DISCRETE APPROXIMATIONS AND REFINED EULER-LAGRANGE CONDITIONS FOR NONCONVEX DIFFERENTIAL INCLUSIONS

By

Boris Mordukhovich

DISCRETE APPROXIMATIONS AND REFINED EULER-LAGRANGE CONDITIONS FOR NONCONVEX DIFFERENTIAL INCLUSIONS [1]

BORIS MORDUKHOVICH [2]

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

February 1993

Abstract. This paper deals with the Bolza problem (P) for differential inclusions subject to general endpoint constraints. We pursue a twofold goal. First, we develop a finite difference method for studying (P) and construct a discrete approximation to (P) ensuring a strong convergence of optimal solutions. Second, we use this direct method for obtaining necessary optimality conditions in a refined Euler-Lagrange form without standard convexity assumptions. In general, we prove necessary conditions for the so-called intermediate relaxed local minimum which takes an intermediate place between the classical concepts of strong and weak minima. In the case of strong minima for Mayer problems or boundary solutions to differential inclusions, this Euler-Lagrange form holds without any relaxation. The results obtained are expressed in terms of nonconvex-valued generalized differentiation constructions for nonsmooth mappings and sets.

Key words. discrete approximations, differential inclusions, nonsmooth analysis, generalized differentiation, Euler-Lagrange conditions

1991 Mathematical Subject Classification. 49K24, 49J52, 49M25

Short Title. Discrete Approximations

---

1. Introduction

This paper is mainly concerned with the following problem of dynamic optimization: minimize the Bolza functional

$$(1.1) \qquad J[x] := \varphi_0(x(a), x(b)) + \int_a^b f(x(t), \dot{x}(t), t) dt$$

over absolutely continuous trajectories $x : [a, b] \to \mathbf{R}^n$ for the differential inclusion

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

subject to geometric (abstract), inequality, and equality type endpoint constraints:

$$(1.3) \qquad (x(a), x(b)) \in \Omega \subset \mathbf{R}^{2n},$$

$$(1.4) \qquad \varphi_i(x(a), x(b)) \leq 0 \quad \text{for} \quad i = 1, 2, \ldots, q,$$

$$(1.5) \qquad \varphi_i(x(a), x(b)) = 0 \quad \text{for} \quad i = q+1, q+2, \ldots, q+r.$$

Here $T := [a, b]$ is a fixed time interval, F is a set-valued mapping (multifunction), and f, φ_i are real-valued functions. We shall label this problem as (P) and call it *the Bolza problem for differential inclusions.*

The problem formulated covers a broad range of other problems in dynamic optimization, in particular, both standard and nonstandard models in optimal control for open-loop and closed-loop control systems (see, e.g., Clarke [9]). On the other hand, problem (P) can be imbeded in the so-called Generalized Problem of Bolza [41, 9] where the function f is allowed to take values in $\bar{\mathbf{R}} := \mathbf{R} \cup \{+\infty\}$. In this paper we prefer to consider the problem in form (1.1)–(1.5) and prove results depending on the specific character of differential inclusions.

The mainstream in studying optimization problems for differential inclusions consists in obtaining necessary conditions for optimality (global or strong local minima). There are different approaches and various results in this area using one or another tool in nonsmooth analysis; we refer the reader to [6–11, 17, 23–25, 28–30, 37, 38, 40, 48] and the bibliography therein. Most of the results are obtained for the *Mayer problem* which corresponds to (1.1)–(1.5) as $f = 0$.

In [25], Loewen and Rockafellar consider the Bolza problem (P) with the explicit endpoint constraints only in form (1.3) (and with additional state constraints). Assuming that the function $f(x, \cdot, t)$ and the sets $F(x, t)$ of admissible velocities are *convex*, they obtain necessary optimality conditions under usual boundedness and Lipschitzness hypotheses but without imposing any constraint qualification such as calmness (cf. Clarke [6–11]). They prove that if $\bar{x}(t)$ solves problem (1.1)–(1.3), then there exist a number $\lambda \geq 0$ and an absolute continuous function $p : [a, b] \to \mathbf{R}^n$, not both zero, such that

$$(1.6) \quad (\dot{p}(t), p(t)) \in \lambda \partial_C f(\bar{x}(t), \dot{\bar{x}}(t), t) + N_C((\bar{x}(t), \dot{\bar{x}}(t)); \mathrm{gph} F(\cdot, t)) \quad \text{a.e.} \quad t \in [a, b],$$

2

(1.7) $$(-\dot{p}(t), \dot{\bar{x}}(t)) \in \partial_C H_\lambda(\bar{x}(t), p(t), t) \quad \text{a.e.} \ \ t \in [a, b],$$

(1.8) $$\langle p(t), \dot{\bar{x}}(t) \rangle = H_\lambda(\bar{x}(t), p(t), t) \quad \text{a.e.} \ \ t \in [a, b],$$

(1.9) $$(p(a), -p(b)) \in \lambda \partial_C \varphi_0(\bar{x}(a), \bar{x}(b)) + N_C((\bar{x}(a), \bar{x}(b)); \Omega)$$

where gph $F(\cdot, t) := \{(x, v) \in \mathbf{R}^{2n} | v \in F(x, t)\}$,

(1.10) $$H_\lambda(x, p, t) := \max\{\langle p, v \rangle - \lambda f(x, v, t) | v \in F(x, t)\},$$

notation N_C and ∂_C stand, respectively, for Clarke's normal cone to a closed set at a given point and the generalized gradient of a locally Lipschitz function ([9]; see Section 4 below).

Condition (1.6) called the Euler-Lagrange inclusion was first obtained by Clarke [6] for the Mayer problem under the calmness assumption which ensures that $\lambda = 1$ (normality) in the transversality inclusion (1.9). The Hamiltonian inclusion (1.7) was proved by Clarke first under the calmness hypothesis and then without it; see [9, 10]. Observe that (1.8) is the Weierstrass-Pontryagin maximum condition which is implied by each of the conditions (1.6) and (1.7) under the convexity assumptions imposed. Note also that, in general, conditions (1.6) and (1.7) are independent; see examples in [23, 26].

Another version of necessary optimality conditions was obtained by Mordukhovich for the Mayer problem (1.1)–(1.5) with $f = 0$ under the convexity of $F(x, t)$ and usual boundedness and Lipschitzness assumptions but without any calmness hypotheses or something similar; see [28–30]. The conditions obtained are stated in the form:

(1.11) $$(\dot{p}(t), \dot{\bar{x}}(t)) \in \mathrm{co}\{(u, v) | (u, p(t)) \in N((\bar{x}(t), v); \mathrm{gph} F(\cdot, t)),$$

$$v \in M(\bar{x}(t), p(t), t)\} \quad \text{a.e.} \ \ t \in [a, b],$$

(1.12) $$(p(a), -p(b)) \in \lambda \partial \varphi_0(\bar{x}(a), \bar{x}(b)) + N((x(a), x(b)); \Omega)$$

where "co" means the convex hull of a set,

(1.13) $$M(x, p, t) := \{v \in F(x, t) | \langle p, x \rangle = H(x, p, t)\},$$

and $H(x, p, t)$ coincides with the Hamiltonian (1.10) as $f = 0$. Here N and ∂ are not Clarke's normal cone and generalized gradient but their *nonconvex* counterparts whose convex closures coincide with the corresponding constructions of Clarke; see Section 4 for more details. These nonconvex constructions were first used in Mordukhovich [27] for obtaining transversality conditions like (1.12) in nonsmooth problems of optimal control.

Observe that condition (1.11) implies both the maximum condition (1.8) and an analogue of the Euler-Lagrange inclusion (1.6) in the form

(1.14) $$\dot{p}(t) \in \mathrm{co}\{u | (u, p(t)) \in N((\bar{x}(t), v); \mathrm{gph} F(\cdot, t)),$$

$$v \in M(\bar{x}(t), p(t), t)\} \quad \text{a.e.} \quad t \in [a, b].$$

In comparison with (1.6) as $f = 0$, condition (1.14) requires less convexification: only to the components involving derivatives of the adjoint function instead of to all components at once. This makes (1.14) essentially stronger than (1.6) in certain situations. In particular, if the maximum set (1.13) is a singelton along $(\bar{x}(t), p(t))$ for a.e. $t \in [a, b]$ (it happens, for instance, if the sets $F(x, t)$ are strictly convex along $\bar{x}(t)$), then (1.14) is reduced to

$$(1.15) \qquad \dot{p}(t) \in \text{co}\{u | (u, p(t)) \in N((\bar{x}(t), \dot{\bar{x}}(t)); \text{gph} F(\cdot, t))\} \quad \text{a.e.} \quad t \in [a, b]$$

which is strictly better than (1.6) as $f = 0$.

So (1.11) appears to be an advanced version of the Euler-Lagrange inclusion and the maximum condition for the Mayer problem involving convex (i.e., convex-valued) differential inclusions (1.2). What relationships exist between (1.11) and the Hamiltonian inclusion (1.7) as $f = 0$ under the usual convexity, boundedness, and Lipschitzness assumptions?

It follows from Rockafellar's dualization result [45] that (1.7) implies (1.11). On the other hand, it has been recently proved by Ioffe (personal communication; see also [20, Section 3.5]) that (1.11) implies (1.7) under the mentioned assumptions. Therefore, version (1.11) of the Euler-Lagrange condition is *equivalent* to the Hamiltonian condition in Clarke's form (with the same adjoint function) for convex differential inclusions. This prolongs the line of equivalency between the Hamiltonian and Euler-Lagrange conditions which is well known for smooth and fully convex problems (see, e.g., [41, 45]). Now one can conclude that any improvement of the necessary optimality conditions in form (1.11) provides a strengthening of the Hamiltonian conditions in Clarke's form for convex differential inclusions.

In the recent paper [26], Loewen and Rockafellar have established that the Mayer problem for convex differential inclusions can be actually reduced to the situation where the sets $F(x, t)$ are strictly convex along the optimal trajectory. In the latter case, conditions (1.14) and (1.15) are equivalent while (1.11) is equivalent to the simultaneous fulfilment of (1.8) and (1.15). In the general convex case, (1.15) implies the maximum condition in (1.8) as $f = 0$; see Proposition 4.12 stated below. Therefore, (1.15) also implies (1.14) as well as (1.11) where the improvement may be proper; see Example 5.2 in [26].

Using the mentioned strict convexification procedure and a Hamiltonian calculus, Loewen and Rockafellar prove in [26] that an optimal solution to the Mayer problem for convex differential inclusions in (1.1)–(1.3) satisfies conditions (1.12), (1.15), and (1.7), (1.8) as $f = 0$ with the same adjoint function $p(t)$. Moreover, they also consider the case of unbounded differential inclusions truncating it to the bounded case under suitable Lipschitzian assumptions and the convexity of $F(x, t)$. Their consideration of the unbounded case leads to improvements of necessary conditions for the Bolza problem with convex velocities.

It is well known that convexity assumptions are really restrictive in some important situations and it makes sense to release them as much as possible. Moreover, if the admissible velocity sets $F(x, t)$ are convex and the multifunction $F(\cdot, t)$ is Lipschitz continuous, then the differential inclusion (1.2) admits a control representation $F(x, t) = g(x, U, t)$ with a

4

Lipschitzian function g and a control set U independent on x; see, e.g., [3, 29]. This is no longer the case when $F(x, t)$ are not convex. So considering nonconvex differential inclusions, one should definitely study them for their own sake.

The primal goal of the present paper is to develop the theory of necessary optimality conditions in the refined Euler-Lagrange form like (1.15) with the transversality conditions like (1.9) for the Mayer and Bolza problems involving *nonconvex* differential inclusions. For brevity, such a pair of necessary conditions is called the (*refined*) *Euler-Lagrange conditions* in the problems under consideration. We consider the case of general endpoint constraints (1.3)–(1.5) and impose standard Lipschitzness and boundedness assumptions on the multifunction F in (1.2). The results obtained below achieve the following advancements in the state of art.

1. We study a new (to the best of our knowledge) concept of local minima for the considered variational problems involving differential inclusions. Previous results for such problems were concerned with strong (or global) minima. In contrast to a strong minimum, we compare a reference feasible trajectory $\bar{x}(\cdot)$ with other feasible ones close to it not only in the C-norm for arcs but also in the L^p-norm $(1 \leq p < \infty)$ for derivatives. This means that we consider a neighborhood of $\bar{x}(\cdot)$ in the Sobolev space $W^{1,p}$ equipped with a natural topology. Such a local minimum takes an intermediate place between the classical weak and strong minima; we shall call it the *intermediate local minimum*. Note that the results obtained in this paper provide a new information even for convex differential inclusions. In particular, they imply the maximum condition for an intermediate local minimum which might not be strong.

2. We obtain refined necessary conditions for the Bolza problem (P) stated above with the Euler-Lagrange inclusion

$$(1.16) \quad \dot{p}(t) \in \mathrm{co}\{u|(u, p(t)) \in \lambda_0 \partial f(\bar{x}(t), \dot{\bar{x}}(t), t) + N((\bar{x}(t), \dot{\bar{x}}(t)); \mathrm{gph}F(\cdot, t))\}$$

and the transversality inclusion

$$(1.17) \qquad (p(a), -p(b)) \in \partial(\sum_{i=0}^{q+r} \lambda_i \varphi_i)(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega)$$

where numbers $\lambda_0, \ldots, \lambda_{q+r}$ and an absolute continuous function $p(t)$ are not all zero, $\lambda_i \geq 0$ for $i = 0, \ldots, q$, and

$$\lambda_i \varphi_i(\bar{x}(a), \bar{x}(b)) = 0 \text{ for } i = 1, \ldots, q.$$

These transversality conditions are proved in Section 6 for locally Lipschitzian functions φ_i not being surprising for experts working in the area. Actually they are related to (1.12) due to calculus rules avaliable for the nonconvex subdifferentials under consideration (Section 4). One can obtain such conditions even in a stronger form without Lipschitzian assumptions; see Section 7.

In general, we prove necessary conditions (1.16), (1.17) for any trajectory $\bar{x}(t)$ which is feasible for the original problem (P) and provides an intermediate local minimun for the

5

so-called *relaxed* problem obtained from (P) by some convexification procedure. Note that in this case, the Euler-Lagrange inclusion (1.16) is expressed in terms of the original data F, f and may be quite different from its counterpart in terms of the convexifications. We discuss effective sufficient conditions when an optimal solution to (P) solves the relaxed problem as well, so the conditions obtained characterize solutions to the original problem of Bolza without any convexity.

3. In the case of the Mayer functional in (1.1)-(1.5), we prove that conditions (1.15), (1.17) are satisfied for every *strong minima* $\bar{x}(t)$ *without any relaxation*. This implies that if $\bar{x}(t)$ solves the original Mayer problem but may not solve the relaxed one, conditions (1.15) and (1.17) still hold. The same Euler-Lagrange inclusion (1.15) is also proved for any *boundary trajectory* of (1.2) without either convexity or relaxation. The latter result essentially strengthens the recent one of Kaskosz and Lojasiewicz [23] who have proved the Euler-Lagrange inclusion in Clarke's form (1.6) as $f = 0$ for boundary trajectories.

Now let us explain the principal method we employ for obtaining the mentioned results. This is a direct method based on *finite difference (discrete) approximations*. Such an approach to variational problems goes back to Euler who used it in 1744 for proving the classical Euler-Lagrange equation in the calculus of variations. (Actually Leibnitz was the first who employed a similar direct method to find the brachistochrone in the very beginning of the calculus of variations; see, e.g., [1]). The basic idea is as follows: 1) to replace (approximate) the original continuous-time variational problem by a "right" sequence of finite dimensional optimization problems which can be solved (studied) effectively; and then 2) passing to the limit with respect to approximation parameters to obtain desired charateristics of the original variational problem.

Finite difference methods appear to be a powerful tool for numerical solutions of infinite dimensional variational problems. We refer the reader to the book of Polak [36] and the survey paper of Dontchev and Lempio [16] devoted to numerical aspects of consistent discrete approximations in optimal control. Some results in the present paper are also concerned with numerical questions. In Section 3 we develop a discrete approximation algorithm for nonconvex differential inclusions with strong convergence properties and error estimates. But our main interest here is to use finite difference approximations as a direct vehicle for obtaining necessary optimality conditions in infinite dimensional problem (P) via a variational analysis of nonsmooth problems in finite dimensions. Two issues are important in this approach:

1) to construct a correct discrete approximation of problem (P) which ensures a desired convergence of optimal solutions for discrete problems to a given local minimum for (P);

2) to choose "right" generalized derivative (normal) constructions for nosmooth mappings and sets suitable to the method. Such constructions need to be appropriate for obtaining optimality conditions in discrete problems and also possess robustness and calculus properties for passing to the limit in the approximation procedure. Let us observe that problem (P) and its discrete counterparts are definitely objects of nonsmooth analysis and optimizations because of a special nature of dynamic constraints like (1.2) even under smooth data in (1.1) and (1.3)-(1.5).

Note that not any differentiation constructions in nonsmooth analysis fit these requirements. For example, Pshenichnyi employed in [38] some tangentially generated constructions related to the contingent cone. Such constructions possess the required properties only in special situations. This allowed him to prove necessary conditions for global (actually strong) minima in autonomous differential inclusions under some restrictive assumptions close to the graph convexity of $F(\cdot, t)$. He used a discrete approximation ensuring the uniform (C-) convergence of optimal trajectories.

In Mordukhovich [28–30], we used generalized normals and derivatives of another nature and somewhat different algorithms for approximating Lipschitzian differential inclusions. These generalized constructions appear in (1.11), (1.12) and possess required robustness and calculus properties which are considered in Section 4. Note that if one employs the convexification of the normal cone in (1.11), i.e., uses Clarke's normal cone to the graph of a Lipschitzian mapping, then such a construction does not ensure the convergence of adjoint functions in discrete approximations (see Remark 4.15).

Though the approximation algorithm in [29] is used for proving the C-convergence of optimal trajectories in discrete approximations, its slight modification provides the *strong L^2-*convergence of the velocities. It was first observed by Smirnov [40] who obtained the refined condition (1.15) for optimal solutions to a Mayer problem involving convex autonomous differential inclusions under some additional assumptions.

In the present paper, we develop the method of discrete approximations to obtain the results mentioned above in the general setting under consideration. The reminder of the paper is organized as follows.

Section 2 is devoted to the concept of intermediate local minimum for the original and relaxed problems of Bolza. We consider sufficient conditions ensuring the "proper relaxation" of (P) when a given minimum for the original problem solves the relaxed problem as well.

Section 3 deals with discrete approximations of problem (P). We provide a construction of discrete approximations and natural assumptions which ensure the strong convergence of optimal solutions with respect to the value function, trajectories, and velocities.

In Section 4 we describe the tools of the generalized differentiation for nonsmooth and set-valued mappings used in the paper to obtain necessary optimality conditions for discrete and differential inclusions. The reader can find there a review of the basic differentiation properties important in the method of discrete approximations.

Section 5 is concerned with necessary optimality conditions for nonsmooth finite dimensional problems. We obtain discrete analogues of the refined Euler-Lagrange conditions (1.16), (1.17) without the convexity operation in (1.16) and any Lipschitzian assumptions on F, f. These results appear to be direct consequences of the Lagrange multiplier rule in nondifferentiable programming with many geometric constraints.

Section 6 is devoted to the limiting procedure in discrete approximations allowing us to prove conditions (1.16) and (1.17) for an intermediate relaxed local minimun in (P). Under relaxation stability, the results obtained characterize optimal solutions to the original problem without imposing convexity.

In the final Section 7, we deal with strong minima for the nonconvex Mayer problem in (1.1)–(1.5) without any relaxation. We prove the refined Euler-Lagrange and transversality conditions (1.15) and (1.17) on the base of results in Section 6 and an approximation procedure involving Ekeland's variational principle. The same approach works to prove (1.15) for any boundary trajectory.

Notation in this paper is standard. The adjoint (transposed) matrix to A is denoted by $A^\star$; the set B is always the unit closed ball of the space in question. Some special symbols are introduced and explained in Section 4.

2. Intermediate Local Minimum and Relaxation

Recall that we consider problem (P) stated above in the class of absolutely continuous functions $x : [a, b] \to \mathbf{R}^n$ (arcs) satisfying constraints (1.2)–(1.5). Any solution to (1.2) is called an (*original*) *trajectory* for the differential inclusion, and any trajectory satisfying constraints (1.3)–(1.5) is called a *feasible solution* to problem (P). Let us introduce a notion of local minimum for (P) studied in the paper.

2.1. DEFINITION. The arc $\bar{x}(\cdot)$ is called an *intermediate local minimun* (i.l.m.) of rank $p \in [1, \infty)$ for (P) if $\bar{x}(\cdot)$ is a feasible solution to (P) and there exist numbers ϵ, $\varepsilon > 0$ such that $J[\bar{x}] \leq J[x]$ for any other feasible solution $x(\cdot)$ to (P) satisfying

$$(2.1) \qquad |x(t) - \bar{x}(t)| < \epsilon \text{ for all } t \in [a, b],$$

$$(2.2) \qquad \int_a^b |\dot{x}(t) - \dot{\bar{x}}(t)|^p dt < \varepsilon \text{ a.e. } t \in [a, b].$$

If (2.2) is fulfilled, then instead of (2.1) one can obviously set $|x(a) - \bar{x}(a)| < \eta$ with $\eta > 0$ and $\epsilon \geq \eta + \varepsilon(b - a)$. Relationships (2.1), (2.2) mean that we consider a neighborhood of $\bar{x}(\cdot)$ in the Sobolev space $W^{1,p}$ of absolute continuous functions $x : [a, b] \to \mathbf{R}^n$ equipped with a natural norm.

If there is only requirement (2.1) in Definition 2.1, then one gets a *strong* local minimum (with respect to the C-norm). This actually corresponds to the L^1-weak topology for derivatives instead of the strong (L^p-norm) topology in (2.2). Obviously any *optimal solution* to problem (P) (global minimum) provides a strong local minimun (and, therefore, an i.l.m.) for (P). As we know, all available necessary optimality conditions for differential inclusions work for strong local minima but it is not clear a priori whether they hold for i.l.m.

If instead of (2.3) one sets the more restrictive requirement

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

then we have a *weak* local minimum in the framework of Definition 2.1. This corresponds to considering a neighborhood of $\bar{x}(\cdot)$ in the space $W^{1,\infty}$ with the L^∞-norm for derivatives (or the C^1-norm for continuous differentiable functions in classical variational problems).

Therefore, the notion of i.l.m. introduced takes (for any rank $p \in [1, \infty)$) an intermediate place between the familiar concepts of strong and weak minima. Note that some aspects of this setting are related to the Lavrentiev phenomenon in the calculus of variation; see, e.g., [5, 24].

Now we construct an extension of the original problem (P) in the line well known in the calculus of variations and optimal control (cf., e.g., [4, 7, 14, 21, 49, 51]). Let

$$(2.3) \qquad f_F(x, v, t) := f(x, v, t) + \delta(v, F(x, t))$$

where $\delta(v, \Lambda) = 0$ if $v \in \Lambda$ and $\delta(v, \Lambda) = \infty$ if $v \notin \Lambda$ (the indicator function). Denote by $\hat{f}_F(x, v, t)$ the convexification (the second conjugate function) for f_F in the v variable, i.e., the largest convex function majorized by $f_F(x, \cdot, t)$ for each x and t. Along with the original problem (P), we consider its *relaxation* (R) as follows:

$$(2.4) \qquad \text{minimize} \;\; \hat{J}[x] := \varphi_0(x(a), x(b)) + \int_a^b \hat{f}_F(x(t), \dot{x}(t), t) dt$$

over absolutely continuous functions on $[a, b]$ under endpoint constraints (1.3)–(1.5). Note that if $\hat{J}[x] < \infty$, then $x(\cdot)$ satisfies the convexified differential inclusion

$$(2.5) \qquad \dot{x}(t) \in \mathrm{co} F(x(t), t) \;\; \text{a.e.} \;\; t \in [a, b].$$

Any trajectory for (2.5) is called a *relaxed trajectory* for (1.2). It is well known that under natural assumptions involving Lipschitzness of F in x, the following *approximation property* holds: Every relaxed trajectory $x(\cdot)$ can be uniformly in $[a, b]$ approximated by original trajectories $x_k(\cdot)$ starting from the same initial state (but may be not satisfied endpoint constraints) such that

$$(2.6) \qquad \varliminf \int_a^b f(x_k(t), \dot{x}_k(t), t) dt \leq \int_a^b \hat{f}_F(x(t), \dot{x}(t), t) dt \;\; \text{as} \;\; k \to \infty.$$

2.2. DEFINITION. The arc $\bar{x}(\cdot)$ is called an *intermediate relaxed local minimum* (i.r.l.m.) of rank $p \in [1, \infty)$ for the original problem (P) if $\bar{x}(\cdot)$ is a feasible solution to (P) providing an intermediate local minimum of rank p for the relaxed problem (R) with $J[\bar{x}] = \hat{J}[\bar{x}]$.

It is essential in Definition 2.2 that $\bar{x}(\cdot)$ is an original trajectory for (1.2). Obviously, there is no difference between i.r.l.m. and i.l.m. if problem (P) is *convex* in the following sence: for each $t \in [a, b]$ and x around $\bar{x}(t)$, the function f is convex in v on the convex set $F(x, t)$. One can see that the refined Euler-Lagrange inclusion obtained for i.r.l.m. may be different from its counterpart for arbitrary i.l.m. in the relaxed problem. This happens because the normal cone to the graph of F and to the graph of $\mathrm{co} F$ is not the same.

On the other hand, we cannot guarantee, in principal, that necessary conditions for i.r.l.m. will work for arbitrary i.l.m. (or strong minima) for the original problem. Nevertheless, the latter holds in rather general settings without any convexity assumptions. Actually

9

this is related to the property of "hidden convexity" inherent in continuous-time systems like (1.2). In the present paper, we use only one result in this direction going back to the classical Bogoljubov theorem [4].

2.3. PROPOSITION. *Let $\bar{x}(\cdot) \in W^{1,\infty}[a,b]$ be a strong minimum for problem (1.1), (1.3)–(1.5) where the integrant $f(x,v,t)$ is continuous in (x,v) around $(\bar{x}(t), \dot{\bar{x}}(t))$ iniformly in $[a,b]$ and measurable in t. Then $\bar{x}(\cdot)$ is a strong minimum for the relaxed problem (1.3)– (1.5), (2.4) as $f_F = f$ with $J[\bar{x}] = \hat{J}[\bar{x}]$.*

Proof. According to the version of Bogoljubov's theorem in [22, Section 9.2.4], for any $x(\cdot) \in W^{1,\infty}[a,b]$ one can find a sequence of $x_k(\cdot) \in W^{1,\infty}[a,b]$ such that $x_k(a) = x(a)$, $x_k(b) = x(b)$, $x_k(\cdot)$ converges to $x(\cdot)$ uniformly in $[a,b]$, and (2.6) is fulfilled as $f_F = f$. If $\bar{x}(\cdot)$ is not a strong minimum for the relaxed problem (2.4), (1.3)–(1.5) or/and $\hat{J}[\bar{x}] < J[\bar{x}]$, then there exists a function $x(\cdot) \in W^{1,\infty}[a,b]$ with $\hat{J}[x] < J[\bar{x}]$ such that $x(\cdot)$ belongs to the C-neighborhood of $\bar{x}(\cdot)$ and satisfies constraints (1.3)–(1.5). This contradicts the strong minimality of $\bar{x}(\cdot)$ in the original problem, thanks to the Bogoljubov approximation for $x(\cdot)$. □

There are several generalizations and analogues of Bogoljubov's theorem having many important applications to optimal control systems and differential inclusions; see, e.g., [7, 14, 21, 29, 49] and references therein. They lead to the property of *relaxation stability* (or proper relaxation) when an optimal solution to the original problem solves the relaxed problem as well with the same optimal value. For problems (P) involving differential inclusions with endpoint constraints at either $t = a$ or $t = b$, such a relaxation stability follows directly from the approximation property for relaxed trajectories stated above.

For problems with general endpoint constraints, the relaxation stability is ensured by the *calmness* property in Clarke [7, 9]. The latter property is fulfilled for "almost all" endpoint constraints (at least of inequality type) and shows that the relaxation stability may fail only for ill-posed problems where small perturbations of boundary conditions produce proportionally unbounded variations of the minimum. According to Clarke [9], 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 [50].

For special classes of problems (P) with arbitrary endpoint constraints, the relaxation stability holds without any calmness or normality assumptions. In particular, let differential inclusion (1.2) be represented in the *linear* form

$$\dot{x}(t) \in F_1(t)x(t) + F_2(t) \quad \text{a.e.} \quad t \in [a,b]$$

where the multifunctions F_1 and F_2 are integrable in $[a,b]$, F_1 is convex-valued while F_2 is not. If, moreover, the function f in (1.1) is convex in v, then any of such problems (P) possesses the property of relaxation stability. This can be proved by using Aumann's theorem about set-valued integrals; cf. arguments in [29, Theorem 19.7]. The same situation holds for

problems (P) involving nonlinear one-dimensional differential inclusions; see Remark 19.2 in [29].

3. Discrete Approximations

In this section we construct a sequence of discrete approximations for the original problem of Bolza such that optimal solutions to discrete approximations converge in $W^{1,p}$ to a given i.r.l.m. for problem (P).

First let us consider a fixed original trajectory $\bar{x}(\cdot)$ for (1.2) and prove that it can be approximated by trajectories for corresponding discrete inclusions. To do this, we assume that the multifunction $F(x,t)$ is bounded and locally Lipschitzian around $\bar{x}(\cdot)$ and it is also Hausdorff continuous in t a.e. on $[a,b]$. More precisely, we impose the following hypotheses:

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

$$(3.1) \qquad F(x,t) \subset m_F B \quad \forall (x,t) \in U \times [a,b],$$

$$(3.2) \qquad F(x_1,t) \subset F(x_2,t) + l_F |x_1 - x_2| B \quad \forall x_1, x_2 \in U, \ t \in [a,b].$$

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

Following Dontchev and Farkhi [15], let us consider the so-called *averaged modulus of continuity* for the multifunction $F(x,t)$ in $t \in [a,b]$ when $x \in U$. This modulus $\tau(F;h)$ depending on the parameter $h > 0$ is defined as follows:

$$(3.3) \qquad \tau(F;h) := \int_a^b \sigma(F;t,h)dt$$

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

$$\omega(F;x,t,h) := \sup\{\mathrm{haus}(F(x,t'),F(x,t''))|t',t'' \in [t-h/2,t+h/2] \cap [a,b],$$

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

It is proved in [15] that if $F(x,\cdot)$ is Hausdorff continuous for a.e. $t \in [a,b]$ uniformly in $x \in U$, then $\tau(F;h) \to 0$ as $h \to 0$.

Note that in the case of single-valued functions $f(t)$ not depending on x, the construction $\tau(f;h)$ in (3.3) has been developed in Sendov and Popov [39] under the name of "averaged modulus of smoothness". It has been proved in [39] that $\tau(f;h) \to \infty$ as $h \to 0$ if and only if f is Riemann integrable on $[a,b]$. The latter is equivalent to f being continuous for a.e. $t \in [a,b]$. In this paper, we shall use the name "average modulus of continuity" for both single-valued and multi-valued cases.

11

Now let us construct a finite difference (discrete) approximation for the given differential inclusion using the replacement of the derivative in (1.2) by the *Euler finite difference*

$$\dot{x}(t) \approx [x(t+h) - x(t)]/h.$$

For any natural number $k = 1, 2, \ldots$, we consider a uniform grid $T_k := \{t_j | j = 0, 1, \ldots, k\}$ with $t_0 = a$, $t_k = b$ and stepsize

$$h_k := (b - a)/k = t_{j+1} - t_j \quad (j = 0, \ldots, k-1).$$

The associate discrete inclusion is as follows:

$$(3.4) \qquad x_{j+1}^k \in x_j^k + h_k F(x_j^k, t_j), \quad j = 0, \ldots, k-1, \quad x_0^k = \bar{x}(a).$$

3.1. THEOREM. *Let $\bar{x}(\cdot)$ be a trajectory for (1.2) under hypotheses* (H1) *and* (H2). *Then there is a sequence* $\{z_j^k | j = 0, \ldots, k\}$, $k = 1, 2, \ldots$, *of solutions to discrete inclusions* (3.4) *such that the functions*

$$(3.5) \qquad v^k(t) := (z_{j+1}^k - z_j^k)/h_k, \quad t \in [t_j, t_{j+1}), \quad j = 0, \ldots, k-1,$$

converge to $\dot{\bar{x}}(\cdot)$ as $k \to \infty$ in the norm topology of $L^1[a, b]$.

Proof. Let $\{w^k(\cdot)\}$, $k = 1, 2, \ldots$, be an arbitrary sequence of functions in $[a, b]$ such that $w^k(t)$ are constant in $[t_j, t_{j+1})$ for every $j = 0, \ldots, k-1$ and $w^k(t)$ converge to $\dot{\bar{x}}(t)$ as $k \to \infty$ in the norm of $L^1[a, b]$. Such a sequence always exists because of the density of step-functions in the space $L^p[a, b]$ for any $p \in [1, \infty]$. Taking $p = \infty$ and using (3.1), one gets

$$(3.6) \qquad |w^k(t)| \le m_F + 1 \quad \forall t \in [a, b] \quad \text{as} \quad k \to \infty.$$

In the arguments and estimates below, we shall use the number

$$(3.7) \qquad \xi_k := \int_a^b |\dot{\bar{x}}(t) - w^k(t)| dt \to 0 \quad \text{as} \quad k \to \infty.$$

Let us define the discrete functions $\{y_j^k | j = 0, \ldots, k\}$ as follows:

$$(3.8) \qquad y_{j+1}^k = y_j^k + h_k w_j^k, \quad j = 0, \ldots, k-1, \quad y_0^k = \bar{x}(a)$$

where $w_j^k := w^k(t_j)$, $j = 0, \ldots, k-1$. Note that the functions

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

are piecewise linear extensions of (3.8) on the interval $[a, b]$ with

$$(3.9) \qquad |y^k(t) - \bar{x}(t)| \le (b - a)\xi_k \quad \forall t \in [a, b].$$

12

Therefore, $y^k(t) \in U$ for all $t \in [a, b]$ if k is big enough.

Denote by $\mathrm{dist}(w, F)$ the Euclidean distance between the point w and the closed set F. It is well known that the Lipschitz condition (3.2) is equivalent to

$$\mathrm{dist}(w, F(x_1, t)) \le \mathrm{dist}(w, F(x_2, t)) + l_F |x_1 - x_2| \quad \forall w \in \mathbf{R}^n, \ x_1, x_2 \in U, \ t \in [a, b].$$

For any $w, x \in \mathbf{R}^n$ and $t_1, t_2 \in [a, b]$, one obviously has

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

Now using (3.3), we get

$$\varsigma_k := \sum_{j=0}^{k-1} h_k \mathrm{dist}(w_j^k, F(y_j^k, t_j)) = \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \mathrm{dist}(w_j^k, F(y_j^k, t_j)) dt \le$$

$$\sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} \mathrm{dist}(w_j^k, F(y_j^k, t)) dt + \tau(F; h_k).$$

It follows from (3.2), (3.6), and (3.9) that

$$\mathrm{dist}(w_j^k, F(y_j^k, t)) \le \mathrm{dist}(w^k(t), F(y^k(t), t)) + l_F(m_F + 1)(t - t_j) \quad \forall t \in [t_j, t_{j+1}),$$

$$\mathrm{dist}(w^k(t), F(y^k(t), t)) \le \mathrm{dist}(w^k(t), F(\bar{x}(t), t)) + l_F |y^k(t) - \bar{x}(t)| \le$$

$$|w^k(t) - \dot{\bar{x}}(t)| + l_F(b - a)\xi_k \quad \text{a.e.} \ t \in [a, b].$$

Therefore, we have the estimate

$$(3.10) \qquad \varsigma_k \le \gamma_k := (1 + l_F(b - a))\xi_k + l_F(b - a)(m_F + 1)h_k/2 + \tau(F; h_F).$$

Note that functions (3.8) are not trajectories for (3.4) because one does not get $w_j^k \in F(y_j^k, t_j)$. Now we shall use w_j^k to define trajectories for (3.4) which are close to y_j^k and have the convergence property stated in this theorem.

Let us construct the desired trajectories $\{z_j^k | j = 0, \dots, k\}$ using the following *proximal algorithm*:

$$(3.11) \qquad z_0^k = \bar{x}(a), \ v_j^k \in F(z_j^k, t_j) \ \text{with} \ |v_j^k - w_j^k| = \mathrm{dist}(w_j^k, F(z_j^k, t_j)),$$

$$z_{j+1}^k = z_j^k + h_k v_j^k, \ j = 0, \dots, k - 1.$$

Note that in (3.11) we take projections of velocities as in [29, 40] instead of projections of states as in [38, 15]. This will allow us to prove a strong convergence of discrete approximations with respect to velocities.

First we prove that algorithm (3.11) keeps $\{z_j^k | j = 0, \dots, k\}$ inside the neighborhood U from (H1) if k is big enough. Indeed, we consider any number k such that $\bar{x}(t) + \eta_k B \subset U$ for all $t \in [a, b]$ where

$$\eta_k := \gamma_k \exp[l_F(b - a)] + (b - a)\xi_k,$$

ξ_k and γ_k are defined in (3.7) and (3.10) respectively. One can see that $\eta_k \to 0$ as $k \to \infty$ because $\tau(F; h_k) \to 0$ under assumption (H2).

By induction, let us show that if $z_m^k \in V$ for all $m = 0, \ldots, j$, then this also holds for $m = j + 1$. Using (3.2), (3.10), and (3.11), one gets

$$|z_{j+1}^k - y_{j+1}^k| \le |z_j^k - y_j^k| + h_k \mathrm{dist}(w_j^k, F(z_j^k, t_j)) \le |z_j^k - y_j^k| + h_k(\mathrm{dist}(w_j^k, F(y_j^k, t_j)) +$$

$$l_F|z_j^k - y_j^k|) \le \ldots \le h_k \sum_{m=0}^{j} (1 + l_F h_k)^{j-m} \mathrm{dist}(w_m^k, F(y_m^k, t_m)) \le$$

$$\exp[l_F(b - a)] \sum_{m=0}^{j} h_k \mathrm{dist}(w_m^k, F(y_m^k, t_m)) \le \gamma_k \exp[l_F(b - a)].$$

Due to (3.9), the latter implies

$$(3.12) \qquad |z_{j+1}^k - \bar{x}(t_{j+1})| \le \gamma_k \exp[l(b - a)] + (b - a)\xi_k := \eta_k$$

which proves that $z_{j+1}^k \in U$ for all $j \in \{0, \ldots, k\}$. Taking this into account, one can extract from the previous arguments the following estimate:

$$(3.13) \qquad \sum_{j=0}^{k} |z_j^k - y_j^k| \le (b - a) \exp[l_F(b - a)] \sum_{j=0}^{k-1} \mathrm{dist}(w_j^k, F(y_j^k, t_j)).$$

Now let us estimate the quantity $\alpha_k := \int_a^b |v^k(t) - w^k(t)| dt$ where the functions $v^k(t)$ are defined in (3.5). Employing (3.10), (3.11) and (3.13), we get

$$\alpha_k = \sum_{j=0}^{k-1} h_k |v_j^k - w_j^k| = \sum_{j=0}^{k-1} h_k \mathrm{dist}(w_j^k, F(z_j^k, t_j)) \le \sum_{j=0}^{k-1} h_k \mathrm{dist}(w_j^k, F(y_j^k, t_j)) +$$

$$l_F \sum_{j=0}^{k-1} h_k |z_j^k - y_j^k| \le \gamma_k (1 + l_F(b - a) \exp[l_F(b - a)]).$$

Thus we obtain the final estimate

$$(3.14) \qquad \alpha_k := \int_a^b |v^k(t) - \dot{\bar{x}}(t)| dt \le \beta_k := \xi_k + \gamma_k (1 + (b - a) \exp[l_F(b - a)]).$$

This ensures the L^1-convergence $v^k(\cdot) \to \dot{\bar{x}}(\cdot)$ due to (3.7) and $\tau(F; h_k) \to 0$ as $k \to \infty$ under (H2). $\square$

3.2. *Remarks.* The result obtained provides a strong approximation with respect to velocities of any absolutely continuous trajectories for the differential inclusion (1.2) by discrete trajectories for its Euler difference counterparts (3.4). Note that the error estimate for velocities (3.14) immediately implies the following estimate

$$|z^k(t) - \bar{x}(t)| \le (b - a)\beta_k \quad \forall t \in [a, b]$$

14

for the corresponding motions $z^k(t) := \bar{x}(a) + \int_a^t v^k(s)ds$ which are piecewise linear extensions of discrete trajectories (3.10). One can see that the *numerical efficiency* of the estimates obtained depends on the evaluation of $\tau(F; h)$ and the approximation accuracy in (3.7).

It has been proved in [15] that $\tau(F; h) = O(h)$ if $F(\cdot, x)$ is of bounded variation on $[a, b]$ uniformly in $x \in U$. Using the technique for averaged moduli of continuity (smoothness) developed in [39], one can obtain effective estimates for ξ_k in (3.7). Indeed, if $\dot{\bar{x}}(\cdot)$ is Riemann integrable on $[a, b]$, then we always get $\xi_k \leq 2\tau(\dot{\bar{x}}; h_k)$ taking $v^k(t) = \dot{\bar{x}}(t_j)$ for $t \in [t_j, t_j + h_k)$ as $j = 0, \ldots, k - 1$.

Now we consider the given original trajectory $\bar{x}(\cdot)$ which appears to be an i.r.l.m. of some rank $p \in [1, \infty)$ for problem (P). One can easily see that under boundedness assumption (3.1), the notion of i.r.l.m. for (P) *does not depend on rank p*. This means that if $\bar{x}(\cdot)$ is an i.r.l.m. of some rank $p \in [1, \infty)$, then it will be an i.r.l.m. of any other rank from $[1, \infty)$. In what follows we shall always take $p = 2$.

Let us construct a sequence of optimization problems (P_k) for discrete inclusions (3.4) such that optimal solutions to (P_k) strongly (in $W^{1,2}[a, b]$) converge to $\bar{x}(\cdot)$ as $k \to \infty$. Take numbers ϵ and ε from the $W^{1,2}$-neighborhood (2.1), (2.2) of the given i.r.l.m. and assume (H1), (H2) along $\bar{x}(\cdot)$. One can always suppose that $\bar{x}(t) + \epsilon/2 \in U$ for all $t \in [a, b]$. Using Theorem 3.1, we approximate $\bar{x}(\cdot)$ by discrete trajectories $\{z_j^k | j = 0, \ldots, k\}$ and compute the numbers η_k in (3.12) and

$$(3.15) \qquad \rho_i^k := \varphi_i(\bar{x}(a), z_k^k) - \varphi_i(\bar{x}(a), \bar{x}(b)) \text{ for } i = 1, \ldots, q,$$

$$(3.16) \qquad \varrho_i^k := |\varphi_i(\bar{x}(a), z_k^k)| \text{ for } i = q + 1, \ldots, q + r.$$

One can see that $|\rho_i^k| \leq l_i \eta_k$ for $i = 1, \ldots, q$ and $\varrho_i^k \leq l_i \eta_k$ for $i = q + 1, \ldots, q + r$ if the corresponding functions $\varphi_i(\bar{x}(a), \cdot)$ are Lipschitz continuous around $\bar{x}(b)$ with constants l_i.

For each $k = 1, 2, \ldots$, we define the discrete optimization problem (P_k) as follows:

$$(3.17) \qquad \text{minimize } J_k[x^k] := \varphi_0(x_0^k, x_k^k) + |x_0^k - \bar{x}(a)|^2 +$$

$$h_k \sum_{j=0}^{k-1} f(x_j^k, (x_{j+1}^k - x_j^k)/h_k, t_j) + \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} |(x_{j+1}^k - x_j^k)/h_k - \dot{\bar{x}}(t)|^2 dt$$

over discrete trajectories $x^k = (x_0^k, x_1^k, \ldots, x_k^k)$ for the difference inclusion (3.4) subject to constraints

$$(3.18) \qquad \varphi_i(x_0^k, x_k^k) \leq \rho_i^k \text{ for } i = 1, \ldots, q,$$

$$(3.19) \qquad -\varrho_i^k \leq \varphi_i(x_0^k, x_k^k) \leq \varrho_i^k \text{ for } i = q + 1, \ldots, q + r,$$

$$(3.20) \qquad (x_0^k, x_k^k) \in \Omega_k := \Omega + \eta_k B,$$

$$(3.21) \qquad\qquad |x_j^k - \bar{x}(t_j)| \leq \epsilon/2 \ \text{ for } \ j = 0, \ldots, k,$$

$$(3.22) \qquad\qquad \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} |(x_{j+1}^k - x_j^k)/h_k - \dot{\bar{x}}(t)|^2 dt \leq \varepsilon/2.$$

Let $x^k(\cdot)$ be the piecewise linear extension of the discrete trajectory $\{x_j^k | j = 0, \ldots, k\}$ on $[a, b]$, and let $\dot{x}^k(\cdot)$ denote the piecewise constant extension of the "velocity" $(x_{j+1}^k - x_j^k)/h_k$. One has

$$\dot{x}^k(t) = (x_{j+1}^k - x_j^k)/h_k = \dot{x}^k(t_j) \ \text{ for any } \ t \in [t_j, t_{j+1}).$$

We shall consider the (strong) $W^{1,2}$-convergence of $x^k(\cdot)$ to some absolutely continuous function $x(\cdot)$ in $[a, b]$. This means that $x^k(a) = x_0^k \to x(a)$ and $\dot{x}^k(\cdot) \to \dot{x}(\cdot)$ in $L^2[a, b]$ as $k \to \infty$. The latter obviously implies that $x^k(\cdot)$ converges to $x(\cdot)$ uniformly in $[a, b]$.

In addition to (H1) and (H2), now we impose the following hypotheses on f, φ_i, and Ω:

(H3) $f(x, v, \cdot)$ is continuous for a.e. $t \in [a, b]$ and bounded uniformly in $(x, v) \in U \times (m_F B)$.

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

$$(3.23) \quad A_\nu(t) := \{(x, v) \in U \times (m_F + \nu)B | \ v \in F(x, t') \ \text{for some} \ t' \in (t - \nu, t]\}$$

uniformly in $t \in [a, b]$.

(H5) The functions φ_i are continuous on $U \times U$ for all $i = 0, \ldots, q + r$; the set Ω is closed around $(\bar{x}(a), \bar{x}(b))$.

3.3. THEOREM. *Let $\bar{x}(\cdot)$ be an i.r.l.m. for problem (P) and let hypotheses (H1)–(H5) be fulfilled. Then any sequence $\{\bar{x}^k(\cdot)\}$, $k = 1, 2, \ldots$, of optimal solutions to (P_k) converges to $\bar{x}(\cdot)$ in the space $W^{1,2}[a, b]$ as $k \to \infty$.*

Proof. First let us prove that for each k big enough, the discrete trajectory $\{z_j^k | j = 0, \ldots, k\}$ constructed in Theorem 3.1 is a feasible solution to (P_k). We need to check that this trajectory satisfies all the constraints (3.18)–(3.22). For the case of (3.18)–(3.20), it follows directly from (3.12) and (3.15), (3.16). Taking k such that $\eta_k \leq \epsilon/2$, we get (3.21) from (3.12). By virtue of (3.1) and (3.14), constraint (3.22) for $x^k = z^k$ is reduced to

$$\int_a^b |v^k(t) - \dot{\bar{x}}(t)|^2 dt \leq 2 m_F \alpha_k \leq 2 m_F \beta_k \leq \varepsilon/2.$$

The latter is fulfilled for big numbers k due to the expression for β_k in (3.14). Therefore, x^k is a feasible solution to (P_k) for all k big enough. According to the classical Weierstrass theorem, we can conclude that there is an optimal solution $\bar{x}^k$ to (P_k) for such k under the assumptions made.

Let us prove that for any sequence of optimal solutions x^k to (P_k) one has

(3.24) $$\overline{\lim} \, J_k[\bar{x}^k] \le J[\bar{x}] \quad \text{as} \ k \to \infty.$$

If it is not true, then there exists a sequence $\mathcal{N}$ of natural numbers $k \to \infty$ such that

(3.25) $$\varphi_0(\bar{x}(a), \bar{x}(b)) + \int_a^b f(\bar{x}(t), \dot{\bar{x}}(t), t)dt < J_k[\bar{x}^k] \quad \forall k \in \mathcal{N}.$$

Using again Theorem 3.1, we approximate $\bar{x}(\cdot)$ by discrete trajectories z^k which are proved to be feasible for (P_k). Due to the continuity of φ_0 one has

$$\varphi_0(z_0^k, z_k^k) \to \varphi_0(\bar{x}(a), \bar{x}(b)) \quad \text{as} \ k \to \infty.$$

In expression (3.17) for $J_k[z^k]$, the second term vanishes; the forth term tends to zero as $k \to \infty$ because of (3.5) and (3.14). Let us prove that

(3.26) $$\sigma_k := h_k \sum_{j=0}^{k-1} f(z_j^k, (z_{j+1}^k - z_j^k)h_k, t_j) \to \int_a^b f(\bar{x}(t), \dot{\bar{x}}(t), t) \quad \text{as} \ k \to \infty$$

under assumptions (H1), (H3), and (H4). Note that (H3) implies $\tau(f; h_k) \to 0$ as $k \to \infty$ for modulus (3.3). It what follows we use the sign "$\sim$" for expressions which are equivalent as $k \to \infty$. Due to (3.5), (3.12), and (3.14) one gets

$$\sigma_k = \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} f(z_j^k, v^k(t), t_j)dt \sim \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} f(z_j^k, v^k(t), t)dt + \tau(f; h_k) \sim$$

$$\sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} f(\bar{x}(t), v^k(t), t)dt = \int_a^b f(\bar{x}(t), v^k(t), t)dt \sim \int_a^b f(\bar{x}(t), \dot{\bar{x}}(t), t)dt.$$

The last statement holds by virtue of the classical Lebesque limiting theorem because $\{v^k(\cdot)\}$ contains a subsequence converging for a.e. $t \in [a, b]$.

Due to (3.25) and (3.26) we have $J_k[z^k] < J_k[\bar{x}^k]$ for some $k \to \infty$. This contradicts the optimality of $\bar{x}^k$ in (P_k) and proves (3.24).

In the arguments above we haven't actually used the property of $\bar{x}(\cdot)$ to be an i.r.l.m. for (P). Now let us prove that in the latter case, inequality (3.24) implies that

(3.27) $$\lim_{k \to \infty} [c_k := |\bar{x}^k(a) - \bar{x}(a)|^2 + \int_a^b |\dot{\bar{x}}^k(t) - \dot{\bar{x}}(t)|^2 dt] = 0,$$

i.e., $\bar{x}^k(\cdot)$ converges to $\bar{x}(\cdot)$ in the norm of $W^{1,2}[a, b]$. Suppose that it is not true and consider a limiting point $c > 0$ of the sequence $\{c_k\}$ in (3.27). Let, for simplicity, $c = \lim c_k$ for all $k \to \infty$.

Because of (3.21) and (3.22) we claim the existence of an absolutely continuous function $\tilde{x}(t)$ in $[a, b]$ such that $\bar{x}^k(\cdot) \to \tilde{x}(\cdot)$ uniformly in $[a, b]$ and $\dot{\bar{x}}^k(\cdot) \to \dot{\tilde{x}}(\cdot)$ weakly in $L^2[a, b]$ as $k \to \infty$ (we take all k without loss of generality). According to the classical Masur theorem,

17

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 $L^2[a,b]$. Hence it contains a subsequence converging to $\dot{\tilde{x}}(\cdot)$ for a.e. $t \in [a,b]$.

Using these facts and taking into account that

$$h_k \sum_{j=0}^{k-1} f(\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k, t_j) \sim \int_a^b f(\bar{x}^k(t), \dot{\bar{x}}^k(t), t)dt \quad \text{as} \quad k \to \infty$$

and also the definition of $\hat{f}_F$ for (2.3), we get

$$(3.28) \qquad \int_a^b \hat{f}_F(\tilde{x}(t), \dot{\tilde{x}}(t), t)dt \leq \varliminf h_k \sum_{j=0}^{k-1} f(\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)h_k, t_j) \quad \text{as} \quad k \to \infty.$$

where $\tilde{x}(\cdot)$ satisfies the confexified differential inclusion (2.5).

Observe that the integral functional

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

is lower semicontinuous in the weak topology of $L^2[a,b]$ due to the convexity of the integrant in v. Since

$$\sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} |(\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k - \dot{\tilde{x}}(t)|^2 dt = \int_a^b |\dot{\bar{x}}^k(t) - \dot{\tilde{x}}(t)|^2 dt,$$

the latter implies that

$$(3.29) \quad \int_a^b |\dot{\tilde{x}}(t) - \dot{\tilde{x}}(t)|^2 dt \leq \varliminf \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} |(\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k - \dot{\tilde{x}}(t)|^2 dt \quad \text{as} \quad k \to \infty.$$

Now passing to the limit in (3.18)–(3.22) as $k \to \infty$ and using (3.29) as well as (H5), we get that $\tilde{x}(\cdot)$ satisfies constraints (1.3)–(1.5) and

$$|\tilde{x}(t) - \bar{x}(t)| \leq \epsilon/2 \quad \text{for} \quad t \in [a,b], \quad \int_a^b |\dot{\tilde{x}}(t) - \dot{\bar{x}}(t)|^2 dt \leq \varepsilon/2.$$

The latter means that $\tilde{x}(\cdot)$ belongs to the given neighborhood of $\bar{x}(\cdot)$ in $W^{1,2}[a,b]$. Moreover, (3.28) implies

$$(3.30) \qquad \varphi_0(\tilde{x}(a), \tilde{x}(b)) + \int_a^b \hat{f}(\tilde{x}(t), \dot{\tilde{x}}(t), t)dt + c \leq \varliminf J_k[x^k] \quad \text{as} \quad k \to \infty.$$

Due to (3.24), (3.30), and $c > 0$ we get $\hat{J}[\tilde{x}] < J[\bar{x}]$. But this is impossible because $\bar{x}(\cdot)$ is an i.r.l.m. for (P). Therefore, one has $c = 0$ which establishes (3.27) and ends the proof of the theorem. $\square$

The convergence theorem proved allows us to make a bridge between variational problems for differential inclusions and dynamic optimization problems in finite dimensions. The latter can be reduced to finite dimensional problems of mathematical programming with

many geometric constraints. The mathematical programming problems obtained in this way appear to be objects of nonsmooth optimization even in the case of smooth initial data in the original problem (P).

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

4. Tools of Variational Analysis

This section is concerned with tools of generalized differentiation appropriate for the main objectives of the research. We define the basic concepts of generalized normals and derivatives and review some of their important properties which are broadly used in the paper. Most of these results with detailed proofs and discussions can be found in the book of Mordukhovich [29, Chapter 1]. We also refer the reader to Clarke [11], Ioffe [18–20], Mordukhovich [26–35], Rockafellar [42, 46], and the forthcoming book of Rockafellar and Wets [47] for related and additional material.

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

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

$$(4.1) \qquad \Pi(x, \Omega) := \{\omega \in \mathrm{cl}\,\Omega \text{ such that } |x - \omega| = \mathrm{dist}(x, \Omega)\}$$

be the (multivalued) *Euclidean projector* of x on $\mathrm{cl}\,\Omega$. In the following definition, "cone" denotes the conic hull of a set and "lim sup" means the well-known *Kuratowski-Painlevé upper limit* for multifunctions, i.e., the collection of all limiting points (see, e.g., [3, p. 41]).

4.1. DEFINITION [27]. Given $\bar{x} \in \mathrm{cl}\,\Omega$, the closed cone

$$(4.2) \qquad N(\bar{x}; \Omega) := \limsup_{x \to \bar{x}} [\mathrm{cone}(x - \Pi(x, \Omega))]$$

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

The normal cone defined (or equivalent constructions) is widely used in the literature, sometimes under different names: the "approximate normal cone" [18], the "limiting proximal normal cone" [42], the "prenormal cone" [11], etc. If Ω is convex, then (4.2) is reduced to the normal cone of convex analysis. In general, the normal cone (4.2) is frequently *nonconvex* and its convex closure coincides with the *Clarke normal cone*:

$$(4.3) \qquad N_C(\bar{x}; \Omega) = \mathrm{clco}\, N(\bar{x}; \Omega).$$

Furthermore, (4.2) always admits the representation

$$(4.4) \qquad N(\bar{x}; \Omega) = \limsup_{x \to \bar{x}} \hat{N}(x; \Omega)$$

where the cone

$$\hat{N}(x;\Omega) := \{x^\star \in \mathbf{R}^n | \overline{\lim}_{x'(\in\Omega)\to x}\langle x^\star, x' - x\rangle/|x' - x| \le 0\}$$

appears to be *dual* to the well-known Bouligand *contingent cone*; see, e.g., [3, Chapter 4].

Note that the normal cone (4.2) is not dual to any tangent cone because it is not convex. Despite its nonconvexity, this construction possesses many nice properties important for applications. First, we observe that the normal cone (4.2) 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 the nonconvex normal cone and related generalized differential objects (see below) enjoy rich calculi which are even better than ones for corresponding convex-valued constructions. The progress in this direction has been achieved by using a *variational approach* instead of convex analysis.

Now let us consider an extended-real-valued function $f : \mathbf{R}^n \to \bar{\mathbf{R}}$ and define its subdifferential associated with the normal cone (4.2) to the *epigraph* $\text{epi} f := \{(x,\mu) \in \mathbf{R}^{n+1} | \mu \ge f(x)\}$.

4.2. DEFINITION [27]. Let $\bar{x} \in \text{dom} f := \{x \in \mathbf{R}^n | -\infty < f(x) < \infty\}$. The set

$$(4.5) \qquad \partial f(\bar{x}) := \{x^\star \in \mathbf{R}^n | (x^\star - 1) \in N((\bar{x}, f(\bar{x})); \text{epi} f)\}$$

is called the *subdifferential* of f at $\bar{x}$. If $\bar{x} \notin \text{dom} f$, we put $\partial f(\bar{x}) = \emptyset$.

There are various analytic (limiting) representations of the subdifferential (4.5) in terms of other subgradient mappings. We refer to [11, 18–20, 29–35, 42, 46, 47] for more information about these questions. Therein one can find subdifferential constructions equivalent to (4.5) under different names (e.g., the approximate subdifferential, the presubdifferential, the set of limiting proximal subgradients). Observe that for continuous functions, the subdifferential (4.5) appears to be the upper limit (robust regularization) of the subdifferential mapping used in the theory of *viscosity solution* [12]. Note that the set (4.5) may be not convex for simple nonconvex functions such as $f(x) = -|x|$, $x \in \mathbf{R}$, where $\partial f(0) = \{-1,1\}$.

For the case of Lipschitz continuous functions, we have the following well-known result (see, e.g., [11, Propositions 1.1 and 1.2] and [29, Theorem 2.1]).

4.3. PROPOSITION. *Let the function f be locally Lipschitzian around $\bar{x}$ with a Lipschitz constant l_f. Then*

$$\partial f(\bar{x}) \ne \emptyset, \quad |x^\star| \le l_f \ \forall x^\star \in \partial f(\bar{x}), \quad \text{and}$$

$$(4.6) \qquad \partial_C f(\bar{x} = \text{co}\partial f(\bar{x})$$

for the generalized gradient of Clarke.

It is easy to see that

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

for the indicator $\delta(\cdot, \Omega)$ of the set Ω. We can get another representation of the normal cone in terms of the subdifferential of the Lipschitz continuous distance function. The following result is proved in [29, Proposition 2.7].

4.4. PROPOSITION. *For an arbitrary set Ω, one has*

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

Among the most important advantages of constructions (4.2), (4.5) and ones related to them, one has a rich calculus under general assumptions. We refer to [11, 18, 19, 29, 31, 34, 42, 47] for various results in this direction. For applications in this paper, we need the two following calculus rules.

4.5. PROPOSITION. *Let f_1 and f_2 be lower semicontinuous functions one of which is locally Lipschitzian around $\bar{x}$. Then*

$$(4.7) \qquad\qquad \partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x})$$

when equality holds if one of these functions is strictly differentiable at $\bar{x}$ (in particular, belongs to C^1).

Different proofs of inclusion (4.7) (but using somewhat close approximation ideas) can be found, e.g., in [11, Proposition 1.5], [18, Theorem 4], [29, Theorem 4.1], and [31, Theorem 3.1]; see also references therein. The case of equality follows from (4.7) by employing it to the sum $(f_1 + f_2) + (-f_2)$ where f_2 is strictly differentiable. The next "chain rule" in the form of equality is proved in [29, Theorem 4.7].

4.6. PROPOSITION. *Let $(\varphi \circ g)(x) := \varphi(g(x))$ be a composition of the function $g : \mathbf{R}^n \rightarrow \mathbf{R}^m$ Lipschitz continuous around $\bar{x}$ and the function $\varphi : \mathbf{R}^m \rightarrow \mathbf{R}$ strictly differentiable at $\bar{y} := g(\bar{x})$. Then one has*

$$\partial(\varphi \circ g)(\bar{x}) = \partial\langle\nabla\varphi(\bar{y}), g\rangle(\bar{x})$$

where $\langle\nabla\varphi(\bar{y}), g\rangle(x) := \langle\nabla\varphi(\bar{y}, g(x)\rangle$.

Now we consider generalized differentiation constructions for multifunctions and nonsmooth mappings connected with the normal cone (4.2) to their *graphs*. Pick an arbitrary multifunction $F : \mathbf{R}^n \Rightarrow \mathbf{R}^m$ with the nonempty graph

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

4.7. DEFINITION [28]. Let the point $(\bar{x}, \bar{y})$ belong to the closure of gph$F$. The multi-function $D^\star F(\bar{x}, \bar{y})$ from $\mathbf{R}^m$ into $\mathbf{R}^n$ defined by

$$(4.8) \qquad D^\star F(\bar{x}, \bar{y})(y^\star) := \{x^\star \in \mathbf{R}^n | (x^\star, -y^\star) \in N((\bar{x}, \bar{y}); \mathrm{gph} F)\}$$

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

One can see that $D^\star F(\bar{x}, \bar{y})(\cdot)$ is a positive homogeneous multifunction with closed values. These values may be not convex by virtue of the nonconvexity in (4.2). Due to the latter property, the coderivative (4.8) is *not dual* to any tangentially generated *derivatives* of multifunctions (see, e.g., [3, Chapter 5]). On the other hand, the normal cone representation (4.4) implies that the coderivative (4.8) appears to be the *robust regularization* of the dual construction to Aubin's *contingent derivative* [2, 3].

Let us observe that the subdifferential (4.5) of an extended-real-valued function f can be treated as the coderivative of the special epigraphical multifunction

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

Indeed, (4.5) means that $\partial f(\bar{x}) = D^\star E_f(\bar{x}, f(\bar{x}))(1)$. The next proposition (see [29, Proposition 2.1]) allows us to replace epi$f$ by gph$f$ in (4.5) for the case of lower semicontinuous functions.

4.8. PROPOSITION. *Let the function* $f : \mathbf{R}^n \to \bar{\mathbf{R}}$ *be l.s.c. around* $\bar{x} \in$ dom$f$. *Then one has* $\partial f(\bar{x}) = D^\star f(\bar{x})(1)$.

Now let us consider the multifunction F of a special form which graph is

$$(4.9) \qquad \mathrm{gph} F := \{(x, y) \in \mathbf{R}^n \times \mathbf{R}^m | x \in \Omega, \ g(x) - y \in \Delta\}$$

where $g : \mathbf{R}^n \to \mathbf{R}^m$. For each $x \in \mathbf{R}^n$ and $y^\star \in \mathbf{R}^m$, define the (generalized) *scalarization function*:

$$(4.10) \qquad s(x, y^\star) := \langle y^\star, g(x) \rangle + \delta(x, \Omega).$$

The following result proved in [29, Theorem 3.3] provides the calculation of the coderivative for the multifunction (4.9).

4.9. PROPOSITION. (i) *Let the set* (4.9) *be closed around the point* $(\bar{x}, \bar{y}) \in$ gph$F$. *Then one has the implication*

$$[D^\star F(\bar{x}, \bar{y})(y^\star) \neq \emptyset] \implies y^\star \in N(g(\bar{x}) - \bar{y}; \Delta).$$

(ii) *If g is Lipschitz continuous around $\bar{x}$, then*

(4.11) $$D^\star F(\bar{x}, \bar{y})(y^\star) = \begin{cases} \partial_x s(\bar{x}, y^\star) & \text{for } y^\star \in N(g(\bar{x}) - \bar{y}; \Delta) \\ \emptyset & \text{otherwise} \end{cases}$$

(4.12)

Setting $\Omega = \mathbf{R}^n$, $\Delta = \{0\}$, and $\langle y^\star, g \rangle(x) = \langle y^\star, g(x) \rangle$, we get from (4.11) the following scalarization formula proved in [18, Proposition 8].

4.10. COROLLARY. *Let $g : \mathbf{R}^n \to \mathbf{R}^m$ be Lipschitz continuous around $\bar{x}$. Then*

(4.13) $$D^\star g(\bar{x}(y^\star) = \partial \langle y^\star, g \rangle(\bar{x}) \quad \forall y^\star \in \mathbf{R}^m.$$

Now using (4.6), (4.12), and a chain rule in [9, Theorem 2.6.6], one has the following relationship between the coderivative (4.8) and Clarke's generalized Jacobian $J_C g(\bar{x})$ of locally Lipschitz functions [9, p. 70].

4.11. COROLLARY. *Let $g : \mathbf{R}^n \to \mathbf{R}^m$ be Lipschitz continuous around $\bar{x}$. Then*

(4.14) $$\text{co} D^\star g(\bar{x})(y^\star = \text{co} \partial \langle y^\star, g \rangle(\bar{x}) = (J_C g(\bar{x}))^\star y^\star \quad \forall y^\star \in \mathbf{R}^m.$$

Coming back to the coderivative (4.8) for general multifunctions, let us formulate a useful result [29, Theorem 3.1] in the convex-valued case. In particular, this result ensures 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., [3, p. 39].

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

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

It turns out that the coderivative introduced enjoys a rich calculus under natural (e.g., Lipschitzian) assumptions in general settings. We refer the reader to our recent paper [34] for various results in this direction. The results obtained cover not only the comprehensive calculus for the first order subdifferentials (4.5) and relative constructions but also provide useful calculus rules for the second order objects generated by (4.8). The latter constructions appear to be important for the sensitivity analysis of generalized equations and variational inequalities; see [33, 35].

To conclude this section, let us present applications of the coderivative (4.8) to dual characterizations of Lipschitzian properies of multifunctions. These results play a *crucial role*

for the approximation methods in the paper ensuring the convergence of adjoint functions; see the proofs of Theorem 6.1 and 7.1.

Recall that the multifunction $F : \mathbf{R}^n \Rightarrow \mathbf{R}^m$ of closed graph is said to be *pseudo-Lipschitzian* around the point $(\bar{x}, \bar{y}) \in \mathrm{gph}F$ if there exist a neighborhood U of $\bar{x}$, a neighborhood V of $\bar{y}$, and a constant $l \geq 0$ such that

$$(4.15) \qquad F(x') \cap V \subset F(x) + l|x' - x|B \quad \forall x, \; x' \in U.$$

This definition goes back to Aubin [2] who imposed an additional condition $F(x) \cap V \neq \emptyset$ for all $x \in U$; see also [3, Definition 1.4.5]. The pseudo-Lipschitzian property in the sense of (4.14) was studied by Rockafellar [43] who proved its equivalence to the locally Lipschitzness of the distance function $\mathrm{dist}(v, F(x))$ in both variables. The next characterizations were obtained by Mordukhovich (see [32, Theorem 5.7] and [33, Theorem 3.2] for different proofs).

4.13. PROPOSITION. *The following conditions are equivalent*:

(i) *F is pseudo-Lipschitzian around $(\bar{x}, \bar{y})$*;

(ii) *there exist a neighborhood U of $\bar{x}$, a neighborhood V of $\bar{y}$, and a constant $l \geq 0$ such that*

$$(4.16) \qquad \sup\{|x^*| : x^* \in D^*F(x, y)(y^*)\} \leq l|y^*|$$

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

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

If F is bounded around $\bar{x}$, then its classical locally Lipschitzian behavior (like in (3.2)) is equivalent to F being pseudo-Lipschitzian around $(\bar{x}, \bar{y})$ for every $\bar{y} \in F(\bar{x})$; see [43, Theorems 2.1 and 2.2]. Therefore, we get criteria for the classical Lipschitzian property of multifunctions.

4.14. COROLLARY. *Let the multifunction F be bounded around the point $\bar{x}$ with $F(\bar{x}) \neq \emptyset$. Then the following properties are equivalent*:

(i) *F is locally Lipschitzian around $\bar{x}$*;

(ii) *there are a neighborhood U of $\bar{x}$ and a number $l \geq 0$ such that estimate (4.15) holds if $x \in U$ and $y^* \in \mathbf{R}^m$*;

(iii) *$D^*F(\bar{x}, \bar{y})(0) = \{0\} \; \forall \bar{y} \in F(\bar{x})$*.

4.15. *Remark.* If one replaces the normal cone (4.2) in the coderivative construction (4.8) by the Clarke normal cone (4.3), then such a counterpart D_C^*F of the coderivative does not ensure estimate (4.15) and the "null-condition" (iii) for Lipschitzian multifunctions in many important situations. This is related to the fact that Clarke's normal cone to any *Lipschitzian manifold* (which is a set locally representable as the *graph* of a Lipschitz continuous vector function) is actually a *linear subspace*; see Rockafellar [44]. It turns out

that Lipschitzian manifolds include not only graphs of Lipschitz continuous functions but also maximal monotone operators, subdifferential mappings for convex and saddle functions, etc. For such objects, estimate (4.15) in terms of $D_C^\star F$ and the corresponding "null-condition" are fulfilled in fact only for "strictly smooth" multifunctions. We refer to [44] and [35] for more information about these and related properties.

5. Necessary Conditions for Discrete Approximations

In this section we obtain necessary optimality conditions in discrete approximation problems (P_k) for each $k = 1, 2, \ldots$. These conditions will be derived from a *generalized Lagrange multiplier rule* for finite dimensional problems in mathematical programming with many geometric constraints, nonsmooth inequality constraints, and smooth equality constraints.

Let $\phi_i : \mathbf{R}^d \to \bar{\mathbf{R}}$ for $i = 0, \ldots, s$ and $g_j : \mathbf{R}^d \to \mathbf{R}^n$ for $j = 0, \ldots, m$. Consider the following problem (MP):

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

$$(5.2) \qquad \phi_i(z) \le 0 \text{ for } i = 1, \ldots, s,$$

$$(5.3) \qquad g_j(z) = 0 \text{ for } j = 0, \ldots, m,$$

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

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

$$(5.5) \qquad z_j^\star \in N(\bar{z}; \Delta_j) \text{ for } j = 0, \ldots, l,$$

$$(5.6) \qquad \mu_i \ge 0 \text{ for } i = 0, \ldots, s,$$

$$(5.7) \qquad \mu_i \phi_i(\bar{z}) = 0 \text{ for } i = 1, \ldots, s,$$

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

The proof of this result based on the *metric approximation method* can be found in Mordukhovich [28, Theorem 1] and [29, Corollary 7.5.1]. Note that this method allows us to obtain necessary conditions for (MP) in more general forms in the presence of nonsmooth

equality and inequality constraints without Lipschitzian assumptions; see [29, Section 7] and Section 7 below.

Now we employ Proposition 5.1 and calculus rules for the generalized differential constructions in Section 4 to prove necessary optimality conditions for finite difference problems (P_k) in the following *Euler-Lagrange form*. Considering problem (P_k) in (3.4), (3.18)–(3.22) for any fixed $k = 1, 2, \ldots$, we denote

$$F_j(\cdot) := F(\cdot, t_j) \text{ and } f_j(\cdot, \cdot) := f(\cdot, \cdot, t_j) \text{ as } j = 0, \ldots, k-1.$$

5.2. THEOREM. *Let $\bar{x}^k = (\bar{x}_0^k, \ldots, \bar{x}_k^k)$ be an optimal solution to problem (P_k). Assume that the sets Ω and $\mathrm{gph}F_j$ are closed and the functions φ_i and f_j are Lipschitz continuous around the points $(\bar{x}_0^k, \bar{x}_k^k)$ and $(\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k)$ respectively for all $i = 0, \ldots, q+r$ and $j = 0, \ldots, k-1$. Then there exist real numbers $\lambda_0^k, \ldots, \lambda_{q+r}^k$ and a vector $p^k = (p_0^k, \ldots, p_k^k) \in \mathbf{R}^{(k+1)n}$, not all zero, such that*

(5.9)
$$\lambda_i^k \geq 0 \text{ for } i = 0, \ldots, q,$$

(5.10)
$$\lambda_i^k(\varphi_i(\bar{x}_0^k, \bar{x}_k^k) - \rho_i^k) = 0 \text{ for } i = 1, \ldots, q,$$

(5.11)
$$(p_0^k + 2\lambda_0^k(\bar{x}(a) - \bar{x}_0^k), -p_k^k) \in \partial(\sum_{i=0}^{q+r} \lambda_i^k \varphi_i)(\bar{x}_0^k, \bar{x}_k^k) + N((\bar{x}_0^k, \bar{x}_k^k); \Omega_k),$$

(5.12)
$$((p_{j+1}^k - p_j^k)/h_k, p_{j+1}^k - \lambda_0^k \theta_j^k/h_k) \in \lambda_0^k \partial f_j(\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k) +$$

$$N((\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k); \mathrm{gph}F_j) \text{ for } j = 0, \ldots, k-1$$

where

(5.13)
$$\theta_j^k := 2\int_{t_j}^{t_{j+1}} (\dot{\bar{x}}(t) - (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k)dt.$$

Proof. Let us introduce a new variable $z = (x_0, \ldots, x_k, y_0, \ldots, y_{k-1}) \in \mathbf{R}^{(2k+1)n}$ and consider the following problem of mathematical programming:

(5.14)
$$\text{minimize } \phi_0(z) := \varphi_0(x_0, x_k) + |x_0 - \bar{x}(a)|^2 + h_k \sum_{j=0}^{k-1} f_j(x_j, y_j) +$$

$$\sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} |y_j - \dot{\bar{x}}(t)|^2 dt \text{ subject to}$$

(5.15)
$$\phi_i(z) := \varphi_i(x_0, x_k) - \rho_i^k \leq 0 \text{ for } i = 1, \ldots, q,$$

(5.16)
$$\phi_i(z) := \varphi_i(x_0, x_k) - \varrho_i^k \leq 0 \text{ for } i = q+1, \ldots, q+r,$$

26

$$(5.17) \qquad \phi_{r+i}(z) := -\varphi_i(x_0, x_k) - \varrho_i^k \le 0 \text{ for } i = q+1, \ldots, q+r,$$

$$(5.18) \qquad \phi_{q+2r+j+1}(z) := |x_j - \bar{x}(t_j)| - \epsilon/2 \le 0 \text{ for } j = 0, \ldots, k,$$

$$(5.19) \qquad \phi_{q+2r+k+2}(z) := \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} |y_j - \dot{\bar{x}}(t)|^2 dt - \varepsilon/2 \le 0,$$

$$(5.20) \qquad g_j(z) := x_{j+1} - x_j - h_k y_j = 0 \text{ for } j = 0, \ldots, k-1,$$

$$(5.21) \quad z \in \Delta_j := \{(x_0, \ldots, y_{k-1}) \in \mathbf{R}^{(2k+1)n} | y_j \in F_j(x_j)\} \text{ for } j = 0, \ldots, k-1,$$

$$(5.22) \qquad z \in \Delta_k := \{(x_0, \ldots, y_{k-1}) \in \mathbf{R}^{(2k+1)n} | (x_0, x_k) \in \Omega_k\}.$$

It is easy to see that problem (5.14)–(5.22) defined is equivalent to the discrete approximation problem (P_k) in (3.4), (3.17)–(3.22). On the other hand, (5.14)–(5.22) is a problem (MP) in (5.1)–(5.4) with $d = (2k+1)n$, $s = q + 2r + k + 2$, $m = k - 1$, $l = k$, and the specified functions ϕ_i, g_j and sets Δ_j. Now we employ Proposition 5.1 for the optimal solution $\bar{z} = \bar{z}^k := (\bar{x}_0^k, \ldots, \bar{x}_k^k, (\bar{x}_1^k - \bar{x}_0^k)/h_k, \ldots, (\bar{x}_k^k - \bar{x}_{k-1}^k)/h_k)$ to (5.14)–(5.22) where $\bar{x}^k = (\bar{x}_0^k, \ldots, \bar{x}_k^k)$ is a given optimal solution to (P_k).

According to this result, one gets real numbers $(\mu_0, \ldots, \mu_{q+2r+k+2})$ as well as vectors $\psi_j \in \mathbf{R}^n$ $(j = 0, \ldots, k-1)$ and $z_j^* = (x_{0j}^*, \ldots, x_{kj}^*, y_{0j}^*, \ldots, y_{j\,k-1}^*) \in \mathbf{R}^{2k+1)n}$ $(j = 0, \ldots, k)$, not all zero, such that conditions (5.5)–(5.8) are fulfilled for the initial data in (5.14)–(5.22). Note that these μ_i, ψ_j, and z_j^* depend on k but we shall omit the index "k" for simplicity, considering k big enough.

First let us observe that thanks to Theorem 3.3, $\phi_i(\bar{z}^k) < 0$ as $i = q + 2r + 1, \ldots, q + 2r + k + 2$ for all big k. This implies $\mu_i = 0$ for $i = q + 2r + 1, \ldots, q + 2r + k + 2$ by virtue of the corresponding complemetary slackness conditions in (5.7).

Without loss of generality we suppose that $\varrho_i^k > 0$ for $i = q + 1, \ldots, q + r$. Then complementary slackness for (5.16) and (5.17) implies that

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

Now denoting

$$(5.23) \quad \lambda_i^k := \mu_i \text{ for } i = 0, \ldots, q \ \& \ \lambda_i^k := \mu_i - \mu_{r+i} \text{ for } i = q + 1, \ldots, q + r,$$

we get conditions (5.9), (5.10) and ensure that $(\lambda_0^k, \ldots, \lambda_{q+r}^k)$ are not equal to zero simultaneously with $(\psi_0, \ldots, \psi_{k-1})$ and $(z_0^*, \ldots, z_k^*)$ for k big enough.

It follows from the structure of the sets Δ_j in (5.21) and (5.22) that conditions (5.5) are equivalent to

$$(5.24) \qquad (x_{jj}^*, y_{jj}^*) \in N((\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k); \text{gph} F_j) \ \&$$

$$x_{ij}^\star = y_{ij}^\star \ \text{if} \ i \neq j \ \forall j = 0, \ldots, k-1;$$

(5.25) $\qquad (x_{0k}^\star, x_{kk}^\star) \in N((\bar{x}_0^k, \bar{x}_k^k); \Omega_k) \ \& \ x_{ik}^\star = y_{ik}^\star = 0 \ \text{otherwise.}$

Taking this into account and using Proposition 4.5, we get from (5.7), (5.14)–(5.17), and (5.19) the following relationships:

(5.26) $\qquad -x_{00}^\star - x_{0k}^\star = u_0 + 2\lambda_0^k(\bar{x}_0^k - \bar{x}(a)) + \lambda_0^k h_k \vartheta_0 - \psi_0,$

(5.27) $\qquad -x_{jj}^\star = \lambda_0^k h_k \vartheta_j + \psi_{j-1} - \psi_j \ \text{for} \ j = 1, \ldots, k-1,$

(5.28) $\qquad -x_{kk}^\star = u_k + \psi_{k-1},$

(5.29) $\qquad -y_{jj}^\star = \lambda_0^k h_k w_j + \lambda_0^k \theta_j^k - h_k \psi_j \ \text{for} \ j = 0, \ldots, k-1$

where θ_j^k is defined in (5.13),

(5.30) $\qquad (u_0, u_k) \in \partial(\sum_{i=0}^{q+r} \lambda_i^k \varphi_i)(\bar{x}_0^k, \bar{x}_k^k), \ \text{and}$

(5.31) $\qquad (\vartheta_i, w_j) \in \partial f_j(\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k) \ \text{for} \ j = 0, \ldots, k-1.$

Now denoting

$$p_0^k := x_{0k}^\star + 2\lambda_0^k(\bar{x}_0^k - \bar{x}(a)) \ \& \ p_j := \psi_{j-1} \ \text{for} \ j = 1, \ldots, k,$$

one can conclude that relationships (5.24)–(5.31) imply conditions (5.11) and (5.12) where $(\lambda_0^k \ldots, \lambda_{q+r}^k)$ and $(p_0^k, \ldots, p_k^k)$ are not equal to zero simultaneously. This ends the proof of the theorem. $\quad \square$

5.3. COROLLARY. *In addition to the assumptions of Theorem 5.2, let us suppose that for each $j = 0, \ldots, k-1$, the multifunction F_j is pseudo-Lipschitzian around $(\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k)$. Then conditions (5.9)–(5.12) are fulfilled with $(\lambda_0^k, \ldots, \lambda_{q+r}^k, p_k^k) \neq 0$, i.e., one can set*

(5.32) $\qquad |p_k^k| + \sum_{i=0}^{q} \lambda_i^k + \sum_{i=q+1}^{q+r} |\lambda_i^k| = 1 \ \forall k = 1, 2, \ldots.$

Proof. If $\lambda_0^k = 0$, then (5.12) is represented as

(5.33) $\quad (p_{j+1}^k - p_j^k)/h_k \in D^\star F_j((\bar{x}_j^k, (x_{j+1}^k - \bar{x}_j^k)/h_k)(-p_{j+1}^k) \ \text{for} \ j = 0, \ldots, k-1$

in terms of the coderivative (4.8). By virtue of (5.33), $p_k^k = 0$ implies that $p_j^k = 0$ for all $j = 0, \ldots, k-1$, according to Proposition 4.13(iii). This proves the corollary. $\quad \square$

6. Necessary Conditions for the Bolza Problem

Now we come back to the original Bolza problem (P) and prove necessary optimality conditions for an i.r.l.m. in the refined Euler-Lagrange form. To accomplish it, it remains to pass to the limit in the necessary conditions for discrete approximation problems (P_k) as $k \to \infty$ (Section 5), taking into account the $W^{1,2}$-convergence of discrete optimal solutions (Section 3) and some properties of the generalized differential constructions in Section 4.

Here we keep assumptions (H1)–(H3) but instead of the continuity hypothethes in (H4) and (H5), we assume the corresponding Lipschitz continuity. Namely:

(H4′) There exist numbers $\nu > 0$ and $l_f \leq 0$ such that the function $f(\cdot, \cdot, t)$ is Lipschitz continuous on the set $A_\nu(t)$ in (3.23) with the constant l_f, i.e.,

$$|f(x_1, v_1, t) - f(x_2, v_2, t)| \leq l_f |(x_1, v_1) - (x_2, v_2)| \quad \forall (x_1, v_1), \ (x_2, v_2) \in A_\nu(t), \ t \in [a, b].$$

(H5′) The functions φ_i are Lipschitz continuous on $U \times U$ for all $i = 0, \ldots, q + r$; the set Ω is closed around $(\bar{x}(a), \bar{x}(b))$.

In what follows, we denote by $\partial f = \partial f(\cdot, \cdot, t)$ the subdifferential (4.5) of the function $f(x, v, t)$ with respect to (x, v) under fixed t. Similarly, $N((\cdot, \cdot); \mathrm{gph} F(\cdot, t))$ means the normal cone (4.2) to the set $\mathrm{gph} F(\cdot, t)$ at a given point $(\cdot, \cdot)$ when t is fixed. Note that the normal cone to the graph of F is related to the generalized derivative (coderivative) of F according to (4.8).

One of the fundamental properties of the generalized differential constructions under consideration is their robustness (upper semicontinuity) with respect to variables of differentiation; see Section 4. This is of a principal importance for the method of discrete approximations. In the limiting procedure below, we also need such a robustness of $\partial f(\cdot, \cdot, t)$ and $N((\cdot, \cdot); \mathrm{gph} F(\cdot, t))$ with respect to the parameter t. More precisely, we impose the following technical assumptions:

(H6) For a.e. $t \in [a, b]$ one has

$$\limsup_{\substack{(x', v') \to (\bar{x}(t), \dot{\bar{x}}(t)) \\ t' \to t, \ t' < t}} \partial f(x', v', t') = \partial f(\bar{x}(t), \dot{\bar{x}}(t), t).$$

(H7) For a.e. $t \in [a, b]$ one has

$$\limsup_{\substack{(x', v') \to (\bar{x}(t), \dot{\bar{x}}(t)) \\ t' \to t, \ t' < t}} N((x', v'); \mathrm{gph} F(\cdot, t')) = N((\bar{x}(t), \dot{\bar{x}}(t)); \mathrm{gph} F(\cdot, t)).$$

Properties (H6) and (H7) are obviously fulfilled if f and F do not depend on t and also if $f = f_1(x, v) + f_2(t)$, $F = F_1(x) + F_2(t)$ with F_2 satisfying (H2). Actually, (H6) and (H7) mean that the continuity in t holds under the generalized differentiation of f and F with

respect to the other variables. In particular, this takes place when f and F are represented as compositions of mappings separated in t and (x, v). Note also that for smooth in (x, v) functions f, (H6) means the classical continuity of $\partial f / \partial x$ and $\partial f / \partial v$ at $(\bar{x}(t), \dot{\bar{x}}(t), t)$.

Now we prove the refined Euler-Lagrange conditions for the original Bolza problem (P).

6.1. THEOREM. *Let $\bar{x}(\cdot)$ be an i.r.l.m. for problem (P) under assumptions* (H1)–(H3), (H4′), (H5′), (H6), *and* (H7). *Then there exist real numbers* $\lambda_0, \ldots, \lambda_{q+r}$ *and an absolutely continuous function* $p : [a, b] \to \mathbf{R}^n$, *not all zero, such that*

$$(6.1) \qquad \lambda_i \geq 0 \ \text{ for } \ i = 0, \ldots, q,$$

$$(6.2) \qquad \lambda_i \varphi_i(\bar{x}(a), \bar{x}(b)) = 0 \ \text{ for } \ i = 1, \ldots, q,$$

$$(6.3) \qquad (p(a), -p(b)) \in \partial(\sum_{i=0}^{q+r} \lambda_i \varphi_i)(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega),$$

$$(6.4) \qquad \dot{p}(t) \in \mathrm{co}\{u | (u, p(t)) \in \lambda_0 \partial f(\bar{x}(t), \dot{\bar{x}}(t), t) + N((\bar{x}(t), \dot{\bar{x}}(t)); \mathrm{gph} F(\cdot, t))\}$$

for a.e. $t \in [a, b]$.

Proof. Let us construct a sequence of discrete approximations (P_k) of problem (P) which approximates $\bar{x}(\cdot)$ in the sense of Theorem 3.3. Now employing Theorem 5.2 for optimal solutions $x^k = (x_0^k, \ldots, x_k^k)$ to (P_k) as $k \to \infty$, we find real numbers $\lambda_0^k, \ldots, \lambda_{q+r}^k$ and vectors $p^k = (p_0^k, \ldots, p_k^k)$ satisfying conditions (5.9)–(5.13), and (5.32). One can always suppose that $\lambda_i^k \to \lambda_i$ as $k \to \infty$ for all $i = 0, \ldots, q + r$.

We shall use notation $\bar{x}^k(t)$ and $p^k(t)$ for piecewise linear extensions of the corresponding discrete functions on $[a, b]$ with their piecewise constant derivatives $\dot{\bar{x}}^k(t)$ and $\dot{p}^k(t)$. We also consider a sequence of the functions

$$\theta^k(t) := \theta_j^k / h_k \ \text{ for } \ t \in [t_j, t_{j+1}), \ \ j = 0, \ldots, k - 1,$$

generated by (5.13). Theorem 3.3 implies that

$$(6.5) \qquad \int_a^b |\theta^k(t)| dt = \sum_{j=0}^{k-1} |\theta_j^k| \leq 2 \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} |\dot{\bar{x}}(t) - (\bar{x}_{j+1}^k - \bar{x}_j^k) / h_k| dt =$$

$$2 \int_a^b |\dot{\bar{x}}(t) - \dot{\bar{x}}^k(t)| dt := \nu_k \to 0 \ \text{ as } \ k \to \infty.$$

Without loss of generality we can suppose that

$$(6.6) \qquad \dot{\bar{x}}^k(t) \to \dot{\bar{x}}(t) \ \text{ and } \ \theta^k(t) \to 0 \ \text{ a.e. } \ t \in [a, b] \ \text{ as } \ k \to \infty.$$

Let us estimate the adjoint functions $p^k(\cdot)$ for big k. According to (5.12) and Definition 4.7 of the coderivative $D^\star F_j$, there exist vectors $(\vartheta_j^k, w_j^k) \in \partial f_i(\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k)$ such that

$$(6.7) \quad (p_{j+1}^k - p_j^k)/h_k - \lambda_0^k \vartheta_j^k \in D^\star F_j(\bar{x}_j^k, (\bar{x}_{j+1}^k - \bar{x}_j^k)/h_k)(\lambda_0^k w_j^k + \lambda_0^k \theta_j^k/h_k - p_{j+1}^k)$$

for all $j = 0, \ldots, k - 1$. Now using (6.5), (3.2), and Corollary 4.14(ii), one has

$$(6.8) \quad |(p_{j+1}^k - p_j^k)/h_k - \lambda_0^k \vartheta_j^k| \leq l_F |\lambda_0^k w_j^k + \lambda_0^k \theta_j^k/h_k - p_{j+1}^k| \quad \text{for} \quad j = 0, \ldots, k - 1.$$

It follows from (H4′) and Proposition 4.3 that

$$(6.9) \qquad\qquad |\vartheta_j^k| \leq l_f \quad \text{and} \quad |w_j^k| \leq l_f \quad \text{for} \quad j = 0, \ldots, k - 1.$$

Using (5.32), (6.6), (6.8), and (6.9), we get

$$(6.10) \qquad\qquad |p_j^k| \leq (1 + h_k l_F)|p_{j+1}^k| + h_k l_f(1 + l_F) + l_F \theta_j^k \leq \ldots \leq$$

$$\exp[l_F(b - a)] + l_f(b - a)(1 + l_F) + l_F \nu_k \quad \forall j = 0, \ldots, k - 1 \quad \text{as} \quad k \to \infty.$$

This means that the adjoint functions $p^k(t)$ are uniformly bounded in $[a, b]$. Employing (6.8) and (6.9) to estimate the derivatives $\dot{p}^k(t)$, one has

$$|\dot{p}(t)| = |(p_{j+1}^k - p_j^k)/h_k| \leq l_f + l_F(l_f + |\theta^k(t)| + |p_j^k|) \quad \text{for} \quad t_j \leq t < t_{j+1}.$$

By virtue of (6.6) and (6.10) this implies that the sequence $\{\dot{p}^k(\cdot)\}$ is weakly compact in $L^1[a, b]$. Therefore, we can find an absolutely continuous function $p(\cdot)$ such that $p^k(\cdot) \to p(\cdot)$ uniformly in $[a, b]$ and $\dot{p}^k(\cdot) \to \dot{p}(\cdot)$ weakly in $L^1[a, b]$ for $k \to \infty$ (as usual we take all $k = 1, 2, \ldots$).

Let us rewrite (5.12) as follows:

$$(6.11) \qquad \dot{p}(t) \in \{u | (u, p^k(t_{j+1})) - \lambda_0^k \theta^k(t)) \in \lambda_0^k \partial f(\bar{x}^k(t_j), \dot{\bar{x}}^k(t), t_j) +$$

$$N((\bar{x}^k(t_j), \dot{\bar{x}}^k(t)); \mathrm{gph} F(\cdot, t_j))\} \quad \text{for} \quad t \in [t_j, t_{j+1}), \ j = 0, \ldots, k - 1.$$

According to the classical results, there is a sequence of convex combinations of $\dot{p}^k(t)$ which converges to $\dot{p}(t)$ for a.e. $t \in [a, b]$. Now passing to the limit in (6.6) as $k \to \infty$ and using (6.11) as well as hypotheses (H6) and (H7), we obtain the Euler-Lagrange inclusion (6.4).

Passing to the limit in (5.9) and (5.10) and taking into account that $\rho_i^k \to 0$ as $k \to \infty$, we get, respectively, the sign and complementary slackness conditions (6.1) and (6.2). Taking the limit in (5.32), one has the normalization condition

$$|p(b)| + \sum_{i=0}^{q} \lambda_i + \sum_{i=q+1}^{q+r} |\lambda_i| = 1$$

which implies that $\lambda_0, \ldots, \lambda_{q+r}$ and $p(\cdot)$ are not equal to zero simultaneously. It follows from (6.10) that if $p(t_0) = 0$ at some point $t_0 \in [a, b]$, then $p(t) \equiv 0$ in $[a, b]$.

It remains to establish the transversality inclusion (6.3). First note that

$$\partial(\sum_{i=0}^{q+r}\lambda_i^k\varphi_i)(\bar{x}_0^k, \bar{x}_k^k) \to \partial(\sum_{i=0}^{q+r}\lambda_i\varphi_i)(\bar{x}(a), \bar{x}(b)) \quad \text{as} \quad k \to \infty$$

due to robustness of the subdifferential (4.5). Then observe that the set Ω_k in (3.20) is represented as

$$(6.12) \qquad \Omega_k = \{(x_0, x_k) \in \mathbf{R}^{2n} |\ \text{dist}((x_0, x_k), \Omega) \le \eta_k\}.$$

Now passing to the limit in (5.11) as $k \to \infty$ and using Proposition 4.4 in (6.12), we obtain (6.3). This completes the proof of the theorem. $\square$

6.2. *Remark.*

The Euler-Lagrange inclusion (6.4) can be expressed in terms of the coderivative $D_x^\star F$ of the multifunction $F(\cdot, t)$ under fixed t. Indeed, one has

$$(6.13) \qquad \dot{p}(t) \in \text{co}\{ \bigcup_{(\vartheta,w)\in\partial f(\bar{x}(t),\dot{\bar{x}}(t),t)} [\lambda_0\vartheta + D_x^\star F(\bar{x}(t), \dot{\bar{x}}(t), t)(\lambda_0 w - p(t))]\}$$

for a.e. $t \in [a, b]$ which is equavalent to (6.4) by virtue of (4.8). Inclusion (6.13) implies the following one:

$$\dot{p}(t) \in \text{co}[\lambda_0\partial_x f(\bar{x}(t), \dot{\bar{x}}(t), t) + D_x^\star F(\bar{x}(t), \dot{\bar{x}}(t), t)(\lambda_0\partial_v f(\bar{x}(t), \dot{\bar{x}}(t), t) - p(t))]$$

in the case when $\partial f(\bar{x}(t), \dot{\bar{x}}(t), t) \subset \partial_x f(\bar{x}(t), \dot{\bar{x}}(t), t) \times \partial_v f(\bar{x}(t), \dot{\bar{x}}(t), t)$ for a.e. $t \in [a, b]$. In particular, it happens when either $f(x, v, t) = f_1(x, t) + f_2(v, t)$ or f possesses some regularity (e.g., f is smooth or convex in (x, v)).

6.3. *Remark.*

The transversality inclusion (6.3) can be expressed in the more general form

$$(6.14) \qquad (p(a), -p(b)) \in \partial L_\Omega(\cdot, \lambda_0, \ldots, \lambda_{q+r})(\bar{x}(a), \bar{x}(b))$$

in terms of the *essential endpoint Lagrangian*

$$(6.15) \qquad L_\Omega(x_0, x_1, \lambda_0, \ldots, \lambda_{q+r}) := \sum_{i=0}^{q+r}\lambda_i\varphi_i(x_0, x_1) + \delta((x_0, x_1), \Omega);$$

see Theorem 7.1 stated below. On the other hand, (6.3) implies, thanks to Proposition 4.5, the following "separated" transversality inclusion

$$(p(a), -p(b)) \in \sum_{i=0}^{q}\lambda_i\partial\varphi_i(\bar{x}(a), \bar{x}(b)) + \sum_{i=q+1}^{q+r}\lambda_i\partial^0\varphi_i(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega)$$

where $\partial^0\varphi(\bar{x}) := \partial\varphi(\bar{x}) \cup [-\partial(-\varphi)(\bar{x})]$ is a "bilateral" differential construction. Note also that, using the procedure developed above, we can obtain transversality conditions under weaker assumptions on φ_i (lower semicontinuity for $i = 0, \ldots, q$ and continuity for

$i = q + 1, \ldots, q + r$). Such conditions are expressed in terms of the normal cone (4.2) to graphical (epigraphical) sets associated with the functions φ_i; cf. [29, Section 7] and the next section.

The necessary optimality conditions in Theorem 6.1 are proved for any i.r.l.m. in the Bolza problem (P). In particular, they hold for any feasible solution to (P) which appears to be a strong local minimum for the relaxed problem. As we pointed out in Section 2, an optimal solution to (P) automatically solves the relaxed problem as well in some common settings. Let us present such a corollary of Theorem 6.1 which is used in the next section for obtaining the refined Euler-Lagrange conditions without any relaxation.

6.4. COROLLARY. *Let the arc $\bar{x}(\cdot)$ provide a strong local minumum for the Bolza problem* (1.1), (1.3)–(1.5) *with no differential inclusion. Suppose that for some number $\mu > 0$ and open set $U \in \mathbf{R}^n$ one has:*

$$(6.16) \qquad \bar{x}(t) \in U \ \ \forall t \in [a, b] \ \text{ and } \ |\dot{\bar{x}}(t)| < \mu \ \text{ a.e. } \ t \in [a, b];$$

$f(x, v, \cdot)$ is bounded and continuous for a.e. $t \in [a, b]$ uniformly in $(x, v) \in U \times (\mu B)$; $f(\cdot, \cdot, t)$ is Lipschitz continuous on $U \times (\mu B)$ uniformly in $t \in [a, b]$; and (H5'), (H6) *hold. Then there exist real numbers $\lambda_0, \ldots, \lambda_{q+r}$ and an absolutely continuous function $p : [a, b] \to \mathbf{R}^n$, not all zero, such that conditions* (6.1)–(6.3) *are fulfilled and*

$$(6.17) \qquad \dot{p}(t) \in \mathrm{co}\{u | (u, p(t)) \in \lambda_0 \partial f(\bar{x}(t), \dot{\bar{x}}(t), t)\} \ \text{ a.e. } \ t \in [a, b].$$

Proof. The boundedness of $\dot{\bar{x}}(t)$ in (6.16) means that $\bar{x}(\cdot) \in W^{1,\infty}[a, b]$. According to Proposition 2.3, $\bar{x}(\cdot)$ is a strong minimum for the relaxed problem (1.3)–(1.5), (2.4). Let us consider the (trivial) differential inclusion

$$(6.18) \qquad \dot{x} \in F(x, t) := \mu B \ \text{ a.e. } \ t \in [a, b].$$

It is obvious that $\bar{x}(\cdot)$ is an i.r.l.m. for problem (P) in (1.3)–(1.5), (6.18) where all the assumptions of Theorem 6.1 are fulfilled. Moreover, $(\bar{x}(t), \dot{\bar{x}}(t)) \in (\mathrm{int} \ \mathrm{gph} F)$ for a.e. $t \in [a, b]$. In this case, (6.4) is equivalent to (6.17). $\square$

7. The Euler-Lagrange Inclusion without Relaxation

In the concluding section of the paper we consider the original problem (P) in (1.1)–(1.5) with $f = 0$. We shall denote this Mayer problem for differential inclusions by (P_M). The main objection of this section is to prove the refined Euler-Lagrange conditions for a *strong* local minimum in (P_M) *without any relaxation*. To accomplish it, we employ the results in the previous section (namely Corollary 6.4) and an additional approximation procedure combining ideas in [8, 23, 29]. The latter procedure allows us to approximate the Mayer problem under consideration by a sequence of nonsmooth Bolza problems without differential

inclusions and any endpoint constraints. In this way we also get the refined Euler-Lagrange inclusion for any boundary trajectories of nonconvex differential inclusions.

Keeping here assumptions (H1), (H2), and (H6) on F around the given trajectory for (1.2), we relax the Lipschitzian assumptions on φ_i. Namely, we assume that

($\overline{\text{H5}}$) The functions φ_i are lower semicontinuous for $i = 0, \ldots, q$ and continuous for $i = q + 1, \ldots, q + r$; the set Ω is closed around $(\bar{x}(a), \bar{x}(b))$.

Consider the closed set

$$(7.1) \qquad \mathcal{E}_\Omega := \{(x_0, x_1, \nu_0, \ldots, \nu_{q+r}) \in \mathbf{R}^{2n+q+r+1} | (x_0, x_1) \in \Omega,$$

$$\varphi_i(x_0, x_1) \le \nu_i \text{ for } i = 0, \ldots, q \ \& \ \varphi_i(x_0, x_1) = \nu_i \text{ for } i = q+1, \ldots, q+r\}.$$

7.1. THEOREM. *Let $\bar{x}(\cdot)$ be a strong minimum for the Mayer problem (P_M) in (1.1)–(1.5) under assumptions (H1), (H2), ($\overline{\text{H5}}$), and (H7). Then there exist a vector $y^\star = (\lambda_0, \ldots, \lambda_{q+r}) \in \mathbf{R}^{q+r+1}$ and an absolutely continuous function $p : [a, b] \to \mathbf{R}^n$, not both zero, such that*

$$(7.2) \qquad \dot{p}(t) \in \mathrm{co} D_x^\star F(\bar{x}(t), \dot{\bar{x}}(t), t)(-p(t)) \text{ a.e } t \in [a, b],$$

$$(7.3) \qquad (p(a), -p(b), -y^\star) \in N((\bar{x}(a), \bar{x}(b), \bar{c}); \mathcal{E}_\Omega)$$

where $\bar{c} := (\varphi_0(\bar{x}(a), \bar{x}(b))$. Condition (7.3) always implies (6.1) and (6.2) being equivalent to the simultaneous fulfilment of (6.1), (6.2), and (6.14) if all φ_i are Lipschitz continuous around $(\bar{x}(a), \bar{x}(b))$.

Proof. Let $\bar{\gamma} := \varphi_0(\bar{x}(a), \bar{x}(b))$. According to the metric approximation method in Mordukhovich [27–29], we consider the parametric functional

$$(7.4) \quad I_\gamma[x] := \mathrm{dist}((x(a), x(b), c), \mathcal{E}_\Omega) \text{ with } c := (\gamma, 0, \ldots, 0) \in \mathbf{R}^{q+r+1}, \ \gamma \in \mathbf{R}$$

on the trajectories for the differential inclusion (1.2). Let $U \in \mathbf{R}^n$ be a bounded neighborhood of the strong minimum $\bar{x}(\cdot)$ where assumptions (H1), (H2), and ($\overline{\text{H5}}$) are fulfilled. For every $\varepsilon > 0$, one has

$$I_\gamma[\bar{x}] \le |\gamma - \bar{\gamma}| < \varepsilon$$

if γ is close to $\bar{\gamma}$. On the other hand,

$$I[x] > 0 \text{ for any } \gamma < \bar{\gamma}$$

whenever the trajectory $x(\cdot)$ for (1.2) belongs to the neighborhood U of the strong minimum $\bar{x}(\cdot)$.

Following Clarke [8], let us consider the set X of all trajectories $x(\cdot)$ for (1.2) satisfying $x(t) \in \mathrm{cl}U$ in $[a, b]$ and let us define a metric in X as follows

$$(7.5) \qquad d(x, y) := |x(a) - y(a)| + \int_a^b |\dot{x}(t) - \dot{y}(t)| dt.$$

One can easily see that X is a complete metric space with metric (7.5) and functional (7.4) is continuous in X for any γ. According to the constructions above, for every $\varepsilon > 0$ we find $\gamma_\varepsilon < \bar{\gamma}$ such that $\gamma_\varepsilon \to \bar{\gamma}$ as $\varepsilon \to 0$, $I_\varepsilon[\bar{x}] < \varepsilon$, and

$$(7.6) \qquad I_\varepsilon[x] > 0 \quad \text{for} \quad I_\varepsilon := I_{\gamma_\varepsilon}$$

where $x(\cdot)$ is any trajectory for (1.2) with $x(t) \in U$ in $[a, b]$. Now one can apply the Ekeland variational principle [13] and claim the existence of $x_\varepsilon(\cdot)$ such that

$$(7.7) \qquad d(x_\varepsilon, \bar{x}) \le \sqrt{\varepsilon} \quad \text{and}$$

$$I_\varepsilon[x] + \sqrt{\varepsilon} d(x, x_\varepsilon) \ge I_\varepsilon[x_\varepsilon] \quad \forall x(\cdot) \in X.$$

Note that (7.7) implies $x_\varepsilon(t) \in U$ for ε small enough, so $I_\varepsilon[x_\varepsilon] > 0$ by virtue of (7.6).

Now for any positive numbers M, ε and the Lipschitz constant l_F in (3.2), we define the functional

$$(7.8) \quad J_\varepsilon^M[x] := I_\varepsilon[x] + \sqrt{\varepsilon} d(x, x_\varepsilon) + M(1 + l_F^2)^{1/2} \int_a^b \mathrm{dist}((x(t), \dot{x}(t)), \mathrm{gph}F(\cdot, t)) dt$$

on the set of all arcs $x(\cdot)$ (not necessarily trajectories for (1.2)) satisfying $x(t) \in U$ in $[a, b]$. We omit the proof of the following lemma which can be furnished by the arguments in Kaskosz and Lojasiewicz [23, Lemmas 1 and 2].

7.2. LEMMA. *There exists a number $M \ge 1$ such that for every $\varepsilon \in (0, 1/M)$ the arc $x_\varepsilon(\cdot)$ provides an unconditional strong local minimum for the functional* (7.8).

Let us continue the proof of Theorem 7.1. Setting $c_\varepsilon := (\gamma_\varepsilon, 0, \ldots, 0)$, we consider any element $(z_{0\varepsilon}, z_{1\varepsilon}, e_\varepsilon) \in \Pi((x_\varepsilon(a), x_\varepsilon(b), c_\varepsilon), \mathcal{E}_\Omega)$ from the Euclidean projector (4.1) of $(x_\varepsilon(a), x_\varepsilon(b), c_\varepsilon)$ on the set $\mathcal{E}_\Omega$. Using this projection and Lemma 7.2, one can conclude that $x_\varepsilon(\cdot)$ provides a strong local minimun for the unconstrained Bolza problem

$$(7.9) \qquad \text{minimize} \ \varphi_\varepsilon(x(a), x(b)) + \int_a^b f_\varepsilon(x(t), \dot{x}(t), t) dt$$

with the endpoint function

$$(7.10) \quad \varphi_\varepsilon(x_0, x_1) := [|x_0 - z_{0\varepsilon}|^2 + |x_1 - z_{1\varepsilon}|^2 + |c_\varepsilon - e_\varepsilon|^2]^{1/2} + \sqrt{\varepsilon}|x_0 - x_\varepsilon(a)|$$

and the integrand

$$(7.11) \qquad f_\varepsilon(x, v, t) := M(1 + l_F^2)^{1/2} \mathrm{dist}((x, v), \mathrm{gph}F(\cdot, t)) + \sqrt{\varepsilon}|v - \dot{x}(t)|.$$

Now we employ Corollary 6.4 in the unconstrained Bolza problem (7.9)–(7.11). Let us observe that the assumptions in Theorem 7.1 ensure the fulfilment of the assumptions in Corollary 6.4 around the solution $x_\varepsilon(\cdot)$ and the first term in (7.10) is smooth around $(x_\varepsilon(a), x_\varepsilon(b))$ by virtue of (7.6).

For each $\varepsilon > 0$, using the result in Corollary 6.4 with $\lambda_0 = 1$ and also Propositions 4.4 and 4.5 for functions (7.10) and (7.11), we find an arc $p_\varepsilon(\cdot)$ such that

$$(7.12) \qquad \dot{p}_\varepsilon(t) \in \mathrm{co}\{u | (u, p(t)) \in N((x_\varepsilon(t), \dot{x}_\varepsilon(t)); \mathrm{gph}F(\cdot, t)) + \sqrt{\varepsilon}(0, B)\}$$

for a.e. $t \in [a, b]$ and

$$(7.13) \qquad p_\varepsilon(a) = (x_\varepsilon(a) - z_{0\varepsilon})/\alpha_\varepsilon + \sqrt{\varepsilon}, \quad -p_\varepsilon(b) = (x_\varepsilon(b) - z_{1\varepsilon})/\alpha_\varepsilon$$

where $\alpha_\varepsilon := [|x_\varepsilon(a) - z_{0\varepsilon}|^2 + |x_\varepsilon(b) - z_{1\varepsilon}|^2 + |c_\varepsilon - e_\varepsilon|^2]^{1/2} > 0$.

Denoting $y_\varepsilon^\star := (e_\varepsilon - c_\varepsilon)/\alpha\varepsilon$, one has

$$(7.14) \qquad |p_\varepsilon(a)|^2 + |p_\varepsilon(b)|^2 + |y_\varepsilon^\star|^2 = 1 \quad \text{and}$$

$$(7.15) \quad (p_\varepsilon(a), -p_\varepsilon(b), -y_\varepsilon^\star) \in \mathrm{cone}[(x_\varepsilon(a), x_\varepsilon(b), c_\varepsilon) - \Pi((x_\varepsilon(a), x_\varepsilon(b), c_\varepsilon), \mathcal{E}_\Omega)].$$

Now let us consider the limiting procedure in (7.12)–(7.15) as $\varepsilon \to 0$. By virtue of (7.5), relationship (7.7) means that $x_\varepsilon(\cdot) \to \bar{x}(\cdot)$ in $W^{1,\infty}[a, b]$ as $\varepsilon \to 0$. This implies that $x_\varepsilon(t) \to \bar{x}(t)$ uniformly in $[a, b]$ and $\dot{x}_\varepsilon(t) \to \dot{\bar{x}}(t)$ for a.e. $t \in [a, b]$.

Further, using (7.12), (7.14), (4.8), and Corollary 4.14(ii), one can conclude (cf. the proof of Theorem 6.1) that there exist an arc $p(\cdot)$ and a vector $y^\star \in \mathbf{R}^{q+r+1}$, not both zero, such that $y_\varepsilon^\star$ converges to $y^\star$, $p_\varepsilon(t)$ converges to $p(t)$ uniformly in $[a, b]$, and a convex combination of $\dot{p}_\varepsilon(t)$ converges to $\dot{p}(t)$ for a.e. $t \in [a, b]$ as $\varepsilon \to 0$ along some subsequence.

Now passing to the limit in (7.12), (7.14), (7.15) and using Definitions 4.1 and 4.7 as well as robustness of the normal cone (4.2), one gets the main conclusions (7.2) and (7.3) of the theorem. Representing $y^\star = (\lambda_0, \ldots, \lambda_{q+r})$, we obtain (6.1) and (6.2) directly from Proposition 4.9(i) where

$$\mathrm{gph}F = \mathcal{E}_\Omega, \quad g = (\varphi_0, \ldots, \varphi_{q+r}), \quad \bar{x} = (\bar{x}(a), \bar{x}(b)), \quad \bar{y} = (0, \ldots, 0),$$

$$\Delta = \{(\mu_0, \ldots, \mu_{q+r}) | \mu_i \leq 0 \text{ for } i = 0, \ldots, q \ \& \ \mu_i = 0 \text{ for } i = q+1, \ldots, q+r\}.$$

If all φ_i are locally Lipschitzian around $(\bar{x}(a), \bar{x}(b))$, then the equivalence of (7.3) to the simultaneous fulfilment of (6.1), (6.2), and (6.14) follows from Proposition 4.9(ii) where the scalarization function (4.10) is reduced to the essential endpoint Lagrangian (6.15) in the case under condideration. This ends the proof of the theorem. $\square$

Let us consider an analogue of the result obtained for the case of *boundary trajectories*. Given a nonempty closed set $A \subset \mathbf{R}^n$, we denote by $R(A)$ the *reachable set* for (1.2) from

A, i.e., the set of all points $x(b)$ where $x(\cdot)$ is a trajectory for (1.2) with $x(a) \in A$.

7.3. THEOREM. *Let $\bar{x}(\cdot)$ be a trajectory for (1.2) with $\bar{x}(a) \in A$ and let assumptions (H1), (H2), (H7) hold. If $g : \mathbf{R}^n \to \mathbf{R}^m$ is a locally Lipschitzian function around $\bar{x}(b)$ and if $g(\bar{x}(b))$ is a boundary point of the set $R(A)$, then there exist an arc $p : [a, b] \to \mathbf{R}^n$ and a unit vector $\psi \in \mathbf{R}^m$ such that $p(\cdot)$ satisfies the refined Euler-Lagrange inclusion (7.2) with the following boundary (transversality) conditions:*

$$(7.16) \qquad p(a) \in N(\bar{x}(a); A), \quad -p(b) \in \partial\langle\psi, g\rangle(\bar{x}(b)).$$

7.4. *Remark.* The result formulated generalizes the recent one in Kaskosz and Lojasiewicz [23] where (7.2) is replaced by Clarke's form of the Euler-Lagrange inclusion (see (1.6) as $f = 0$) and conditions (7.16) are replaced by

$$p(a) \in N_C(\bar{x}(a); A), \quad p(b) \in [J_C g(\bar{x}(b))]^* \psi$$

in terms of Clarke's normal cone and generalized Jacobian; cf. (4.3) and (4.13).

Proof of Theorem 7.3. Following the arguments in [23], for every $\varepsilon > 0$ we can find a vector $c_\varepsilon \in \mathbf{R}^m$ and a trajectory $x_\varepsilon(\cdot)$ of (1.2) with $x_\varepsilon(a) \in A$ such that $|g(x_\varepsilon(b)) - c_\varepsilon| > 0$,

$$c_\varepsilon \to g(\bar{x}(b)), \quad x_\varepsilon(\cdot) \to \bar{x}(\cdot) \text{ in } W^{1,\infty}[a, b] \text{ as } \varepsilon \to 0,$$

and $x_\varepsilon(\cdot)$ provides an unconditional strong local minimum for problem (7.9) with integrand (7.11) and the endpoint functional

$$(7.17) \qquad \varphi_\varepsilon(x_0, x_1) := |g(x_1) - c_\varepsilon| + \sqrt{\varepsilon}|x_0 - x_\varepsilon(a)| + M\text{dist}(x_0, A).$$

Now employing Corollary 6.4 in problem (7.9), (7.11), (7.17) and using Propositions 4.4–4.6 for computing the subdifferential of (7.17), one gets an arc $p_\varepsilon(\cdot)$ and a unit vector $\psi_\varepsilon \in \mathbf{R}^m$ such that (7.12) holds and

$$(7.18) \qquad p_\varepsilon(a) \in \sqrt{\varepsilon}B + N(x_\varepsilon(a); A), \quad -p_\varepsilon(b) \in \partial\langle\psi_\varepsilon, g\rangle(x_\varepsilon(b)).$$

Following the proof of Theorem 7.1 above, we obtain conditions (7.2) and (7.16) by passing to the limit in (7.12) and (7.18) as $\varepsilon \to 0$. $\square$

7.5. *Remark.* Using the method of metric approximation (as in proof of Theorem 7.1), one can extend Theorem 7.3 to a more general setting when $g(\bar{x}(b))$ is a *locally extremal point* of the set $R(A)$ relative to other given sets. We refer to [29, Section 6] and [34, Section 3] for more details about this concept.

References

[1] V. M. ALEKSEEV, V. M. TIKHOMIROV AND S. V. FOMIN, *Optimal Control*, Nauka, Moscow, 1979. (Russian; English transl. in Consultants Bureau, New York, 1987)

[2] J.-P. AUBIN, *Lipschtz behavior of solutions to convex minimization problems*, Math. Oper. Res., 9 (1984), pp. 87–111.

[3] J.-P. AUBIN AND H. FRANKOWSKA, *Set-Valued Analysis*, Birhäuser, Boston, 1990.

[4] N. N. BOGOLJUBOV, *Sur quelques methodes nouvelles dans le calculus des variations*, Ann. Math. Pura Appl., 7 (1930), pp. 249–271.

[5] L. CESARI, *Optimization–Theory and Applications: Problems with Ordinary Differential Equations*, Springer-Verlag, New York, 1983.

[6] F. H. CLARKE, *Optimal solutions to differential inclusions*, J. Optimization Theory Appl., 19 (1976), pp. 469–478.

[7] F.H. CLARKE, *The generalized problem of Bolza*, SIAM J. Control Optim. 14 (1976), pp. 682–699.

[8] F. H. CLARKE, *Necessary conditions for a general control problem*, in Calculus of Variations and Control Theory, D. Russel, ed., Academic Press, New York, 1976, pp. 257–278.

[9] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.

[10] F. H. CLARKE, *Hamiltonian analysis of the generalized problem of Bolza*, Trans. Amer. Math. Soc., 301 (1987), pp. 385–400.

[11] F. H. CLARKE, *Methods of Dynamic and Nonsmooth Optimization* , SIAM Publications, Philadelphia, 1989.

[12] M. G. CRANDALL AND P.-L. LIONS, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277 (1983), pp. 1–42.

[13] I. EKELAND, *On the variational principle*, J. Math. Anal. Appl., 47 (1974), pp. 324–353.

[14] I. EKELAND AND R. TEMAN, *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, 1976.

[15] A. L. DONTCHEV AND E. M. FARKHI, *Error estimates for dicretized differential inclusions*, Computing, 41 (1989), pp. 349–358.

[16] A. L. DONTCHEV AND F. LEMPIO, *Difference methods for differential inclusions: a survey*, SIAM Review, 34 (1992), pp. 263–294.

[17] H. FRANKOWSKA, *The maximum principle for an optimal solution to a differential inclusion with endpoint constraints*, SIAM J. Control Optim., 25 (1987), pp. 145–157.

[18] A. D. IOFFE, *Approximate subdifferentials and applications*, I: *the finite dimensional theory*, Trans. Amer. Math. Soc., 281 (1984), pp. 389–416.

[19] A. D. IOFFE, *Approximate subdifferentials and applications*, III: *the metric theory*, Mathematika, 36 (1989), pp. 1–38.

[20] A. D. IOFFE, *Nonsmooth subdifferentials: their calculus and applications*, Proceed. 1st World Congress of Nonlinear Analysts, to appear.

[21] A. D. IOFFE AND V. M. TIKHOMIROV, *Extensions of variational problems*, Trans. Moscow Math. Soc., 18 (1968), pp. 187–246. (Russian)

[22] A.D. IOFFE AND V. M. TIKHOMIROV, *Theory of Extremal Problems*, Nauka, Moscow, 1974. (Russian; English transl. in North-Holland, Amsterdam, 1979)

[23] B. KASKOSZ AND S. LOJASIEWICZ JR., *Lagrange-type extremal trajectories in differential inclusions*, Systems and Control Lett., 19 (1992), pp. 241–247.

[24] P. D. LOEWEN, *On the Lavrentiev phenomenon*, Canad. Math. Bull., 30 (1987), pp. 102–108.

[25] P. D. LOEWEN AND R. T. ROCKAFELLAR, *The adjoint arc in nonsmooth optimization*, Trans. Amer. Math. Soc., 325 (1991), pp. 39–72.

[26] P. D. LOEWEN AND R. T. ROCKAFELLAR, *Optimal control of unbounded differential inclusions*, SIAM J. Control Optim., to appear.

[27] B. MORDUKHOVICH, *The maximum principle in the problem of time-optimal control with nonsmooth constraints*, J. Appl. Math. Mech., 40 (1976), pp. 960–969.

[28] B. MORDUKHOVICH, *Metric approximations and necessary optimality conditions for general classes of nonsmooth extremal problems*, Soviet Math. Dokl. 22 (1980), pp. 526–530.

[29] B. MORDUKHOVICH, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988. (Russian; English transl. to appear in Wiley-Interscience)

[30] B. MORDUKHOVICH, *On variational analysis of differential inclusions*, in Optimization and Nonlinear Analysis, A. Ioffe et al., eds., Pitman Res. Notes Math. Ser., 244 (1992), pp. 199–213.

[31] B. MORDUKHOVICH, *Sensitivity analysis in nonsmooth optimization*, in Theoretical Aspects of Industrial Design, D. A. Field and V. Komkow, eds., SIAM Publications, Philadelphia, 1992, pp. 32–46.

[32] B. MORDUKHOVICH, *Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc., to appear.

[33] B. MORDUKHOVICH, *Lipschitzian stability of constraint systems and generalized equations*, Nonlinear Anal. Theory Methods Appl., to appear.

[34] B. MORDUKHOVICH, *Generalized differential calculus for nonsmooth and set-valued mappings*, J. Math. Anal. Appl., to appear.

[35] B. MORDUKHOVICH, *Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis*, IMA Preprints Series # 997, July 1992.

[36] E. POLAK, *Computational Methods in Optimization: A Unified Approach*, Academic Press, New York, 1971. (The 2nd revised edition to be published by Academic Press in 1993)

[37] E. S. POLOVINKIN AND G. V. SMIRNOV, *An approach to differentiation of multifunctions and necessary optimality conditions for differential inclusions*, Differn. Equations, 22 (1986), pp. 660–668.

[38] B. N. PSHENICHNYI, *Convex Analysis and Extremal Problems*, Nauka, Moscow, 1980. (Russian)

[39] B. SENDOV AND V. A. POPOV, *The averaged moduli of smoothness* , Wiley-Interscience, New York, 1988.

[40] G. V. SMIRNOV, *Discrete approximations and optimal solutions to differential inclusions*, Kibernetika, 1991, pp. 76–79. (Russian)

[41] R. T. ROCKAFELLAR, *Conjugate convex functions in optimal control and the calculus of variations*, J. Math. Anal. Appl., 32 (1970), pp. 174–222.

[42] R. T. ROCKAFELLAR, *Extensions of subgradient calculus with applications to optimization*, Nonlinear Anal. Theory Methods Appl., 9 (1985), pp. 665–698.

[43] R. T. ROCKAFELLAR, *Lipschitzian properties of multifunctions*, Nonlinear Anal. Theory Methods Appl., 9 (1985), pp. 867–885.

[44] R. T. ROCKAFELLAR, *Maximal monotone relations and the second derivatives of nonsmooth functions*, Ann. Inst. H. Poincaré. Anal. Non Linéare, 2 (1985), pp. 167–184.

[45] R. T. ROCKAFELLAR, *Dualization of subgradient conditions for optimality*, Nonlinear Anal. Theory Methods Appl., to appear.

[46] R. T. ROCKAFELLAR, *Lagrange multipliers and optimality*, SIAM Review, to appear.

[47] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Variational Analysis*, book in preparation.

[48] J. D. L. ROWLAND AND R. B. VINTER, *Dynamic optimization problems with free time and active state constraints*, SIAM J. Control Optim., to appear.

[49] J. WARGA, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

[50] J. WARGA, *Controllability, extremality, and abnormality in nonsmooth optimal control*, J. Optimization Theory Appl., 41 (1983), pp. 239–260.

[51] L. C. YOUNG, *Lectures on the Calculus of Variations and Optimal Control Theory*, Saunders, Philadelphia, 1969.

Recent IMA Preprints

#	Author/s	Title

1041 Neerchal K. Nagaraj and Wayne A. Fuller, Least squares estimation of the linear model with autoregressive errors

1042 H.J. Sussmann & W. Liu, A characterization of continuous dependence of trajectories with respect to the input for control-affine systems

1043 Karen Rudie & W. Murray Wonham, Protocol verification using discrete-event systems

1044 Rohan Abeyaratne & James K. Knowles, Nucleation, kinetics and admissibility criteria for propagating phase boundaries

1045 Gang Bao & William W. Symes, Computation of pseudo-differential operators

1046 Srdjan Stojanovic, Nonsmooth analysis and shape optimization in flow problem

1047 Miroslav Tuma, Row ordering in sparse QR decomposition

1048 Onur Toker & Hitay Özbay, On the computation of suboptimal H^∞ controllers for unstable infinite dimensional systems

1049 Hitay Özbay, H^∞ optimal controller design for a class of distributed parameter systems

1050 J.E. Dunn & Roger Fosdick, The Weierstrass condition for a special class of elastic materials

1051 Bei Hu & Jianhua Zhang, A free boundary problem arising in the modeling of internal oxidation of binary alloys

1052 Eduard Feireisl & Enrique Zuazua, Global attractors for semilinear wave equations with locally distributed nonlinear damping and critical exponent

1053 I-Heng McComb & Chjan C. Lim, Stability of equilibria for a class of time-reversible, $D_n x O(2)$-symmetric homogeneous vector fields

1054 Ruben D. Spies, A state-space approach to a one-dimensional mathematical model for the dynamics of phase transitions in pseudoelastic materials

1055 H.S. Dumas, F. Golse, and P. Lochak, Multiphase averaging for generalized flows on manifolds

1056 Bei Hu & Hong-Ming Yin, Global solutions and quenching to a class of quasilinear parabolic equations

1057 Zhangxin Chen, Projection finite element methods for semiconductor device equations

1058 Peter Guttorp, Statistical analysis of biological monitoring data

1059 Wensheng Liu & Héctor J. Sussmann, Abnormal sub-Riemannian minimizers

1060 Chjan C. Lim, A combinatorial perturbation method and Arnold's whiskered Tori in vortex dynamics

1061 Yong Liu, Axially symmetric jet flows arising from high speed fiber coating

1062 Li Qiu & Tongwen Chen, $\mathcal{H}_2$ and $\mathcal{H}_\infty$ designs of multirate sampled-data systems

1063 Eduardo Casas & Jiongmin Yong, Maximum principle for state-constrained optimal control problems governed by quasilinear elliptic equations

1064 Suzanne M. Lenhart & Jiongmin Yong, Optimal control for degenerate parabolic equations with logistic growth

1065 Suzanne Lenhart, Optimal control of a convective-diffusive fluid problem

1066 Enrique Zuazua, Weakly nonlinear large time behavior in scalar convection-diffusion equations

1067 Caroline Fabre, Jean-Pierre Puel & Enrike Zuazua, Approximate controllability of the semilinear heat equation

1068 M. Escobedo, J.L. Vazquez & Enrike Zuazua, Entropy solutions for diffusion-convection equations with partial diffusivity

1069 M. Escobedo, J.L. Vazquez & Enrike Zuazua, A diffusion-convection equation in several space dimensions

1070 F. Fagnani & J.C. Willems, Symmetries of differential systems

1071 Zhangxin Chen, Bernardo Cockburn, Joseph W. Jerome & Chi-Wang Shu, Mixed-RKDG finite element methods for the 2-D hydrodynamic model for semiconductor device simulation

1072 M.E. Bradley & Suzanne Lenhart, Bilinear optimal control of a Kirchhoff plate

1073 Héctor J. Sussmann, A cornucopia of abnormal subriemannian minimizers. Part I: The four-dimensional case

1074 Marek Rakowski, Transfer function approach to disturbance decoupling problem

1075 Yuncheng You, Optimal control of Ginzburg-Landau equation for superconductivity

1076 Yuncheng You, Global dynamics of dissipative modified Korteweg-de Vries equations

1077 Mario Taboada & Yuncheng You, Nonuniformly attracting inertial manifolds and stabilization of beam equations with structural and Balakrishnan-Taylor damping

1078 Michael Böhm & Mario Taboada, Global existence and regularity of solutions of the nonlinear string equation

1079 Zhangxin Chen, BDM mixed methods for a nonlinear elliptic problem

1080 J.J.L. Velázquez, On the dynamics of a closed thermosyphon

1081 Frédéric Bonnans & Eduardo Casas, Some stability concepts and their applications in optimal control problems

1082 Hong-Ming Yin, $\mathcal{L}^{2,\mu}(Q)$-estimates for parabolic equations and applications

1083 David L. Russell & Bing-Yu Zhang, Smoothing and decay properties of solutions of the Korteweg-de Vries equation on a periodic domain with point dissipation

1084 J.E. Dunn & K.R. Rajagopal, Fluids of differential type: Critical review and thermodynamic analysis

1085 Mary Elizabeth Bradley & Mary Ann Horn, Global stabilization of the von Kármán plate with boundary