gph } F)$ , we establish the differential Euler-Lagrange inclusion (6.4).

It remains to prove the endpoint inclusion (6.5) which combines transversality conditions on  $(p(0), p(\bar{T}))$  with an additional condition on the optimal time interval  $[0, \bar{T}]$ . First let us consider the left-hand side of (5.25) as  $K \rightarrow \infty$ . Obviously

$$(p_K(0) + 2\lambda_{0K}(\bar{x}(0) - \bar{x}_K(0)), -p_K(\bar{T}_K), 2\lambda_{0K}(\bar{T} - \bar{T}_K)) \rightarrow (p(0), -p(\bar{T}), 0). \quad (6.17)$$

Considering  $\varrho_K$  in (5.29) and using (3.2), (5.26) as well as Riemann integrability of  $\dot{\bar{x}}(\cdot)$  and  $\tau(\dot{\bar{x}}, f) \rightarrow 0$  when  $h \rightarrow 0$ , one has

$$\varrho_K = \frac{1}{\bar{T}_K} \sum_{j=0}^{K-1} [j \int_{t_j}^{t_{j+1}} |\dot{\bar{x}}_K(t) - \dot{\bar{x}}(t_j)|^2 dt - (j+1) \int_{t_j}^{t_{j+1}} |\dot{\bar{x}}_K(t) - \dot{\bar{x}}(t_{j+1})|^2 dt] \sim \quad (6.18)$$

$$\frac{1}{\bar{T}_K} \sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} |\dot{\bar{x}}_K(t) - \dot{\bar{x}}(t)|^2 dt = \frac{1}{\bar{T}_K} \int_0^{\bar{T}_K} |\dot{\bar{x}}_K(t) - \dot{\bar{x}}(t)|^2 dt \rightarrow 0 \text{ as } K \rightarrow \infty$$

by virtue of the convergence (3.18) in Theorem 3.2.

Now let us evaluate  $\bar{H}_K$  in (5.27). Using (3.2), (5.26), and the convergences (3.16)–(3.18), (6.8) as well as  $p_K(\cdot) \rightarrow p(\cdot)$  uniformly in  $[0, \bar{T}]$ , we get

$$\bar{H}_K = \frac{1}{\bar{T}_K} \sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} [\langle p_K(t_{j+1}), \dot{\bar{x}}_K(t) \rangle - \lambda_{0K} f(\bar{x}_K(t_j), \dot{\bar{x}}_K(t))] dt \sim \quad (6.19)$$

$$\begin{aligned} & \frac{1}{\bar{T}_K} \sum_{j=0}^{K-1} \int_{t_j}^{t_{j+1}} [\langle p_K(t), \dot{\bar{x}}_K(t) \rangle - \lambda_{0K} f(\bar{x}_K(t), \dot{\bar{x}}_K(t))] dt = \frac{1}{\bar{T}_K} \int_0^{\bar{T}_K} [\langle p_K(t), \dot{\bar{x}}_K(t) \rangle - \\ & \lambda_{0K} f(\bar{x}_K(t), \dot{\bar{x}}_K(t))] dt \rightarrow \frac{1}{\bar{T}} \int_0^{\bar{T}} [\langle p(t), \dot{\bar{x}}(t) \rangle - \lambda_0 f(\bar{x}(t), \dot{\bar{x}}(t))] dt := \bar{H} \text{ as } K \rightarrow \infty \end{aligned}$$

due to the classical Lebesgue limiting theorem.

Finally let us consider the limit of the right-side expression in (5.25). First we note that

$$\text{Limsup}_{K \rightarrow \infty} \partial \left( \sum_{i=0}^{q+r} \lambda_{iK} \varphi_i \right) (\bar{x}_K(0), \bar{x}_K(\bar{T}_K), \bar{T}_K) = \partial \left( \sum_{i=0}^{q+r} \lambda_i \varphi_i \right) (\bar{x}(0), \bar{x}(\bar{T}), \bar{T}) \quad (6.20)$$

due to the robustness of the subdifferential (4.9). Then observe that the set  $\Omega_K$  in (3.14) is represented in the form

$$\Omega_K = \{(x_0, x_K, T) \in \mathbf{R}^{2n+1} \mid \text{dist}((x_0, x_K, T), \Omega) \leq \gamma_K\}. \quad (6.21)$$

Thus the geometric constraint (3.14) can be reduced to an inequality constraint with a Lipschitz continuous function. Now using Proposition 4.14 if  $(\bar{x}_K(0), \bar{x}_K(\bar{T}_K), \bar{T}_K) \in \Omega$  and just the definition of the normal cone (4.1) if  $(\bar{x}_K(0), \bar{x}_K(\bar{T}_K), \bar{T}_K) \notin \Omega$  in (6.21), we get

$$\text{Limsup}_{K \rightarrow \infty} N((\bar{x}_K(0), \bar{x}_K(\bar{T}_K), \bar{T}_K); \Omega_K) = N((\bar{x}(0), \bar{x}(\bar{T}), \bar{T}); \Omega). \quad (6.22)$$

Putting together relationships (6.17)–(6.20) and (6.22) we obtain the endpoint inclusion (6.5) by passing to the limit in its discrete counterpart (5.25). This ends the proof of the theorem.  $\square$

6.2. COROLLARY. *In the assumptions of Theorem 6.1, let the functions  $\varphi_i$  and the set  $\Omega$  do not depend on the  $T$  variable. Then one has all the conclusions of the theorem with changing (6.5) for*

$$(p(0), -p(\bar{T})) \in \partial\left(\sum_{i=0}^{q+r} \lambda_i \varphi_i\right)(\bar{x}(0), \bar{x}(\bar{T})) + N(\bar{x}(0), \bar{x}(\bar{T}); \Omega) \quad \text{and} \quad (6.23)$$

$$\int_0^{\bar{T}} [\langle p(t), \dot{\bar{x}}(t) \rangle - \lambda_0 f(\bar{x}(t), \dot{\bar{x}}(t))] dt = 0. \quad (6.24)$$

*Proof.* This immediately follows from (6.5) and (6.6) under the additional assumption made.  $\square$

6.3. Remark. The integrand

$$H(t) := \langle p(t), \dot{\bar{x}}(t) \rangle - \lambda_0 f(\bar{x}(t), \dot{\bar{x}}(t)), \quad 0 \leq t \leq \bar{T}, \quad (6.25)$$

in (6.6) and (6.24) is the *pseudo-Hamiltonian* of the Bolza problem ( $P$ ) calculated on the optimal solution  $\{\bar{x}(\cdot), \bar{T}\}$  and the corresponding adjoint pair  $\{p(\cdot), \lambda_0\}$ . If we have the *Weierstrass-Pontryagin maximum condition* in the problem under consideration, then

$$H(t) = \max\{\langle p(t), v \rangle - \lambda_0 f(\bar{x}(t), v)\} \quad \text{a.e. } t \in [0, \bar{T}], \quad (6.26)$$

i.e., the values of the pseudo-Hamiltonian (6.25) coincides with the corresponding values of the (maximized) *Hamiltonian*

$$\mathcal{H}(x, p, \lambda_0) := \max\{\langle p, v \rangle - \lambda_0 f(x, v) \mid v \in F(x)\}$$

along optimal processes. The latter always holds (actually it follows from the Euler-Lagrange condition (6.4)) for *convex* problems ( $P$ ) when the sets  $F(x)$  are convex and the function  $f$  is convex in the  $v$  variable.

Moreover, it is well known that for classical autonomous problems in the calculus of variations and optimal optimal control the pseudo-Hamiltonian (6.25) is *constant*:

$$H(t) \equiv c \quad \text{on } [0, \bar{T}] \quad (6.27)$$



where  $c = 0$  if endpoint constraints do not depend on  $T$ . An analogue of the constancy relation (6.27) for the maximized Hamiltonian  $\mathcal{H}$  is proved by Clarke [11] in the framework of his Hamiltonian conditions for Mayer problems involving convex differential inclusions.

One can easily see that in the case of (6.26) and (6.27), the number  $\bar{H}$  in (6.5) and (6.6) is equal to the value of the Hamiltonian at any point  $t \in [0, \bar{T}]$ . In the general nonconvex setting of Theorem 6.1, the number  $\bar{H}$  appears to be the *averaged value* of the pseudo-Hamiltonian on the optimal interval  $[0, \bar{T}]$ .

6.4. *Remark.* It is essential in Theorem 6.1 that  $\{\bar{x}(\cdot), \bar{T}\}$  is a *feasible* solution to the *original nonconvex* problem  $(P)$  but not to its relaxation. Indeed, the Euler-Lagrange inclusion (6.4) for  $(P)$  may be different from its counterpart for the relaxed problem  $(R)$ . It happens because the normal cone to the graph of  $F$  and to the graph of its convex hull are not the same (as well as the subdifferentials for  $f$  and  $\hat{f}_F$ ).

Theorem 6.1 provides necessary optimality conditions in the Bolza problem  $(P)$  for autonomous differential inclusions with the property of relaxation stability. Now we consider a corollary of the theorem where the latter property is automatically fulfilled.

6.5. COROLLARY. *Let  $\{\bar{x}(\cdot), \bar{T}\}$  be an optimal solution to the Bolza problem (2.1), (2.3)–(2.5) with no differential inclusion. Suppose that for some numbers  $\mu > 0$  and open set  $U \subset \mathbf{R}^n$  one has:*

$$\bar{x}(t) \in U \quad \forall t \in [0, \bar{T}] \quad \& \quad |\dot{\bar{x}}(t)| < \mu \quad \text{a.e. } t \in [0, \bar{T}], \quad (6.28)$$

*the set  $\Omega$  is closed, the functions  $\varphi_i$  are locally Lipschitz around  $(\bar{x}(0), \bar{x}(\bar{T}), \bar{T})$ , and the function  $f = f(x, v)$  is Lipschitz continuous around any point  $(x, v) \in U \times (\mu B)$ . Then there exist real numbers  $\lambda_0, \dots, \lambda_{q+r}$  and an absolutely continuous function  $p : [0, \bar{T}] \rightarrow \mathbf{R}^n$ , not all zero, such that conditions (6.2), (6.3), and (6.5) are fulfilled and*

$$\dot{p}(t) \in \text{co}\{u \mid (u, p(t)) \in \lambda_0 \partial f(\bar{x}(t), \dot{\bar{x}}(t))\} \quad \text{a.e. } t \in [0, \bar{T}]. \quad (6.29)$$

*Proof.* It follows from (6.28) that  $\bar{x}(\cdot) \in W^{1, \infty}[0, \bar{T}]$ . Using Proposition 2.2, we conclude that the problem under consideration possesses the property of relaxation stability. In fact, this problem is equivalent to the problem  $(P)$  with the (trivial) differential inclusion

$$\dot{x} \in F(x) := \mu B \quad t \in [0, T] \quad (6.30)$$

where  $(\bar{x}(t), \dot{\bar{x}}(t)) \in (\text{int gph } F)$  for a.e.  $t \in [0, \bar{T}]$ . Obviously, all the assumptions in Theorem 6.1 are fulfilled for problem (2.1), (2.3)–(2.5), (6.30) and the Euler-Lagrange inclusion (6.4) is equivalent to (6.29). This proves the corollary.  $\square$

In conclusion of this section, we consider problem  $(P)$  with a non-autonomous differential inclusion where  $F(x, t)$  is Lipschitz continuous with respect to both variables. Such a problem can be reduced to the autonomous case treated above.

6.6. THEOREM. *Let  $\{\bar{x}(\cdot), \bar{T}\}$  be an optimal solution to problem  $(P)$  in (2.1)–(2.5) which possesses the property of relaxation stability. In addition to (H1) and (H7), we suppose that  $F$  is Lipschitz continuous in  $(t, x)$  jointly,  $\varphi_i$  are Lipschitz continuous around  $(\bar{x}(0), \bar{x}(\bar{T}), \bar{T})$ , and  $f$  is Lipschitz continuous around any point of the set*

$$\tilde{A} := \{(t, x, v) \in \mathbf{R}^{2n+1} \mid (t, x) \in [0, \bar{T}] \times U, v \in F(t, x)\}.$$

*Then there exist real numbers  $\lambda_0, \dots, \lambda_{q+r}$  as well as absolutely continuous functions  $p^0 : [0, \bar{T}] \rightarrow \mathbf{R}$  and  $p : [0, \bar{T}] \rightarrow \mathbf{R}^n$  such that one has (6.2), (6.3),*

$$\{\lambda_0, \dots, \lambda_{q+r}, p(\cdot)\} \neq 0; \tag{6.31}$$

$$(\dot{p}^0(t), \dot{p}(t)) \in \text{co}\{(u^0, u) \in \mathbf{R}^{n+1} \mid (u^0, u, p(t)) \in \lambda_0 \partial f(t, \bar{x}(t), \dot{\bar{x}}(t)) + \tag{6.32}$$

$$N((t, \bar{x}(t), \dot{\bar{x}}(t)); \text{gph } F) \text{ a.e. } t \in [0, \bar{T}];$$

$$(p(0), -p(\bar{T}), -p^0(\bar{T})) \in \partial(\sum_{i=0}^{q+r} \lambda_i \varphi_i)(\bar{x}(0), \bar{x}(\bar{T}), \bar{T}) + \tag{6.33}$$

$$N((\bar{x}(0), \bar{x}(\bar{T}), \bar{T}); \Omega), \quad p^0(0) = 0;$$

$$\int_0^{\bar{T}} p^0(t) dt = - \int_0^{\bar{T}} H(t) dt \tag{6.34}$$

where the pseudo-Hamiltonian  $H(t)$  is defined in (6.25) with  $f = f(t, \bar{x}(t), \dot{\bar{x}}(t))$ .

*Proof.* Let us denote  $t := x^0$  and introduce a new state variable  $z = (x^0, x)$ , corresponding endpoint variables  $z^0 = (x_0^0, x_0)$  and  $z_T = (x_T^0, x_T)$ , and a velocity variable  $w = (\vartheta, v)$  in  $\mathbf{R}^{n+1}$ . We define

$$\tilde{F}(z) := (1, F(x^0, x)), \quad \tilde{f}(z, w) := f(x^0, x, v), \tag{6.35}$$

$$\tilde{\varphi}_i(z_0, z_T) := \varphi_i(x_0, x_T, x_T^0), \quad \tilde{\Omega} := \{(z_0, z_T) \mid x_0^0 \in \mathbf{R}, (x_0, x_T, x_T^0) \in \Omega\}. \quad (6.36)$$

Then one can see that the arc  $\bar{z}(t) = \{(t, \bar{x}(t)), 0 \leq t \leq \bar{T}\}$  is an optimal solution to the  $n + 1$ -dimensional autonomous Bolza problem:

$$\text{minimize } \tilde{J}[z, T] := \tilde{\varphi}_0(z(0), z(T)) + \int_0^T \tilde{f}(z(t), \dot{z}(t)) dt$$

over all trajectories of the differential inclusion

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

subject to the endpoint constraints

$$\tilde{\varphi}_i(z(0), z(T)) \leq 0 \quad \text{for } i = 1, \dots, q;$$

$$\tilde{\varphi}_i(z(0), z(T)) = 0 \quad \text{for } i = q + 1, \dots, q + r;$$

$$(z(0), z(T)) \in \tilde{\Omega} \in \mathbf{R}^{2n+2}.$$

Under the hypotheses of the theorem, the autonomous problem formulated satisfies all the assumptions of Corollary 6.2. Using this corollary and taking into account the special structure of the data (6.35) and (6.36), we get conditions (6.32) and (6.33) directly from (6.4) and (6.23). In this way, (6.34) follows from (6.24) and (6.25), and one can ensure

$$\{\lambda_0, \dots, \lambda_{q+r}, p^0(\cdot), p(\cdot)\} \neq 0. \quad (6.37)$$

Suppose that (6.31) does not hold, i.e.,  $\lambda_0, \dots, \lambda_{q+r}$  are equal to zero simultaneously with  $p(\cdot)$ . Then due to (4.4) and Corollary 4.8, (6.32) implies that  $\dot{p}^0(t) = \dot{p}(t) = 0$  a.e.  $t \in [0, \bar{T}]$ . By virtue of the transversality condition  $p^0(0) = 0$ , now we can conclude that  $p^0(t) \equiv 0$  on  $[0, \bar{T}]$ . This contradicts (6.37) and completes the proof of the theorem.  $\square$

6.7. *Remark.* The additional assertion (6.27) in the framework of Corollary 6.2 would imply that

$$p^0(t) \equiv -H(t) \quad \text{on } [0, \bar{T}]$$

and conditions (6.32) and (6.33) are reduced to the forms

$$(-\dot{H}(t), \dot{p}(t)) \in \text{co}\{(u^0, u) \in \mathbf{R}^{n+1} \mid (u^0, u, p(t)) \in \lambda_0 \partial f(t, \bar{x}(t), \dot{\bar{x}}(t)) +$$

$$N((t, \bar{x}(t), \dot{\bar{x}}(t)); \text{gph } F) \text{ a.e. } t \in [0, \bar{T}];$$

$$(p(0), -p(\bar{T}), H(\bar{T})) \in \partial\left(\sum_{i=0}^{q+r} \lambda_i \varphi_i\right)(\bar{x}(0), \bar{x}(\bar{T}), \bar{T}) + N((\bar{x}(0), \bar{x}(\bar{T}), \bar{T}); \Omega).$$

## 7. Concluding Remarks

Now we discussed some possible improvements and generalizations of the principal results obtained in this paper.

7.1. *Remark.* In the main body of the paper, we have used a discrete approximation procedure to obtain necessary optimality conditions in the constrained problem of Bolza for differential inclusions with free time. This method allows us to prove the refined Euler-Lagrange conditions for the original problem possessing the property of relaxation stability. An important question consists of discovering settings when the *property of relaxation stability can be removed*.

In the previous paper [43], we removed the latter property for the case of general (non-convex) Mayer problems with fixed time. To accomplish it, we developed another procedure for approximating the original problem by a parametric family of unconstrained (but non-smooth) problems of Bolza with no differential inclusion. This procedure is based on the metric approximation method in Mordukhovich [31, 33, 35] and the usage of Ekeland's variational principle as in Clarke [11] and Kaskosz and Lojasiewicz [27]. In this way, besides removing the property of relaxation stability, we obtained [43] more general transversality conditions with and without Lipschitzian assumptions on  $\varphi_i$ .

To develop the latter approximation procedure for non-fixed time problems, one needs to employ necessary optimality conditions as in Corollary 6.5 for unconstrained problems of Bolza with optimal solutions belonging to  $W^{1,\infty}$ . However, the results available in Corollary 6.5 are proved so far for the case of optimal solutions whose derivatives are Riemann integrable. This difference seems to be minor from the practical viewpoint but it plays an important role for the realization of the mentioned approximation procedure.

Thus any progress in justifying the results in Corollary 6.5 for the case of  $W^{1,\infty}$ -optimal solutions allows us to remove the property of relaxation stability in the necessary conditions

of Theorem 6.1 for general problems of Mayer. The same is true for *local controllability* and related results; cf. [43].

7.2. *Remark.* The discrete approximation procedure and necessary optimality conditions are proved above for global optimal solutions of the Bolza problem ( $P$ ). Obviously, they hold for *strong local minima* in the space of trajectories  $x(\cdot) = \{x(t), 0 \leq t \leq T\}$  with the  $C$ -norm

$$\|x(\cdot)\|_C := T + \max_{t \in [0, T]} |x(t)|.$$

Moreover, the method developed enables us to study the so-called *intermediate local minimum* with respect to the  $W^{1,p}$ -norm

$$\|x(\cdot)\|_{W^{1,p}} := T + \max_{t \in [0, T]} |x(t)| + \left( \int_0^T |\dot{x}(t)|^p dt \right)^{1/p}$$

for any  $p \in [1, \infty)$ . This notion has been considered in [43] for fixed-time problems involving differential inclusions. Clearly, it takes an intermediate place between the classical concepts of strong and weak ( $p = \infty$ ) local minima.

The reader can check that the constructions of the paper allow us to keep the results obtained for feasible solutions to the original problem ( $P$ ) which provide an intermediate local minimum for the relaxed problem ( $R$ ) with the same value of the cost functional; cf. [43] in the setting of fixed-time problems.

7.3. *Remark.* Developing the approach of the paper, we can consider a generalization of the Bolza problem ( $P$ ) where constraints (2.3) and (2.4) are replaced by the following:

$$\varphi_i(x(0), x(T), T) + \int_0^T f_i(t, x(t), \dot{x}(t)) dt \leq 0 \quad \text{for } i = 1, \dots, q; \quad (7.1)$$

$$\varphi_i(x(0), x(T), T) + \int_0^T f_i(t, x(t), \dot{x}(t)) dt = 0 \quad \text{for } i = q + 1, \dots, q + r. \quad (7.2)$$

Note that new constraints (7.1) and (7.2) unify the previous endpoint constraints (2.3) and (2.4) with the *isoperimetric* type constraints in variational problems.

The constructions of Section 3 allow us to build a sequence of discrete approximations for problem (2.1), (2.2), (2.5), (7.1), and (7.2) whose solutions strongly  $W^{1,2}$ -converge to the reference optimal solution to the original problem under the corresponding property of

relaxation stability. Following the procedure in Sections 5 and 6, we obtain an analogue of the main Theorem 6.1 where conditions (6.3) (6.4), and (6.6) are changed for, respectively,

$$\lambda_i[\varphi_i(\bar{x}(0), \bar{x}(\bar{T}), \bar{T}) - \int_0^{\bar{T}} f_i(\bar{x}(t), \dot{\bar{x}}(t))dt] = 0 \text{ for } i = 1, \dots, q;$$

$$\dot{p}(t) \in \text{co}\{u \mid (u, p(t)) \in \partial(\sum_{i=0}^{q+r} \lambda_i f_i(\bar{x}(t), \dot{\bar{x}}(t)) + N((\bar{x}(t), \dot{\bar{x}}(t)); \text{gph } F))\}$$

for a.e.  $t \in [0, \bar{T}]$ ; and

$$\bar{H} = \frac{1}{\bar{T}} \int_0^{\bar{T}} [\langle p(t), \dot{\bar{x}}(t) \rangle - \sum_{i=0}^{q+r} \lambda_i f_i(\bar{x}(t), \dot{\bar{x}}(t))] dt$$

with  $f_0 := f$ .

7.4. *Remark.* The discrete approximation approach allows us to extend the results of this paper to problems of Bolza where the functions  $\varphi_i$  depend not only on endpoint state positions and a varying time interval but also on *intermediate times*  $\tau_s \in (0, T)$  and *intermediate states*  $x(\tau_s)$ ; cf. Clarke and Vinter [13, 14] in the framework of optimal multiprocesses. Our results for such (nonconvex) problems combine the refined Euler-Lagrange inclusion (6.4) with a new condition in the line of (6.5) involving jumps of the adjoint arc.

## References

1. J.-P. Aubin and A. Cellina, *Differential Inclusions*, Springer-Verlag, Berlin, 1984.
2. J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.
3. L. D. Berkovitz, *Optimal Control Theory*, Springer-Verlag, New York, 1974.
4. N. N. Bogoljubov, *Sur quelques methods nouvelles dans le calcul des variations*, Ann. Math. Pura Appl. **7** (1930), 249–271.
5. J. M. Borwein and D. M. Zhuang, *Verified necessary and sufficient conditions for regularity of set-valued and single-valued maps*, J. Math. Anal. Appl. **134** (1988), 441–459.

6. J. V. Burke, *Calmness and exact penalization*, SIAM J. Control Optim. **29** (1991), 493–497.
7. F. H. Clarke, *Admissible relaxation in variational and control problems*, J. Math. Anal. Appl. **51** (1975), 557–576.
8. F. H. Clarke, *A new approach to Lagrange multipliers*, Math. Oper. Res. **1** (1976), 165–174.
9. F. H. Clarke, *The generalized problem of Bolza*, SIAM J. Control Optim. **14** (1976), 682–699.
10. F. H. Clarke, *Optimal solutions to differential inclusions*, J. Optim. Theory Appl. **19** (1976), 469–478.
11. F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.
12. F. H. Clarke, *Methods of Dynamic and Nonsmooth Optimizations*, SIAM Publications, Philadelphia, PA, 1989.
13. F. H. Clarke and R. B. Vinter, *Optimal multiprocesses*, SIAM J. Control Optim. **27** (1989), 1072–1091.
14. F. H. Clarke and R. B. Vinter, *Applications of optimal multiprocesses*, SIAM J. Control Optim. **27** (1989), 1048–1071.
15. F. H. Clarke, P. D. Loewen, and R. B. Vinter, *Differential inclusions with free time*, Ann. Inst. H. Poincaré Anal. Non Linéaire **5** (1989), 573–593.
16. M. G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order differential equations*, Bull. Amer. Math. Soc. **27** (1992), 1–67.
17. A. L. Dontchev, *Discrete approximations in optimal control*, this volume.
18. A. L. Dontchev and E. M. Farkhi, *Error estimates for discretized differential inclusions*, Computing **41** (1989), 349–358.

19. A. L. Dontchev and F. Lempio, *Difference methods for differential inclusions: a survey*, SIAM Review **34** (1992), 263–294.
20. A. L. Dontchev and T. Zolezzi, *Well-Posed Optimization Problems*, Springer-Verlag, Berlin, 1993.
21. I. Ekeland and R. Teman, *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, 1976.
22. A. D. Ioffe, *Approximate subdifferentials and applications. I: The finite dimensional theory*, Trans. Amer. Math. Soc. **281** (1984), 389–416.
23. A. D. Ioffe, *Approximate subdifferentials and applications. III: The metric theory*, Mathematika **36** (1989), 1–38.
24. A. D. Ioffe, *Nonsmooth subdifferentials: their calculus and applications*, Proceed. 1st World Congress of Nonlinear Analysts, Tampa, FL, August 1992. (to appear)
25. A. D. Ioffe and V. M. Tikhomirov, *Extensions of variational problems*, Trans. Moscow Math. Soc. **18** (1968), 207–273.
26. A. D. Ioffe and V. M. Tikhomirov, *Theory of Extremal Problems*, North-Holland, Amsterdam, 1979.
27. B. Kaskosz and S. Lojasiewicz Jr., *Lagrange-type extremal trajectories in differential inclusions*, Systems Control Lett. **19** (1992), 241–247.
28. P. D. Loewen, *Optimal Control via Nonsmooth Analysis*, CRM Monograph Series, University of Montreal, Montreal, 1993.
29. P. D. Loewen and R. T. Rockafellar, *Optimal control of unbounded differential inclusions*, SIAM J. Control Optim. (to appear)
30. A. A. Lyapounov, *On complete additive vector functions*, Izvest. AN SSSR, Ser. Math. **4** (1940), 465–478. (Russian)



31. B. Mordukhovich, *Maximum principle in the problem of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech. **40** (1976), 960–969.
32. B. Mordukhovich, *On difference approximations of optimal control systems*, J. Appl. Math. Mech. **42** (1978), 452–461.
33. B. Mordukhovich, *Metric approximations and necessary optimality conditions for general classes of nonsmooth extremal problems*, Soviet Math. Dokl. **22** (1980), 526–530.
34. B. Mordukhovich, *On necessary conditions for an extremum in nonsmooth optimization*, Soviet Math. Dokl. **32** (1985), 215–220.
35. B. Mordukhovich, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988. (Russian; English translation to appear in Wiley-Interscience.)
36. B. Mordukhovich, *Optimization and approximation of differential inclusions*, Cybernetics **24** (1988), 781–788.
37. B. Mordukhovich, *Necessary optimality conditions for nonsmooth control problems with non-fixed time*, Different. Equations **25** (1989), 290–299.
38. B. Mordukhovich, *Sensitivity analysis in nonsmooth optimization*, in Theoretical Aspects of Industrial Design (D. A. Field and V. Komkov, eds.), SIAM Publications, Philadelphia, PA, 1992, 32–47.
39. B. Mordukhovich, *On variational analysis of differential inclusions*, in Optimization and Nonlinear Analysis (A. Ioffe, L. Marcus, and S. Reich, eds.), Pitman Res. Notes Math. Ser. **244** (1992), 199–213.
40. B. Mordukhovich, *Complete characterizations of openness, metric regularity, and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc. (to appear)
41. B. Mordukhovich, *Generalized differential calculus for nonsmooth and set-valued mappings*, J. Math. Anal. Appl. (to appear)
42. B. Mordukhovich, *Stability theory for parametric generalized equations and variational inequalities*, IMA Preprint Series # 997, July 1992.

43. B. Mordukhovich, *Discrete approximations and refined Euler-Lagrange conditions for nonconvex differential inclusions*, IMA Preprint Series # 1115, March 1993.
44. J.-P. Penot, *Metric regularity, openness and Lipschitzian behavior for multifunctions*, *Nonlinear Anal. Theory Methods Appl.* **13** (1989), 629–643.
45. E. Polak, *On the use of consistent approximations in the solution of semi-infinite optimization and optimal control problems*, *SIAM J. Control Optim.* (to appear)
46. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko, *The Mathematical Theory of Optimal Processes*, Wiley-Interscience, New York, 1962.
47. B. N. Pshenichnyi, *Convex Analysis and Extremal Problems*, Nauka, Moscow, 1980. (Russian)
48. R. T. Rockafellar, *Extensions of subgradient calculus with applications to optimization*, *Nonlinear Anal. Theory Methods Appl.* **9** (1985), 665–698.
49. R. T. Rockafellar, *Lipschitzian properties of multifunctions*, *Nonlinear Anal. Theory Methods Appl.* **9** (1985), 867–885.
50. R. T. Rockafellar, *Maximal monotone relations and the second derivatives of nonsmooth functions*, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **2** (1985), 167–184.
51. R. T. Rockafellar, *Dualization of subgradient conditions for optimality*, *Nonlinear Anal. Theory Methods Appl.* **20** (1993), 627–646.
52. R. T. Rockafellar, *Lagrange multipliers and optimality*, *SIAM Review* **35** (1993), 183–238.
53. R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, book in preparation.
54. J. D. L. Rowland and R. B. Vinter, *Dynamic optimization problems with free time and active state constraints*, *SIAM J. Control Optim.* **31** (1993), 677–697.
55. B. Sendov and V. A. Popov, *The Averaged Moduli of Smoothness*, Wiley-Interscience, New York, 1988.

56. G. V. Smirnov, *Discrete approximations and optimal solutions to differential inclusions*, Kibernetika, 1991, 76–79. (Russian)
57. J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.
58. J. Warga, *Controllability, extremality, and abnormality in nonsmooth optimal control*, J. Optim. Theory Appl. **41** (1983), 239–260.
59. L. C. Young, *Lectures on the Calculus of Variations and Optimal Control Theory*, Saunders, Philadelphia, PA, 1969.

#	Author/s	Title
1090	<b>Tongwen Chen &amp; Li Qiu</b> , $\mathcal{H}_\infty$ design of general multirate sampled-data control systems	
1091	<b>Satyanad Kichenassamy &amp; Walter Littman</b> , Blow-up surfaces for nonlinear wave equations, I	
1092	<b>Nahum Shimkin</b> , Asymptotically efficient adaptive strategies in repeated games, Part I: certainty equivalence strategies	
1093	<b>Caroline Fabre, Jean-Pierre Puel &amp; Enrique Zuazua</b> , On the density of the range of the semigroup for semilinear heat equations	
1094	<b>Robert F. Stengel, Laura R. Ray &amp; Christopher I. Marrison</b> , Probabilistic evaluation of control system robustness	
1095	<b>H.O. Fattorini &amp; S.S. Sritharan</b> , Optimal chattering controls for viscous flow	
1096	<b>Kathryn E. Lenz</b> , Properties of certain optimal weighted sensitivity and weighted mixed sensitivity designs	
1097	<b>Gang Bao &amp; David C. Dobson</b> , Second harmonic generation in nonlinear optical films	
1098	<b>Avner Friedman &amp; Chaocheng Huang</b> , Diffusion in network	
1099	<b>Xinfu Chen, Avner Friedman &amp; Tsuyoshi Kimura</b> , Nonstationary filtration in partially saturated porous media	
1100	<b>Walter Littman &amp; Baisheng Yan</b> , Rellich type decay theorems for equations $P(D)u = f$ with $f$ having support in a cylinder	
1101	<b>Satyanad Kichenassamy &amp; Walter Littman</b> , Blow-up surfaces for nonlinear wave equations, II	
1102	<b>Nahum Shimkin</b> , Extremal large deviations in controlled I.I.D. processes with applications to hypothesis testing	
1103	<b>A. Narain</b> , Interfacial shear modeling and flow predictions for internal flows of pure vapor experiencing film condensation	
1104	<b>Andrew Teel &amp; Laurent Praly</b> , Global stabilizability and observability imply semi-global stabilizability by output feedback	
1105	<b>Karen Rudie &amp; Jan C. Willems</b> , The computational complexity of decentralized discrete-event control problems	
1106	<b>John A. Burns &amp; Ruben D. Spies</b> , A numerical study of parameter sensitivities in Landau-Ginzburg models of phase transitions in shape memory alloys	
1107	<b>Gang Bao &amp; William W. Symes</b> , Time like trace regularity of the wave equation with a nonsmooth principal part	
1108	<b>Lawrence Markus</b> , A brief history of control	
1109	<b>Richard A. Brualdi, Keith L. Chavey &amp; Bryan L. Shader</b> , Bipartite graphs and inverse sign patterns of strong sign-nonsingular matrices	
1110	<b>A. Kersch, W. Morokoff &amp; A. Schuster</b> , Radiative heat transfer with quasi-monte carlo methods	
1111	<b>Jianhua Zhang</b> , A free boundary problem arising from swelling-controlled release processes	
1112	<b>Walter Littman &amp; Stephen Taylor</b> , Local smoothing and energy decay for a semi-infinite beam pinned at several points and applications to boundary control	
1113	<b>Srdjan Stojanovic &amp; Thomas Svobodny</b> , A free boundary problem for the Stokes equation via nonsmooth analysis	
1114	<b>Bronislaw Jakubczyk</b> , Filtered differential algebras are complete invariants of static feedback	
1115	<b>Boris Mordukhovich</b> , Discrete approximations and refined Euler-Lagrange conditions for nonconvex differential inclusions	
1116	<b>Bei Hu &amp; Hong-Ming Yin</b> , The profile near blowup time for solution of the heat equation with a nonlinear boundary condition	
1117	<b>Jin Ma &amp; Jiongmin Yong</b> , Solvability of forward-backward SDEs and the nodal set of Hamilton-Jacobi-Bellman Equations	
1118	<b>Chaocheng Huang &amp; Jiongmin Yong</b> , Coupled parabolic and hyperbolic equations modeling age-dependent epidemic dynamics with nonlinear diffusion	
1119	<b>Jiongmin Yong</b> , Necessary conditions for minimax control problems of second order elliptic partial differential equations	
1120	<b>Eitan Altman &amp; Nahum Shimkin</b> , Worst-case and Nash routing policies in parallel queues with uncertain service allocations	
1121	<b>Nahum Shimkin &amp; Adam Shwartz</b> , Asymptotically efficient adaptive strategies in repeated games, part II: Asymptotic optimality	
1122	<b>M.E. Bradley</b> , Well-posedness and regularity results for a dynamic Von Kármán plate	
1123	<b>Zhangxin Chen</b> , Finite element analysis of the 1D full drift diffusion semiconductor model	
1124	<b>Gang Bao &amp; David C. Dobson</b> , Diffractive optics in nonlinear media with periodic structure	
1125	<b>Steven Cox &amp; Enrique Zuazua</b> , The rate at which energy decays in a damped string	
1126	<b>Anthony W. Leung</b> , Optimal control for nonlinear systems of partial differential equations related to ecology	
1127	<b>H.J. Sussmann</b> , A continuation method for nonholonomic path-finding problems	
1128	<b>Yung-Jen Guo &amp; Walter Littman</b> , The null boundary controllability for semilinear heat equations	
1129	<b>Q. Zhang &amp; G. Yin</b> , Turnpike sets in stochastic manufacturing systems with finite time horizon	
1130	<b>I. Györi, F. Hartung &amp; J. Turi</b> , Approximation of functional differential equations with time- and state-dependent delays by equations with piecewise constant arguments	

- 1131 **I. Györi, F. Hartung & J. Turi**, Stability in delay equations with perturbed time lags
- 1132 **F. Hartung & J. Turi**, On the asymptotic behavior of the solutions of a state-dependent delay equation
- 1133 **Pierre-Alain Gremaud**, Numerical optimization and quasiconvexity
- 1134 **Jie Tai Yu**, Resultants and inversion formula for  $N$  polynomials in  $N$  variables
- 1135 **Avner Friedman & J.L. Velázquez**, The analysis of coating flows in a strip
- 1136 **Eduardo D. Sontag**, Control of systems without drift via generic loops
- 1137 **Yuan Wang & Eduardo D. Sontag**, Orders of input/output differential equations and state space dimensions
- 1138 **Scott W. Hansen**, Boundary control of a one-dimensional, linear, thermoelastic rod
- 1139 **Robert Lipton & Bogdan Vernescu**, Homogenization of two phase emulsions with surface tension effects
- 1140 **Scott Hansen & Enrique Zuazua**, Exact controllability and stabilization of a vibrating string with an interior point mass
- 1141 **Bei Hu & Jiongmin Yong**, Pontryagin Maximum principle for semilinear and quasilinear parabolic equations with pointwise state constraints
- 1142 **Mark H.A. Davis**, A deterministic approach to optimal stopping with application to a prophet inequality
- 1143 **M.H.A. Davis & M. Zervos**, A problem of singular stochastic control with discretionary stopping
- 1144 **Bernardo Cockburn & Pierre-Alain Gremaud**, An error estimate for finite element methods for scalar conservation laws
- 1145 **David C. Dobson & Fadi Santosa**, An image enhancement technique for electrical impedance tomography
- 1146 **Jin Ma, Philip Protter, & Jiongmin Yong**, Solving forward-backward stochastic differential equations explicitly — a four step scheme
- 1147 **Yong Liu**, The equilibrium plasma subject to skin effect
- 1148 **Ulrich Hornung**, Models for flow and transport through porous media derived by homogenization
- 1149 **Avner Friedman, Chaocheng Huang, & Jiongmin Yong**, Effective permeability of the boundary of a domain
- 1150 **Gang Bao**, A uniqueness theorem for an inverse problem in periodic diffractive optics
- 1151 **Angelo Favini, Mary Ann Horn, & Irena Lasiecka**, Global existence and uniqueness of regular solutions to the dynamic von Kármán system with nonlinear boundary dissipation
- 1152 **E.G. Kalnins & Willard Miller, Jr.**, Models of  $q$ -algebra representations:  $q$ -integral transforms and “addition theorems”
- 1153 **E.G. Kalnins, V.B. Kuznetsov & Willard Miller, Jr.**, Quadrics on complex Riemannian spaces of constant curvature, separation of variables and the Gaudin magnet
- 1154 **A. Kersch, W. Morokoff & Chr. Werner**, Selfconsistent simulation of sputtering with the DSMC method
- 1155 **Bing-Yu Zhang**, A remark on the Cauchy problem for the Korteweg-de Vries equation on a periodic domain
- 1156 **Gang Bao**, Finite element approximation of time harmonic waves in periodic structures
- 1157 **Tao Lin & Hong Wang**, Recovering the gradients of the solutions of second-order hyperbolic equations by interpolating the finite element solutions
- 1158 **Zhangxin Chen**,  $L^p$ -posteriori error analysis of mixed methods for linear and quasilinear elliptic problems
- 1159 **Todd Arbogast & Zhangxin Chen**, Homogenization of compositional flow in fractured porous media
- 1160 **L. Qiu, B. Bernhardsson, A. Rantzer, E.J. Davison, P.M. Young & J.C. Doyle**, A formula for computation of the real stability radius
- 1161 **Maria Inés Tropsarevsky**, Adaptive control of linear discrete time systems with external disturbances under inaccurate modelling: A case study
- 1162 **Petr Klouček & Franz S. Rys**, Stability of the fractional step  $\Theta$ -scheme for the nonstationary Navier-Stokes equations
- 1163 **Eduardo Casas, Luis A. Fernández & Jiongmin Yong**, Optimal control of quasilinear parabolic equations
- 1164 **Darrell Duffie, Jin Ma & Jiongmin Yong**, Black’s consol rate conjecture
- 1165 **D.G. Aronson & J.L. Vazquez**, Anomalous exponents in nonlinear diffusion
- 1166 **Ruben D. Spies**, Local existence and regularity of solutions for a mathematical model of thermomechanical phase transitions in shape memory materials with Landau-Ginzburg free energy
- 1167 **Pu Sun**, On circular pipe Poiseuille flow instabilities
- 1168 **Angelo Favini, Mary Ann Horn, Irena Lasiecka & Daniel Tataru**, Global existence, uniqueness and regularity of solutions to a Von Kármán system with nonlinear boundary dissipation
- 1169 **A. Dontchev, Tz. Donchev & I. Slavov**, On the upper semicontinuity of the set of solutions of differential inclusions with a small parameter in the derivative
- 1170 **Jin Ma & Jiongmin Yong**, Regular-singular stochastic controls for higher dimensional diffusions — dynamic programming approach
- 1171 **Alex Solomonoff**, Bayes finite difference schemes
- 1172 **Todd Arbogast & Zhangxin Chen**, On the implementation of mixed methods as nonconforming methods for second order elliptic problems
- 1173 **Zhangxin Chen & Bernardo Cockburn**, Convergence of a finite element method for the drift-diffusion semiconductor device equations: The multidimensional case
- 1174 **Boris Mordukhovich**, Optimization and finite difference approximations of nonconvex differential inclusions with free time