

Provided for non-commercial research and educational use. Not for reproduction, distribution or commercial use.
--

Serdica

Mathematical Journal

Сердика

Математическо списание

The attached copy is furnished for non-commercial research and education use only.

Authors are permitted to post this version of the article to their personal websites or institutional repositories and to share with other researchers in the form of electronic reprints.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to third party websites are prohibited.

For further information on
Serdica Mathematical Journal
which is the new series of
Serdica Bulgaricae Mathematicae Publicationes
visit the website of the journal <http://serdica.math.bas.bg>
or contact: Editorial Office
Serdica Mathematical Journal
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Telephone: (+359-2)9792818, FAX:(+359-2)971-36-49
e-mail: serdica@math.bas.bg

ON NECESSARY CONDITIONS IN THE GENERALIZED BOLZA PROBLEM

Alexander D. Ioffe

Communicated by N. Ribarska

9-32 *To the memory of Asen Dontchev and Andrzej Granas*

ABSTRACT. The paper is devoted to proving necessary optimality conditions for weak/intermediate/strong minimum of generalized Bolza problem.

1. Introduction. We shall consider the problem of minimizing the functional

$$J(x(\cdot)) = \ell(x(0), x(T)) + \varphi(x(\cdot)) + \int_0^T L(t, x(t), \dot{x}(t)) dt$$

on the space $W^{1,1}$ of absolutely continuous $\mathbb{R}^n$ -valued functions $x(\cdot)$ on $[0, T]$. Here

- ℓ is a (possibly extended real valued) function on $\mathbb{R}^n \times \mathbb{R}^n$ which we shall assume lower semicontinuous;

2020 *Mathematics Subject Classification.* 49K15, 49K21, 49K27.

Key words: Bolza problem, Pontryagin maximum principle, variational analysis.

- φ is a locally Lipschitz function on the space of continuous mappings from $[0, T]$ into $\mathbb{R}^n$ and
- L is (in general also extended real valued) integrand whose properties will be specified in the statements of the results to be proved.

The problem was first considered by Rockafellar in 1971 [18] with both ℓ and $L(t, \cdot)$ convex. The middle term $\varphi(\cdot)$ was absent in [18] and actually in all publications up to [12]. Since the very beginning the connection between the generalized Bolza problem and problems of optimal control was well recognized. In the original works of Rockafellar [18] and Clarke [1, 2] and in the subsequent publications the study of necessary optimality conditions in the Bolza problem were all based on applications of suitable results for optimal control problems (or occasionally to even more general constrained optimization problems). But in [12] it was shown that, vice versa, the knowledge of necessary conditions in the generalized Bolza problem, even with ℓ satisfying the Lipschitz condition near the solution, can be effectively used to get necessary conditions in a very general class of optimal control problems. In particular, the φ term was introduced in [12] to make it possible to apply Bolza problems to analysis of optimal control problems with state constraints. We plan to return later to this connection, but here we shall be dealing exclusively with the Bolza problem with almost no mention of optimal control.

In what follows we shall fix some $\bar{x}(\cdot) \in W^{1,1}$ which will be assumed a local minimum of J in one or another sense. Namely, we shall consider necessary conditions for three types of a minimum of J : first the weak minimum with respect to the $W^{1,\infty}$ -topology, then the “intermediate” minimum with respect to the $W^{1,1}$ -topology and, finally, the strong minimum with respect to the topology of uniform convergence of $x(\cdot)$.

The necessary condition for the weak minimum (obviously necessary for the other two types of minimum) consists of two parts: an adjoint inclusion associated with the two last terms of J and transversality condition associated with ℓ . The adjoint inclusion that appears in necessary conditions for the first two types of minimum is usually called the inclusion in the Euler–Lagrange form. The necessary condition for the $W^{1,1}$ -minimum contains in addition a Weierstrass-type result which, unlike the classical Weierstrass condition, is valid only with values of the third variable having certain additional properties.

If L is convex with respect to the last argument, the Euler–Lagrange inclusion implies a Hamiltonian condition stated in terms of the Hamiltonian

$H(t, x, p)$ which is the Young-Fenchel conjugate of $L(t, x, \cdot)$:

$$H(t, x, p) = \sup_y (\langle p, y \rangle - L(t, x, y)).$$

But under certain additional assumptions the Hamiltonian condition, as a necessary condition for a strong minimum, can be also proved when $L(t, x, \cdot)$ is not convex, using relaxation of the original problem (replacement of L by its convexification with respect to the last argument). It should be said that the Hamiltonian condition appears in this case too as a consequence of the condition in the Euler-Lagrange form so that the latter can be viewed as the main result.

The plan of the paper is the following. The next section contains the statements of the main results and comments (with sufficient details) concerning interrelations of our theorems with earlier results. In the third section we have collected all necessary technical material basically associated with (a) the theory of subdifferentials and (b) relaxation. The role of relaxation has been just explained and the theory of subdifferential is the key element of any analysis of minima. In the fourth section we prove the theorem containing necessary conditions for the $W^{1,1}$ -minimum, and in the last fifth section we prove all other theorems stated in the next section.

Notation:

$p \in [1, \infty]$, $\|\cdot\|_p$ is the norm in L_p , $\|\cdot\|_{1,p}$ is the norm in $W^{1,p}$;

$\langle \cdot, \cdot \rangle$ stands for the inner product in $\mathbb{R}^n$;

$B(x, r)$ is the ball of radius r around x ;

∂ stands for the convex or limiting subdifferential in $\mathbb{R}^n$;

∂_G is the G -subdifferential in Banach spaces;

∂_c is Clarke's generalized gradient;

conv stands for "convex hull";

α^+ and α^- stand respectively for $\max\{0, \alpha\}$ and $\max\{0, -\alpha\}$;

$\mathcal{L} \otimes \mathcal{B}(R^k)$ is the σ -algebra generated by products of Lebesgue measurable subsets of $[0, T]$ and Borel subsets of $\mathbb{R}^k$.

2. Statements of main results. Let an $\bar{x}(\cdot) \in W^{1,1}$ be given such that $J(\bar{x}(\cdot)) < \infty$. Here are the main assumptions on the components of J :

(H₁) ℓ is a lower semicontinuous (possibly extended real valued) function on $\mathbb{R}^n \times \mathbb{R}^n$ finite at $(\bar{x}(0), \bar{x}(T))$;

- (H₂) φ is Lipschitz continuous in a neighborhood of $\bar{x}(\cdot)$ in the topology of uniform convergence;
- (H₃) $L(t, x, u)$ is a $\mathcal{L} \otimes \mathcal{B}(\mathbb{R}^{2n})$ -measurable function on $[0, T] \times \mathbb{R}^n \times \mathbb{R}^n$ with values in $(-\infty, \infty]$ which is lower semicontinuous as a function of the last two arguments for almost every t ;
- (H₄) There are positive-valued measurable $\varepsilon_0(t)$ and $\gamma_0(t)$ such that the set $A(t, x) = (\text{dom } L(t, x, \cdot) \cap B(\bar{x}(t), \gamma_0(t)))$ does not actually depend on x on $B(\bar{x}(t), \varepsilon_0(t))$ (so we can denote in just $A(t)$), and $A(t) = B(\bar{x}(t), \gamma_0(t))$ if ℓ is not Lipschitz in any neighborhood of $(\bar{x}(0), \bar{x}(T))$.
- (H₅) There is a summable positive-valued $k_0(t)$ such that the function $\sup\{|L(t, \bar{x}(t), y)| : y \in A(t)\}$ is summable and for almost every t and the function $L(t, \cdot, y)$ is $k_0(t)$ -Lipschitz on $B(\bar{x}(t), \varepsilon_0(t))$ for any $y \in A(t)$.

The first and the third assumptions are more or less standard for most of publications relating to the generalized Bolza problem.

Before stating the theorems, let us agree to say that *the problem is singular at $\bar{x}(\cdot)$* if there is a nonzero $b \in \mathbb{R}^n$ such that $(0, b, -b)$ belongs to the normal cone to the graph of ℓ at $(\ell(\bar{x}(0), \bar{x}(T)), \bar{x}(0), \bar{x}(T))$ and for almost every t the vector $(0, 0, b) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ belongs to the limiting normal cone to the graph of the function $L(t, \cdot, \cdot)$ at $(L(t, \bar{x}(t), \dot{\bar{x}}(t)), \bar{x}(t), \dot{\bar{x}}(t))$. It is clear from the definition that singularity is rather a rare phenomenon. In particular, the problem is not singular if ℓ is Lipschitz near $(\bar{x}(0), \bar{x}(T))$.

We start with necessary conditions for a weak minimum. In the statements of all four theorems below ∂ stands for the limiting subdifferential for functions on $\mathbb{R}^n$ (see the next section for the definition).

Theorem 2.1. *Assume (H₁)–(H₅). Let $\bar{x}(\cdot)$ be a local minimum of J in the $W^{1,\infty}$ -topology. If the problem is not singular at $\bar{x}(\cdot)$, then there are a Radon measure ν , a measurable $q(t)$ and a function $p(t)$ of bounded variation, all three with values in $\mathbb{R}^n$, such that $\nu \in \partial_G \varphi(\bar{x}(\cdot))$ and the relations*

$$(1) \quad q(t) \in \text{conv}\{q : (q, p(t)) \in \partial L(t, \cdot, \cdot)(\bar{x}(t), \dot{\bar{x}}(t))\}, \text{ a.e.};$$

$$(2) \quad p(t) = - \int_t^T \nu(d\tau) - \int_t^T q(\tau) d\tau;$$

hold along with the transversality condition

$$(3) \quad (p(0), -p(T) - \nu(\{T\})) \in \partial \ell(\bar{x}(0), \bar{x}(T)).$$

We need one more assumption to state the theorem containing necessary conditions for $W^{1,1}$ -minima.

- (H₆) There is a set valued mapping $Q(t)$ from $[0, T]$ into $\mathbb{R}^n$, whose graph belongs to $\mathcal{L} \otimes \mathcal{B}(\mathbb{R}^n)$, such that $\dot{\bar{x}}(t) \in Q(t)$ a.e. and for any $u \in Q(t)$ it is possible to find some $\varepsilon > 0$ (depending on u and t) such that $L(t, \cdot, u)$ is continuous on $B(\bar{x}(t), \varepsilon)$.

The last assumption implies in particular that $Q(t)$ is a subset of the domain of $L(t, \bar{x}(t), \cdot)$ for almost every t .

Theorem 2.2. *Assume in addition to (H₁)–(H₆) that there is an $\varepsilon_0 > 0$ such that $\varepsilon_0(t) \geq \varepsilon_0$ almost everywhere. Let $\bar{x}(\cdot)$ be a local minimum of J in the $W^{1,1}$ -topology. If the problem is not singular at $\bar{x}(\cdot)$, then there are a Radon measure ν , a measurable $q(t)$ and a function $p(t)$ of bounded variation such that the relations (1)–(3) hold along with the Weierstrass condition*

$$(4) \quad L(t, \bar{x}(t), u) - L(t, \bar{x}(t), \dot{\bar{x}}(t)) - \langle p(t), u - \dot{\bar{x}}(t) \rangle dt \geq 0, \quad \forall u \in Q(t), \quad \text{a.e.}$$

The relations (1) and (2) form what is usually called the *Euler–Lagrange adjoint inclusion*. In particular, if φ is absent in J (or equivalently, is identical zero), then $\nu = 0$, $p(\cdot)$ is absolutely continuous and the adjoint inclusion assumes the form

$$(5) \quad \dot{p}(t) \in \text{conv}\{q : (q, p(t)) \in \partial L(t, \cdot, \cdot)(\bar{x}(t), \dot{\bar{x}}(t))\}.$$

The inclusion reduces to the standard Euler equation if L satisfies the classical differentiability assumptions.

The next two theorems describe situations when a necessary condition can be stated in the Hamiltonian form.

Theorem 2.3. *Assume in addition to the assumptions of Theorem 2.1 that $L(t, x, \cdot)$ is convex for all t and x . Then the conclusion of Theorem 2.1 holds with (1) replaced by*

$$(6) \quad q(t) \in \text{conv}\{q : (-q, \dot{\bar{x}}(t)) \in \partial H(t, \cdot, \cdot)(\bar{x}(t), p(t))\}, \quad \text{a.e.}$$

Things are more complicated when $L(t, x, \cdot)$ is not convex. To state the corresponding result we shall assume that L is finite-valued, so that (H₄) is automatically satisfied. Let $\hat{L}$ be the convexification of L with respect to the last argument:

$$\hat{L}(t, x, y) = \inf \left\{ \sum_{i=1}^{n+1} \alpha_i L(t, x, y_i) : \alpha_i \geq 0, \sum \alpha_i = 1, \sum \alpha_i y_i = y \right\},$$

and let $\hat{L}_N(t, x, y)$ stand for the same infimum but under an additional condition $\|y_i - \dot{\bar{x}}(t)\| \leq N$, $i = 1, \dots, n+1$. Furthermore, given some $\delta > 0$ and $\gamma > 0$, set

$$N_{\delta, \gamma}(t) = \inf\{N : \hat{L}_N(t, x, y) = \hat{L}(t, x, y) \quad \text{if} \quad \|x - \bar{x}(t)\| < \delta, \|y - \dot{\bar{x}}(t)\| < \gamma\}.$$

Note that the definition does not exclude the possibility that $N_{\delta, \gamma}(t) = \infty$ (as $\inf \emptyset = \infty$ according to the standard convention). Let further

$$K_\delta(N, t) = \inf\{k : |L(t, x, y) - L(t, x', y)| \leq k\|x - x'\| \\ \text{if } x, x' \in B(\bar{x}(t), \delta), y \in B(\dot{\bar{x}}(t), N)\}.$$

Here again, we do allow N to be equal to $+\infty$. Clearly

$$K_\delta(N, t) = \sup\left\{\frac{|L(t, x, y) - L(t, x', y)|}{\|x - x'\|} : x, x' \in B(\bar{x}(t), \delta), \right. \\ \left. y \in B(\dot{\bar{x}}(t), N), x \neq x'\right\}.$$

Our additional assumption is the following:

(H₇) There are some $\bar{\delta} > 0$ and a positive-valued measurable function $\bar{\gamma}(t)$ such that the function $t \rightarrow K_{\bar{\delta}}(N_{\bar{\delta}, \bar{\gamma}(t)}(t), t)$ is summable.

We observe that (H₇) implies (H₆) with $Q(t) = B(\bar{x}(t), \bar{\gamma}(t))$ and (H₅) with $\varepsilon_0(t) \equiv \bar{\delta}$, $\gamma_0(t) = \bar{\gamma}(t)$ and $k_0(t) = K_{\bar{\delta}}(N_{\bar{\delta}, \bar{\gamma}(t)}(t), t)$.

Theorem 2.4. *Let (H₁)–(H₃) be satisfied along with (H₇). If under these assumptions, $\bar{x}(\cdot)$ is a strong local minimum of J and the problem is not singular at $\bar{x}(\cdot)$, then there are a measure $\nu \in \partial\varphi(\bar{x}(\cdot))$, a function $p(t)$ of bounded variation and a summable $q(t)$, all three $\mathbb{R}^n$ -valued and defined on $[0, T]$, such that the relations (2), (3) and (6) are satisfied with these ν , $p(\cdot)$ and $q(\cdot)$.*

The key result is Theorem 2.2. We shall see that the proof of the first theorem is actually a simplification of the proof of Theorem 2.2 that works under somewhat weaker assumptions, and the last two theorems are consequences of the two first. That is why we prove the second theorem first.

We refer to e.g. [2, 14, 12, 8, 13, 15, 16, 17, 20] for earlier results concerning the Euler-Lagrange inclusions for the generalized Bolza problem. Originally it was stated in a less precise form:

$$(7) \quad (\dot{p}(t), p(t)) \in \partial_c L(t, \cdot, \cdot)(\bar{x}(t), \dot{\bar{x}}(t)),$$

(for the case when φ is absent in the functional) in Clarke's 1976 paper [2] (and in some subsequent publications). In the form (1) it first appeared in [13, 17]. (Recall that J did not contain the φ -term in all publications prior to [12].)

Theorem 2.2 improves the main result of [12] in several respects in addition to what is connected with the presence of a non-Lipschitz off-integral term in the functional. First of all this concerns characterization of possible values of the third variable u for which the Weierstrass inequality may hold (which was perhaps the main contribution of [12] to our knowledge of necessary conditions for Bolza problems). Specifically, it was proved in [12] that the Weierstrass condition holds for elements of a set $D(t)$ such that for some $\varepsilon > 0$, the same for all t and $u \in D(t)$, the function $L(t, \cdot, u)$ is continuous on $B(\bar{x}(t), \varepsilon)$. Here we show that the same result can be proved under the assumption that ε may depend on t and u no matter how.

The results of other mentioned publications, sometimes even in a stronger form, follow from Theorems 2.1 and 2.2 one way or another. (The only exception is the main theorem of [17]. We shall explain below that the part of this theorem relating to the non-degenerate case follows from Theorem 2.1.) Verification is sufficiently simple for most of them. But in certain cases verification requires some effort. For instance, in [15] L (assumed continuous and not depending on t) is supposed to satisfy $\partial_c L(\bar{x}(t), \dot{\bar{x}}(t)) \subset R(t)B$ with some summable $R(t) \geq 0$ for almost every t . This means that for almost every t the function L is $(R(t) + 1)$ -Lipschitz in a neighborhood of $(\bar{x}(t), \dot{\bar{x}}(t))$. Thus all we need to apply e.g. Theorem 2.1 is to verify that there is a positive-valued measurable $\gamma(\cdot)$ such that for almost every t the corresponding neighborhood contains the ball of radius $\gamma(t)$ around $(\bar{x}(t), \dot{\bar{x}}(t))$. But under the assumption the latter does not need much work. It follows that under the assumptions of [15] the Euler-Lagrange adjoint inclusion in the more precise form (5) must also hold.

In certain earlier publications (namely [2, 17, 14]) the Lipschitz-type assumptions on L look, and actually are in certain respects, weaker than (H_5) . Indeed, in [2, 14] it is the mapping $x \rightarrow \text{epi } L(t, x, \cdot)$ that is assumed $k(t)$ -Lipschitz with a summable $k(\cdot)$ in a neighborhood of $\bar{x}(t)$ (along with a calmness assumption that guarantee a certain level of regularity of the problem), not the integrand itself. In [17] the Lipschitz-type assumption is even weaker: it is only assumed that the epigraphic mapping has the Aubin property near $(\bar{x}(t), \dot{\bar{x}}(t))$. But $L(t, x, \cdot)$ is assumed convex in [17].

In these cases reduction to problems considered here is much less straightforward. Let us start with [2, 14]. Consider the problem

$$\begin{aligned}
& \text{minimize} \quad \alpha(T) + \beta(T) + md((x(T), u(T)), C) \\
(P_0) \quad & \text{s.t.} \quad \alpha(0) = 0, \quad \beta(0) \geq \ell(x(0), u(0)); \\
& \quad (\dot{x}, \dot{\alpha}) \in \text{epi } L(t, x, \cdot), \quad \dot{\beta} = 0, \quad \dot{u} = 0,
\end{aligned}$$

where $C = \{(x, u) : x = u\}$. Set $\mathbf{X} = \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$. It is shown in [2] that under the calmness assumption $\bar{\mathbf{x}}(t) = (\bar{x}(\cdot), \bar{\alpha}(\cdot), \bar{\beta}(\cdot), \bar{u}(\cdot))$, where $\bar{\alpha}(t) = L(t, \bar{x}(t), \dot{\bar{x}}(t))$, $\bar{\beta}(t) \equiv 0$ and $\bar{u}(t) \equiv 0$, solves the problem if m is sufficiently large.

Set $E(t, x) = \text{epi } L(t, x, \cdot) \times \{0\} \times \{0\}$. Under the assumptions of [2, 14], this set-valued mapping is $k(t)$ -Lipschitz as a function of x for almost every t . Let $M(\mathbf{x})$ be the set of solutions of the differential inclusion $\dot{\mathbf{x}} \in E(t, x)$ with $\mathbf{x}(0) = \mathbf{x}$. As follows from [3, Theorem 3.1.6] there is a $K > 0$ such that for all $\mathbf{x}(t) = (x(t), \alpha(t), \beta, u)$ sufficiently close to $\bar{\mathbf{x}}(t)$ in the $W^{1,1}$ -topology

$$(8) \quad d(\mathbf{x}(\cdot), M(\mathbf{x}(0))) \leq K \int_0^T d(\dot{\mathbf{x}}(t), E(t, x(t))) dt.$$

Using this inequality and the fact that there is no constraint on $\mathbf{x}(T)$ in (P_0) , we can show (see e.g. the proof of Theorem 3.3 in [12]), again if K_0 is sufficiently large, that $\bar{\mathbf{x}}(\cdot)$ is a local minimum of

$$\begin{aligned}
& \alpha(T) + \beta(T) + m\|x(T) - w(T)\| + K_0 \left(|\alpha(0)| + d((x(0), w(0)), \beta(0)), \text{epi } \ell \right) \\
& \quad + \int_0^T (d((\dot{x}(t), \dot{\alpha}(t)), \text{epi } L(t, x(t), \cdot)) + \|\dot{w}\| + |\dot{\beta}|) dt \Big).
\end{aligned}$$

We can simplify the functional, since we know the minimum, and replace it by

$$\begin{aligned}
& \alpha(T) + \beta + K_0 \left(|\alpha(0)| + d((x(0), x(T), \beta), \text{epi } \ell) \right. \\
& \quad \left. + \int_0^T d((\dot{x}(t), \dot{\alpha}(t)), \text{epi } L(t, x(t), \cdot)) dt \right)
\end{aligned}$$

with $\beta = \text{const}$ (and $w(t) \equiv x(T)$). This functional satisfies all assumptions of Theorem 2.2 and application of the theorem to the functional gives all relations claimed in [2, 14] plus the Weierstrass condition.

The situation is somewhat different with [17]. The absence of the calmness condition does not allow to use (P_0) any more. Still, the problem of minimization

of J can be equivalently rewritten as a slightly different (and looking even simpler) optimal control problem:

$$(P_1) \quad \begin{array}{ll} \text{minimize} & \alpha(T) + \beta \\ \text{s.t} & \alpha(0) = 0, \beta \geq \ell(x(0), x(T)); \quad (\dot{\alpha}(t), \dot{x}(t)) \in \text{epi } L(t, x(t), \cdot). \end{array}$$

If we set $\mathbf{x} = (x, \alpha)$ and $E(t, x) = \text{epi } L(t, x, \cdot)$ we still can get (8) with $M(\mathbf{x})$ replaced by $\{\mathbf{x} : \alpha(0) = 0, \alpha(T) \geq \ell(x(0), x(T))\}$ thanks to the fact that E has the Aubin property near $((\dot{x}(t), \dot{\alpha}(t)), \bar{x}(t))$ [10]. But passage to an unconstrained Bolza problem is no longer possible unless we assume that the constraints in (P_1) are regular in the sense of variational analysis (see e.g. the proof of Theorem 3.3 in [12]). The corresponding Bolza problem is the same as above. Thus the “regular” part of the main result of [17] follows from Theorem 2.2.

3. Preliminaries.

3.1. Subdifferentials. We shall use several types of subdifferentials in statements and proofs, mainly for functions in $\mathbb{R}^n$. The first is the *proximal subdifferential* ∂_p making sense for functions on a Hilbert space X : if f is a function on X finite at x then $y \in \partial_p f(x)$ if there is a $c > 0$ such that

$$f(x+h) - f(x) \geq \langle y, h \rangle - c\|h\|^2$$

for all h of a neighborhood of zero.

Proposition 3.1 ([13], Theorem 2). *Let X be a Hilbert space, and let $f_1, \dots, f_k$ be lower semicontinuous functions on X with values in $(-\infty, \infty]$ finite at $\bar{x}$. Assume that the following uniform lower semicontinuity property holds:*

(ULC) *there is a $\delta > 0$ such that for any k sequences (x_{im}) , $m = 1, 2, \dots$, $i = 1, \dots, k$ in the δ -ball around $\bar{x}$ satisfying $\|x_{im} - x_{jm}\| \rightarrow 0$ as $m \rightarrow \infty$ for any $i, j = 1, \dots, k$ we can find u_m , $m = 1, 2, \dots$ such that $\|x_{im} - u_m\| \rightarrow 0$ and*

$$\liminf_{m \rightarrow \infty} \sum_{i=1}^k (f_i(x_{im}) - f_i(u_m)) \geq 0$$

Set $f(x) = \sum_i f_i(x)$. Then for any $x^* \in \partial_p f(\bar{x})$ and any $\varepsilon > 0$ there are u_i and $u_i^* \in \partial_p f_i(u_i)$, $i = 1, \dots, k$, such that

$$|f_i(x) - f_i(u_i)| < \varepsilon, \quad \|x - u_i\| < \varepsilon, \quad \left\| \sum u_i^* - x^* \right\| < \varepsilon.$$

The condition of the proposition is trivially satisfied if all functions but for at most one are Lipschitz in a neighborhood of $\bar{x}$. Indeed, if say, $f_2, \dots, f_k$ are Lipschitz near $\bar{x}$, we can take $u_m = x_{1m}$. It follows that, verifying (ULC) we can exclude Lipschitz functions from consideration.

Proposition 3.2 ([6, Chapter 3, Theorem 5.10]). *Let $\varphi(t, x)$ be a $\mathcal{L} \otimes \mathcal{B}(\mathbb{R}^n)$ -measurable function on $[0, T] \times \mathbb{R}^n$ which is lower semicontinuous in x . Consider on $L^2(0, T)$ the functional*

$$f(x(\cdot)) = \int_0^T \varphi(t, x(t)) dt$$

and assume that $\xi(\cdot) \in \partial_p f(x(\cdot))$. Then $\xi(t) \in \partial_p \varphi(t, \cdot)(x(t))$ almost everywhere.

Note that it is assumed in [6] that φ is globally Lipschitz as a function of x . But the assumption actually plays little role in the proof. Namely, it is used (in the proof of Proposition 5.5 in [6]) to extract measurable selection of set-valued mappings $t \rightarrow \{x : \alpha \leq \varphi(t, x) \leq b\}$ which is still possible under our assumptions.

Let f be a lsc function on $\mathbb{R}^n$ which is finite at x . The set

$$\partial f(x) = \limsup_{u \rightarrow x} \partial_p f(u).$$

is the *limiting subdifferential* of f at x .

Proposition 3.3 ([13, Lemma 3]). *Let $\varphi(x, y)$, for $x, y \in \mathbb{R}^n$, be l.s.c., and let $(w_m, v_m) \in \partial_p \varphi(x_m, y_m)$, where $x_m \rightarrow \bar{x}$, $y_m \rightarrow \bar{y}$, $v_m \rightarrow \bar{v}$, $\varphi(x_m, y_m) \rightarrow \varphi(\bar{x}, \bar{y})$ and w_m are uniformly bounded. Let u_m be convex combinations of w_k for $k \geq m$ converging to a certain u . Then $u \in \text{conv}\{w : (w, \bar{v}) \in \partial \varphi(\bar{x}, \bar{y})\}$.*

It is possible to extend the concept of limiting subdifferential to the so called *G-subdifferential* for functions on arbitrary Banach spaces. Below we describe the definition of *G-subdifferential* only for Lipschitz functions on separable Banach spaces. (The general definition is more complicated and we do not need it here. See [11] for details.) First we define the *Dini-Hadamard subdifferential* ∂^- :

$$\partial^- f(x) = \{x^* \in X^* : \langle x^*, h \rangle \leq d^- f(x; h), \forall h \in X\},$$

where

$$d^- f(x; h) = \liminf_{\lambda \rightarrow +0} \lambda^{-1} (f(x + \lambda h) - f(x)).$$

The G -subdifferential is now defined as the limiting subdifferential but with ∂^- instead of ∂_p :

$$\partial_G f(x) = \limsup_{u \rightarrow x} \partial^- f(u).$$

For functions on $\mathbb{R}^n$ a slightly modified definition applies for all lsc functions and the result coincides with the limiting subdifferential. (The modification concerns the definition of $d^- f(x; h)$. Namely, for a non-Lipschitz function on $\mathbb{R}^n$ this is the lower limit of $\lambda^{-1}(f(x + \lambda u) - f(x))$ when $\lambda \rightarrow +0$ and $u \rightarrow h$.) The proposition below gives an estimate for the limiting subdifferentials of some specific class of marginal functions.

Proposition 3.4. *Given a lower semicontinuous function φ on $\mathbb{R}^n \times \mathbb{R}^m$ and a convex continuous function ψ on $\mathbb{R}^m$. Set*

$$f(x, y) = \inf_v (\varphi(x, y + v) + \psi(v)).$$

and denote by $S(x, y)$ the set of v at which the infimum is attained for given (x, y) . Assume that $S(x, y) \neq \emptyset$ for every (x, y) in a neighborhood of a certain $(\bar{x}, \bar{y})$. Let $\bar{v} \in S(\bar{x}, \bar{y})$ and the following property holds: if $(x_k, y_k) \rightarrow (\bar{x}, \bar{y})$, $v_k \in S(x_k, y_k)$ and $f(x_k, y_k) \rightarrow f(\bar{x}, \bar{y})$, then $v_k \rightarrow \bar{v}$.

Let finally $(\xi, \eta) \in \partial f(\bar{x}, \bar{y})$. Then

$$(\xi, \eta) \in \partial \varphi(\bar{x}, \bar{y} + \bar{v}), \quad \text{and} \quad -\eta \in \partial \psi(\bar{v}).$$

Proof. Let $(x_k, y_k, \xi_k, \eta_k) \rightarrow (\bar{x}, \bar{y}, \xi, \eta)$ with $f(x_k, y_k) \rightarrow f(\bar{x}, \bar{y})$ and $(\xi_k, \eta_k) \in \partial^- f(x_k, y_k)$. This means that

$$d^- f(x_k, y_k)(h, e) = \liminf_{\lambda \rightarrow +0} \lambda^{-1} (f(x_k + \lambda h, y_k + \lambda e) - f(x_k, y_k)) \geq \langle \xi_k, h \rangle + \langle \eta_k, e \rangle$$

for all $(h, e) \in \mathbb{R}^n \times \mathbb{R}^m$. Take a $v_k \in S(x_k, y_k)$. Then for any v

$$\begin{aligned} \liminf_{\lambda \rightarrow +0} \lambda^{-1} (\varphi(x_k + \lambda h, y_k + v_k + \lambda(e + v)) + \psi(v_k + \lambda v) - \varphi(x_k, y_k + v_k) - \psi(v_k)) \\ \geq \langle \xi_k, h \rangle + \langle \eta_k, e \rangle = \langle \xi_k, h \rangle + \langle \eta_k, e + v \rangle - \langle \eta_k, v \rangle. \end{aligned}$$

Taking into account that directional derivatives of ψ at any point along any direction do exist (as ψ is convex continuous), we get

$$d^- \varphi(x_k, y_k + v_k)(h, e + v) + \psi'(v_k, v) \geq \langle \xi_k, h \rangle + \langle \eta_k, e + v \rangle - \langle \eta_k, v \rangle$$

for all (h, e, v) . Taking consecutively $v = 0$ and $h = 0$, $e + v = 0$, we conclude that

$$(\xi_k, \eta_k) \in \partial^- \varphi(x_k, y_k + v_k) \quad \text{and} \quad -\eta_k \in \partial \psi(v_k)$$

and the result follows. $\square$

Proposition 3.5 ([9, Theorem 4]). *Let $f(x, y)$ be a lower semicontinuous function on $\mathbb{R}^n \times \mathbb{R}^m$ which is convex in y , finite at $(\bar{x}, \bar{y})$ and Aubin continuous¹ in x near $(\bar{x}, \bar{y})$. Let $f_2^*(x, p)$ stand for the Fenchel conjugate of $f(x, \cdot)$. Let finally, $(\bar{u}, \bar{p}) \in \partial f(\bar{x}, \bar{y})$. Then $-\bar{u} \in \text{conv}\{w : (w, \bar{y}) \in \partial f_2^*(\bar{x}, \bar{p})\}$.*

Proposition 3.6 ([11, Corollary 7.15]). *Let X be a Banach space, and let the functions f_i , $i = 1, \dots, k$ be Lipschitz continuous in a neighborhood of a certain $\bar{x}$. Set $f(x) = \max_i f_i(x)$. and and let $I = \{i : f_i(\bar{x}) = f(\bar{x})\}$. Then $\partial_G f(\bar{x}) \subset \bigcup \sum \alpha_i \partial_G f_i(\bar{x})$, where the union is taken over all tuples $(\alpha_1, \dots, \alpha_k)$ with $\alpha_i \geq 0$, $\sum \alpha_i = 0$ and $\alpha_i = 0$ for $i \notin I$.*

Proposition 3.7 ([11, Proposition 4.60]). *Let X and Y be Banach spaces and $A : X \rightarrow Y$ a linear bounded operator with $\text{Im } A = Y$. Let further f be a function on Y which is Lipschitz continuous in a neighborhood of a certain $\bar{y} = A\bar{x}$. Set $g = f \circ A$. Then*

$$\partial_G g(\bar{x}) = A^*(\partial_G f(\bar{y})).$$

Finally, Clarke's generalized gradient $\partial_c f(x)$ is the convex hull of $\partial f(x)$ if f is a Lipschitz function on $\mathbb{R}^n$. Again, we shall not need ∂_c for more general classes of functions and spaces.

3.2. Relaxation. Consider the functional of calculus of variation $\int_0^T L(t, x(t), \dot{x}(t)) dt$ ($x \in \mathbb{R}^n$) assuming that $L(t, x(t), u(t))$ is measurable when $x(\cdot)$ is continuous and $u(\cdot)$ is measurable. The following is a relaxation result suitable for further applications in this paper.

Proposition 3.8 (relaxation theorem). *Let $x(\cdot) \in W^{1,1}$ and summable $u_i(\cdot)$, $i = 1, \dots, k$ satisfy*

$$\dot{x}(t) = \sum_{i=1}^k \alpha_i(t) u_i(t), \quad \text{a.e.}$$

with some measurable $\alpha_i(t) \geq 0$ such that $\sum \alpha_i(t) = 1$ a.e.. Assume that all functions $L(t, x(t), u_i(t))$ are summable and there are $\varepsilon > 0$ and a summable real-valued $c(t)$ such that for every i and almost every t the function $x \rightarrow L(t, x, u_i(t))$

¹A set-valued mapping $x \rightrightarrows Q(x)$ has the *Aubin property* near $(x, y) \in \text{Graph } Q$ if there are $\varepsilon > 0$ and $K > 0$ such that $Q(x') \cap B(y, \varepsilon) \subset Q(x) + K\|x - x'\|B$ for all x', x'' of a neighborhood of x . A function $f(x, y)$ is *Aubin continuous* in x near (x, y) if the set-valued mapping $x \rightrightarrows \text{epi } f(x, \cdot)$ has the Aubin property near $(x, (y, f(x, y)))$.

is continuous on the ε -neighborhood of $x(t)$ and $L(t, x, u_i(t)) \leq c(t)$ if $\|x - x(t)\| \leq \varepsilon$.

Then there are $x_m(\cdot) \in W^{1,1}$, $m = 1, 2, \dots$ with $\dot{x}_m(t) \in \{u_1(t), \dots, u_k(t)\}$ uniformly converging to $x(\cdot)$ and such that

$$\limsup_{m \rightarrow \infty} \int_0^T L(t, x_m(t), \dot{x}_m(t)) dt \leq \int_0^T \left(\sum_{i=1}^k \alpha_i(t) L(t, x(t), u_i(t)) \right) dt.$$

Moreover, if in addition there is a set $\Delta \subset [0, T]$ of positive measure and a summable function $c_\Delta(t)$ on Δ such that $\sup\{L(t, x, y) : \|x - x(t)\| \leq \varepsilon, \|y - \dot{x}(t)\| \leq \varepsilon\} \leq c_\Delta(t)$ for almost every $t \in \Delta$, then $x_m(\cdot)$ can be chosen to satisfy $x_m(0) = x(0)$, $x_m(T) = x(T)$.

Proof. With no loss of generality we may assume that $c(t) = c_\Delta(t)$ for $t \in \Delta$. Fix an m . Let us break $[0, T]$ into m segments $\Delta_j = [T(j-1)/m, Tj/m]$, $j = 1, \dots, m$ of length T/m . We can further break every Δ_j into k segments Δ_{ji} , $i = 1, \dots, k$ such that the length of Δ_{ji} is $\int_{\Delta_{ji}} \alpha_i(t) dt$.

Now let

$$\alpha_{im}(t) = \begin{cases} 1, & \text{if } t \in \cup_j \Delta_{ji} \\ 0, & \text{otherwise.} \end{cases}$$

and define $x_m(\cdot)$ as follows:

$$x_m(0) = x(0); \quad \dot{x}_m(t) = \sum_{i=1}^k \alpha_{im}(t) u_i(t).$$

It is clear that any $\alpha_{im}(\cdot)$ weak* converge (in L^∞) to $\alpha_i(\cdot)$ when $m \rightarrow \infty$. It follows that $x_m(\cdot)$ uniformly converge to $x(\cdot)$.

We notice next that

$$\int_0^T L(t, x_m(t), \dot{x}_m(t)) dt = \sum_{i=1}^k \int_0^T \alpha_{im}(t) L(t, x_m(t), u_i(t)) dt,$$

so we only need to check that for any i

$$(9) \quad \limsup_{m \rightarrow \infty} \int_0^T \alpha_{im}(t) L(t, x_m(t), u_i(t)) dt \leq \int_0^T \alpha_i(t) L(t, x(t), u_i(t)) dt.$$

Indeed, set $\varphi_r(t) = \sup_{m \geq r} L(t, x_m(t), u_i(t))$. Then $L(t, x(t), u_i(t)) \leq \varphi_r(t) \leq c(t)$ a.e., that is $\varphi_r(\cdot) \in L^1$ and therefore

$$\limsup_{m \rightarrow \infty} \int_0^T \alpha_{im}(t) L(t, x_m(t), u_i(t)) dt \leq \limsup_{m \rightarrow \infty} \int_0^T \alpha_{im}(t) \varphi_r(t) dt = \int_0^T \alpha_i(t) \varphi_r(t) dt.$$

But $\varphi_r(t) \searrow L(t, x(t), u_i(t))$ as $r \rightarrow \infty$ and (9) follows. This completes the proof of the first statement.

To prove the second, choose subsets $\Delta_m \subset \Delta$ whose measures go to zero as $m \rightarrow \infty$. Let now

$$v_{im} = \begin{cases} u_i(t), & \text{if } t \notin \Delta_m, \\ \dot{x}(t), & \text{if } t \in \Delta_m \end{cases}$$

Clearly, $\dot{x}(t) = \sum \alpha_i(t) v_{im}(t)$ almost everywhere, so applying the result of the first statement we shall be able to find for any fixed m a sequence of $x_{lm}(\cdot)$, $l = 1, 2, \dots$ uniformly converging to $x(\cdot)$ and such that

$$\limsup_{l \rightarrow \infty} \int_0^T L(t, x_{lm}(t), \dot{x}_{lm}(t)) dt \leq \int_0^T \left(\sum_{i=1}^k \alpha_i(t) L(t, x(t), u_i(t)) \right) dt.$$

Namely we set $x_{lm}(t) = x(0) + \int_0^t \sum \alpha_{ilm}(s) v_{im}(s) ds$ with some $\alpha_{ilm}(t)$ taking values in $\{0, 1\}$, satisfying $\sum_i \alpha_{ilm}(t) = 1$ and weak* converging to $\alpha_i(\cdot)$ as $l \rightarrow \infty$.

Set $\gamma_{lm} = x_{lm}(T) - x(T)$. Clearly, $\gamma_{lm} \rightarrow 0$ as $l \rightarrow \infty$. If for a given m there is a $l = l(m)$ such that $\|x_{lm}(t) - x(t)\| \leq 1/m$ for all t and $x_{lm}(T) = x(T)$, we set $x_m(\cdot) = x_{lm}(\cdot)$. Otherwise, we modify v_{im} for some l by replacing $\dot{x}(t)$ for $t \in \Delta_m$ by $\dot{x}(t) - \gamma_{lm}/\mu(\Delta_m)$ (where $\mu(\Delta_m)$ is the Lebesgue measure of Δ_m). Take $l = l(m)$ so large that $\|x_{lm}(t)\| + \gamma_{lm} < \varepsilon$ for all t . Then the integral of $L(t, x_{lm}(t), \dot{x}_{lm}(t))$ may change after the modification at most by $2 \int_{\Delta_m} c(t) dt$ which cannot affect the limit relation as $c(\cdot)$ is summable and $\mu(\Delta_m) \rightarrow 0$. $\square$

Remark 3.9. Since $\dot{x}_m(t)$ assumes its values in $\{\dot{x}(t), u_1(t), \dots, u_k(t)\}$ and all these functions are summable, we can be sure that uniform convergence of $x_m(\cdot)$ to $x(\cdot)$ implies weak convergence in $W^{1,1}$.

4. Proof of Theorem 2.2. To begin with, we can assume in the proof that $\bar{x}(t) \equiv 0$. The proof follows the same lines as the proof of Theorem 3.1 in [12] with necessary modifications.

4.1. Reduction of the problem. Suppose we are given a finite number of measurable selections $u_1(\cdot), \dots, u_k(\cdot)$ of $Q(\cdot)$ such that for any i there are an $\varepsilon_i > 0$ and a summable $c_i(t)$ such that the function $x \rightarrow L(t, \cdot, u_i(t))$ is continuous and satisfies $L(t, x, u_i(t)) \leq c_i(t)$ on $B(\bar{x}(t), \varepsilon_i)$. Let now $\delta > 0$ be so small that

$x(\cdot)$ defined by

$$(10) \quad x(0) = 0, \quad \dot{x}(t) = y(t) + \sum_{i=1}^k \alpha_i (u_i(t) - y(t))$$

satisfies $\|x(t)\| < \varepsilon_0$ for all t if $\|y(t)\| \leq \min\{\delta, \gamma_0(t)\}$ a.e. and $0 \leq \alpha_i \leq \delta$.

Let further $L_\delta(t, x, y)$ be equal to $L(t, x, y)$ if $\|x\| < \delta$ and $\|y\| \leq \min\{\delta, \gamma_0(t)\}$ and to $+\infty$ otherwise. Consider the problem of minimizing

$$J_k(x(\cdot), \alpha_1, \dots, \alpha_k) = \ell(x(0), x(T)) + \varphi(x(\cdot)) \\ + \int_0^T (L_\delta(t, x(t), y(t)) + \sum_{i=1}^k \alpha_i (L(t, x(t), u_i(t)) - L(t, x(t), y(t)))) dt,$$

subject to (10) and $0 \leq \alpha_i \leq \delta$.

Our purpose is to show that zero is a local minimum of the problem in $W^{1,1} \times \mathbb{R}^k$. So take some $x(\cdot)$ and $y(\cdot)$ such that (10) holds with some nonnegative $\alpha_1, \dots, \alpha_k$ with $\sum \alpha_i < \delta$ and $L_\delta(t, x(t), y(t))$ is a summable function, that is $y(t) \in A(t)$ a.e. (Otherwise $J_k(x(\cdot), \alpha_1, \dots, \alpha_k) = \infty$.) Then there is a summable $c_0(t)$ such that $L(t, x, y(t)) \leq c_0(t)$ for all $x \in B(\bar{x}(t), \delta)$ almost everywhere on $[0, T]$ and we can apply the relaxation theorem (Proposition 3.8) with $u_0(\cdot) = y(\cdot)$ and $\alpha_i(t) \equiv \alpha_i$.

It follows that (9) holds for a sequence of $x_m(\cdot)$ weakly (in $W^{1,1}$) converging to $x(\cdot)$ and coinciding with $x(\cdot)$ at the ends of the interval. Then

$$\liminf_{m \rightarrow \infty} J(x_m(\cdot)) \leq J_k(x(\cdot), \alpha_1, \dots, \alpha_k)$$

and therefore, if δ is sufficiently small, we can be sure that $x_m(\cdot)$ are also close enough to $\bar{x}(\cdot)$ to guarantee that $J(x_m) \geq J(\bar{x}(\cdot))$ and, consequently, $(\bar{x}(\cdot), 0, \dots, 0)$ is a local minimum of J_k in $W^{1,1} \times \mathbb{R}^k$.

So take such a $\delta > 0$ and set

$$\Lambda(a, y(\cdot), \alpha_1, \dots, \alpha_k)(t) = a + \int_0^t \left(y(\tau) + \sum_{i=1}^k \alpha_i^+ (u_i(\tau) - y(\tau)) \right) d\tau; \\ g_{i\delta}(t) = \sup\{(L(t, x, u_i(t)) - L(t, x, y)) : \|x\| \leq \delta, \|y\| \leq \min\{\delta, \gamma_0(t)\}\}.$$

Consider the following five functionals of $\mathbf{x} = (x(\cdot), a, y(\cdot), \alpha_1, \dots, \alpha_k)$ on

$$\mathbf{X} = W^{1,2} \times \mathbb{R}^n \times L^2 \times \mathbb{R} \times \cdots \times \mathbb{R};$$

$$I_1(\mathbf{x}) = \ell(a, \Lambda(a, y(\cdot), \alpha_1, \dots, \alpha_k)(T));$$

$$I_2(\mathbf{x}) = \varphi(\Lambda(a, y(\cdot), \alpha_1, \dots, \alpha_k)(\cdot));$$

$$I_3(\mathbf{x}) = \int_0^T L_\delta(t, x(t), y(t)) dt;$$

$$I_4(\mathbf{x}) = \int_0^T k_0(t) |x(t) - \Lambda(a, y(\cdot), \alpha_1, \dots, \alpha_l)(t)| dt;$$

$$I_5(\mathbf{x}) = \sum_i^k \alpha_i \int_0^T g_{i\delta}(t) dt + i_k^+(\alpha_1, \dots, \alpha_k),$$

where i_k^+ stands for the indicator of the positive orthant in $\mathbb{R}^k$, that is the function equal to zero if all $\alpha_i \geq 0$ and $+\infty$ otherwise.

Clearly, $(a, y(\cdot)) \rightarrow \Lambda(a, y(\cdot), \alpha_1, \dots, \alpha_k)$ is a bounded linear operator from $\mathbb{R}^n \times L^2$ into the space of continuous functions, hence I_2 and I_4 are Lipschitz functions, as well as I_1 if ℓ is Lipschitz. And the functions $g_{i\delta}(\cdot)$ are summable as $L(t, \bar{x}(t), u_i(t)) - L(t, \bar{x}(t), \dot{\bar{x}}(t)) \leq g_{i\delta}(t) \leq c_i(t) + c_0(t)$.

Set $I = I_1 + I_2 + I_3 + I_4 + I_5$. It is clear that zero is a local minimum of I on $\mathbf{X}$. It is also an easy matter to see that zero remains a local minimum of the functional if we consider $\mathbf{X}$ with the $L^2 \times \mathbb{R}^n \times L^2 \times \mathbb{R} \times \cdots \times \mathbb{R}$ -topology. (that is if we consider $x(\cdot)$ in L^2 rather than in $W^{1,2}$). Thus $0 \in \partial_p I(0)$.

4.2. Analysis. 1. We have to verify that I_j satisfy the uniform lower semicontinuity property of Proposition 3.1 to be able to apply the proposition to the inclusion. We can be sure that the chosen $\delta > 0$ is smaller than all ε_i . Consider first the case when ℓ is not Lipschitz near $(\bar{x}(0), \bar{x}(T))$ and $A(t)$ coincideds with the $\gamma_0(t)$ -ball around zero for any t .

As I_2 and I_4 satisfy the Lipschitz condition, we only have to verify ULC for $I_1 + I_3 + I_5$. So let $\mathbf{x}_{jm} = (x_{jm}(\cdot), a_{jm}, y_{jm}(\cdot), \alpha_{1,jm}, \dots, \alpha_{k,jm})$, $j = 1, 3, 5$; $m = 1, 2, \dots$ be such that $\|\mathbf{x}_{jm}\| < \delta/2$ and $\|\mathbf{x}_{1m} - \mathbf{x}_{jm}\| \rightarrow 0$ for $j = 3, 5$ as $m \rightarrow \infty$.

The integrals $\int_0^T \|y_{1m}(t) - y_{3m}(t)\| dt$ go to zero as $m \rightarrow \infty$.

As by the assumptions $A(t) = \gamma_0(t)B$, we can choose some $\Delta_m \subset [0, T]$ with $\mu(\Delta_m) = \tau_m \rightarrow 0$ and functions $v_m(\cdot)$ on Δ_m with $\|v_m(t)\| \leq \gamma_0(t)$ such

that

$$\int_{\Delta_m} v_m(t)dt + \int_{T \setminus \Delta_m} y_{3m}(t)dt = \int_0^T y_{1m}(t)dt.$$

Let now $w_m(t)$ be equal to $v_m(t)$ on Δ_m and $y_{3m}(t)$ on $T \setminus \Delta_m$. Set finally $\mathbf{u}_m = (x_{3m}(\cdot), a_{1m}, w_m(\cdot), \alpha_{1,1m}, \dots, \alpha_{k,1m})$. Then $I_1(\mathbf{x}_{1m}) = I_1(\mathbf{u}_m)$,

$$I_3(\mathbf{x}_{3m}) - I_3(\mathbf{u}_m) = \int_{\Delta_m} (L(t, x_{3m}(t), y_{3m}(t)) - L(t, x_{3m}(t), v_m(t)))dt \rightarrow 0$$

(as immediately follows from (H₅)) and

$$|I_5(\mathbf{x}_{5m}) - I_5(\mathbf{u}_m)| \leq \sum_{i=0}^k |\alpha_{1m} - \alpha_{5m}| \int_0^T |g_{i\delta}(t)|dt \rightarrow 0.$$

If ℓ is Lipschitz near zero, then $I_1(\mathbf{x}_{1m}) - I_1(\mathbf{x}_{3m}) \rightarrow 0$. In this case we can set $\mathbf{u}_m(\cdot) = \mathbf{x}_{3m}$.

2. It follows from Proposition 3.1 that we can find five sequences $(\mathbf{x}_{jm}) \subset \mathbf{X}$, $j = 1, 2, 3, 4, 5$, $\mathbf{x}_{jm} = (x_{jm}(\cdot), a_{jm}, y_{jm}(\cdot), \alpha_{1,jm}, \dots, \alpha_{k,jm})$, converging to zero as $m \rightarrow \infty$ and $\mathbf{x}_{jm}^* \in \partial_p I_j(\mathbf{x}_{jm})$ such that $I_j(\mathbf{x}_{jm}) \rightarrow I_j(0)$ and $\|\sum_j \mathbf{x}_{jm}^*\| \leq$

m^{-1} . (As I_2 does not depend on $x(\cdot)$ we may assume that $x_2(\cdot) = \Lambda(a_{2m}, y_{2m}(\cdot), \alpha_{1,2m}, \dots, \alpha_{k,2m})$.) In other words, (cf. Propositions 3.2 and 3.7), there are some $(b_{0m}, b_{Tm}) \in \partial_p \ell(x_{1m}(0), x_{1m}(T))$, $\nu_m \in \partial_G \varphi(x_{2m}(\cdot))$, some measurable $\lambda_m(t)$ and $\mu_m(t)$ such that $(\lambda_m(t), \mu_m(t)) \in \partial_p L_\delta(t, \cdot, \cdot)(x_{3m}(t), y_{3m}(t))$, measurable $\mathbb{R}^n$ -valued $\xi_m(\cdot)$ with $\|\xi_m(t)\| \leq k_0(t)$ for almost all t , and $\rho_{im} \leq 0$, $i = 1, \dots, l$ such that the norms of the sums of multipliers corresponding respectively to $x(\cdot)$, a , $y(\cdot)$ and α_i go to zero, that is,

$$\begin{aligned} & \int_0^T |\lambda_m(t) + \xi_m(t)|^2 dt \rightarrow 0; \\ & b_{0m} + b_{Tm} + \int_0^T (\nu_m(dt) - \xi_m(t)dt) \rightarrow 0; \\ (11) \quad & \int_0^T \left| b_{Tm} + \int_t^T \nu_m(d\tau) + \mu_m(t) - \int_t^T \xi_m(\tau)d\tau \right|^2 dt \rightarrow 0; \\ & \int_0^T \left(\left\langle b_{Tm} + \int_t^T \nu_m(d\tau), u_i(t) \right\rangle \right. \\ & \quad \left. - \left\langle \int_t^T \xi_m(\tau)d\tau, u_i(t) \right\rangle + g_{i\delta}(t) \right) dt + \rho_{im} \rightarrow 0, \end{aligned}$$

when $m \rightarrow \infty$. (Recall that proximal subdifferential is smaller than the G -subdifferential, so we may use the G -subdifferential of φ which is more convenient technically.)

We note next that the sequences (b_{0m}) , (b_{Tm}) and $(\mu_m(\cdot))$ may be unbounded if the corresponding functions are not Lipschitz. On the contrary, the sequences (ν_m) , $(\lambda_m(\cdot))$ and $(\xi_m(\cdot))$ are always bounded in the corresponding spaces under the assumptions.

3. If the sequence (b_{0m}) is unbounded, then taking, if necessary, a subsequence, we get $\|b_{0m}\| \rightarrow \infty$. Let us divide each of the four relations in (11) by $\|b_{0m}\|$ and set

$$\tilde{b}_{0m} = \frac{b_{0m}}{\|b_{0m}\|}; \quad \tilde{b}_{Tm} = \frac{b_{Tm}}{\|b_{0m}\|}; \quad \tilde{\mu}_m(t) = \frac{\mu_m(t)}{\|b_{0m}\|}; \quad \tilde{\rho}_{im} = \frac{\rho_{im}}{\|b_{0m}\|}.$$

We may assume that $\tilde{b}_{0m}$ converge to some $b \neq 0$. By the second relation in (11) the sequence (b_{Tm}) is also unbounded and $\tilde{b}_{Tm}$ converge to $-b$, and by the fourth relation $\langle b, \int u_i(t)dt \rangle \geq 0$. Finally, the third relation implies that $\tilde{\mu}(t)$ converge to b almost everywhere.

As $(b_{0m}, b_{Tm}) \in \partial_p \ell(x_{1m}(0), x_{1m}(T))$, we conclude that $(b, -b) \in \partial \ell(0, 0)$. Likewise, as $(\lambda_m(t), \mu_m(t)) \in \partial_p L_\delta(t, \cdot, \cdot)(x_{3m}(t), y_{3m}(t))$ and $\|\lambda_m(t)\| \leq k(t)$ if m is sufficiently large, we have to conclude that the vector $(0, 0, b) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ belongs to the normal cone to the graph of $L(t, \cdot, \cdot)$ at $(L(t, 0, 0), 0, 0)$. Thus, the assumption that the sequence (b_{0m}) is unbounded leads us to the conclusion that (the sequences (b_{Tm}) and $(\mu_m(\cdot))$ are also unbounded and) we have what we called the *singular case*.

4. Assume now that the sequence (b_{0m}) , and hence the sequences (b_{Tm}) and $(\mu_m(\cdot))$, is bounded. Then we can assume, taking if necessary a subsequence, that (b_{0m}, b_{Tm}) converge to some $(b_0, b_T) \in \partial \ell(0, 0)$.

Set $\theta_m(t) = b_{Tm} + \int_t^T \nu_m(d\tau)$ and $\eta_m(t) = \int_t^T \xi_m(\tau)d\tau$, and let $p_m(t) = \eta_m(t) - \theta_m(t)$. Now the second and the third relations in (11) imply that $p_m(0) - b_{0m} \rightarrow 0$, while the first and the third relations imply that the functions under the signs of integral go to zero almost everywhere (at least along some subsequence). Furthermore, as total variations of ν_m are uniformly bounded, we may assume that ν_m weak* converges to some ν such that $\nu \in \partial_G \varphi(0)$. The latter means that $\int_t^T \nu_m(d\tau)$ converge to $\int_t^T \nu(d\tau)$ at every point of continuity of the latter, that is everywhere except maybe countably many points. As all $\xi_m(\cdot)$ are bounded by the same summable function $k_0(t)$, this sequence is relatively compact in the

weak topology of L^1 . By the Eberlein–Smulian theorem, we can assume that the sequence $(\xi_m(\cdot))$ weak converges to a certain $\xi(t)$ and consequently (by the first relation in (11)) the sequence of $\lambda_m(\cdot)$ weak converges to the function $-\xi(\cdot)$ which we shall denote $q(t)$. Hence the functions $\int_t^T \lambda_m(\tau) d\tau$ uniformly converge to $\int_t^T q(\tau) d\tau$ and $p_m(t)$ converge to some $p(\cdot)$ at every t at which $p(t)$ is continuous and also at zero as immediately follows from the definition. Thus $p(0) = b_0$. This proves the transversality condition

Furthermore, as $\lambda_m(\cdot)$ weak converge to $q(\cdot)$, a certain sequence of convex combinations of $\lambda_m(\cdot)$, say $q_k(\cdot) = \sum_i \beta_{ik} \lambda_{m_{ik}}(\cdot)$ with $\beta_{ik} > 0$, $\sum_i \beta_{ik} = 1$, $m_{ik} \geq m_k$, $m_k \rightarrow \infty$ as $k \rightarrow \infty$, norm converges to $q(\cdot)$ in L^1 , and we may assume that $q_k(t) \rightarrow q(t)$ a.e. As follows from the the third relation in (11), for almost every t

$$(\lambda_m(t), p_m(t)) \in \partial_p L_\delta(t, \cdot, \cdot)(x_{2m}(t), y_{2m}(t)) + \delta_m(t)B$$

with some $\delta_m(t) \rightarrow 0$ a.e.. We may assume of course that both $x_{2m}(t)$ and $y_{2m}(t)$ converge to zero almost everywhere. The required relation for $q(\cdot)$ is now an immediate consequence of Proposition 3.3.

Let $\mathcal{P}$ stand for the collection of triples $(q(\cdot), \nu, p(\cdot))$ such that $\nu \in \partial_G \varphi(\bar{x}(\cdot))$ and $q(\cdot)$ and $p(\cdot)$ satisfy (1)–(3). As φ is Lipschitz and $|q(t)| \leq k_0(t)$ a.e., $\mathcal{P}$ is a compact set in the product of the corresponding weak topologies. The above argument shows that for any $\delta > 0$ there is a $(q(\cdot), \nu, p(\cdot)) \in \mathcal{P}$ such that, as follows from the last relation in (11),

$$\int_0^T (-\langle p(t), u_i(t) \rangle + g_{i\delta}(t)) dt \leq 0, \quad i = 1, \dots, k.$$

Taking into account that $\lim_{\delta \rightarrow 0} g_{i\delta}(t) = L(t, 0, u_i(t)) - L(t, 0, 0)$, we conclude (in view of compactness of $\mathcal{P}$) that for some $(q(\cdot), \nu, p(\cdot)) \in \mathcal{P}$

$$(12) \quad \int_0^T (L(t, 0, u_i(t)) - L(t, 0, 0) - \langle p(t), u_i(t) \rangle) dt \geq 0, \quad i = 1, \dots, k.$$

5. Thus for any finite collection $\mathcal{U} = (u_1(\cdot), \dots, u_k(\cdot))$ of measurable selections of $Q(\cdot)$ such that for any i there are an $\varepsilon_i > 0$ and a summable $c_i(t)$ such that the function $x \rightarrow L(t, \cdot, u_i(t))$ is continuous and satisfies $L(t, x, u_i(t)) \leq c_i(t)$ on $B(\bar{x}(t), \varepsilon_i)$ the set $\mathcal{P} = \mathcal{P}(\mathcal{U})$ of triples $(q(\cdot), \nu, p(\cdot))$ satisfying the conclusion

of the theorem with the Weierstrass condition replaced by (12) is nonempty. As we have mentioned, this set is compact in suitable topologies. On the other hand if $\mathcal{U}$ is a subset of $\mathcal{U}'$, then $\mathcal{P}(\mathcal{U}') \subset \mathcal{P}(\mathcal{U})$. Therefore the intersection of all $\mathcal{P}(\mathcal{U})$ is nonempty.

This means that there is a nontrivial triple $(q(\cdot), \nu, p(\cdot))$ satisfying the relations (1)–(3) and such that (12) remains valid if we replace $u_i(\cdot)$ by any measurable selection $u(\cdot)$ of $Q(\cdot)$ such that for some $\varepsilon > 0$ the function $L(t, \cdot, u(t))$ is continuous on εB and for any x of the ball $L(t, x, u(t)) \leq c(t)$ a.e., where $c(t)$ is a summable function.

Recall that $0 \in Q(t)$ for almost every t by (H_6) . Hence, given a measurable selection $u(\cdot)$ of $Q(\cdot)$ and a measurable $\Delta \subset [0, T]$, the function equal $u(t)$ on Δ and zero outside of Δ is also a measurable selection of $Q(\cdot)$. It follows that

$$\int_{\Delta} (L(t, 0, u(t)) - L(t, 0, 0) - \langle p(t), u(t) \rangle) dt \geq 0.$$

The Weierstrass condition (4) is an immediate consequence of the last inequality. Indeed, assume that it fails to be valid on a set of positive measure on $[0, T]$. By the Aumann selection theorem (see e.g. [20]) we can choose a measurable selection $u(t)$ of $Q(\cdot)$ such that

$$L(t, 0, u(t)) - L(t, 0, 0) - \langle p(t), u(t) \rangle < 0$$

on a set Δ of positive measure contrary to (12). So we have to verify that (taking a smaller Δ , if necessary) we can guarantee that for some $\varepsilon > 0$ and summable $c(t)$ the function $u(\cdot)$ satisfies the required properties.

To this end, consider the functions

$$\lambda(t, \varepsilon) = \sup\{L(t, x, u(t)) : \|x\| \leq \varepsilon\};$$

$$\mu(t, \varepsilon) = \lim_{\delta \rightarrow 0} \sup\{|L(t, x, u(t)) - L(t, x', u(t))| : \|x - x'\| \leq \delta, x, x' \in \varepsilon B\}.$$

Both functions are measurable as functions of t and non-decreasing as functions of ε . Moreover, as follows from (H_5) , (H_6) , $\lambda(t, \varepsilon) \searrow L(t, 0, u(t))$ and $\mu(t, \varepsilon) = 0$ if $\varepsilon > 0$ is sufficiently small for almost every $t \in \Delta$. Hence for a sufficiently small $\varepsilon > 0$ there is a subset $\Delta' \subset \Delta$ of positive measure such that $\lambda(t, \varepsilon)$ is summable on Δ' and $\mu(t, \varepsilon) = 0$ if $t \in \Delta'$. The latter means that $L(t, \cdot, u(t))$ is continuous on εB for $t \in \Delta'$. As we can assume without loss of generality that $\varepsilon \leq \varepsilon_0$ (see (H_5)), it remains to extend $u(t)$ to the whole of $[0, T]$ by setting $u(t) = 0$ outside of the set and set $c(t)$ equal to $\lambda(t, \varepsilon)$ if $t \in \Delta'$ and to $L(t, 0, 0) + k_0(t)\varepsilon$ outside of Δ' . This completes the proof of the theorem.

5. Proofs of other theorems.

5.1. Proof of Theorem 2.1. As we have mentioned, the proof of the theorem is a simplification of the proof of Theorem 2.2. The main difference that allows to go through under weaker assumptions on L is that in the proof of Theorem 2.1 we do not need relaxation.

Indeed, take a $\delta > 0$ such that $J(x(\cdot)) \geq J(\bar{x}(\cdot))$ if $\|x(t) - \bar{x}(t)\| \leq \delta$ and $\|\dot{x}(t) - \dot{\bar{x}}(t)\| \leq \delta$ almost everywhere, and let as before $L_\delta(t, x, y)$ be the function coinciding with $L(t, x, y)$ if $\|x - \bar{x}(t)\| \leq \delta$ and $\|y - \dot{\bar{x}}(t)\| \leq \min\{\delta, \gamma_0(t)\}$ and equal to ∞ otherwise. Thus $\bar{x}(\cdot)$ is the minimum of the functional

$$J_\delta(x(\cdot)) = \ell(x(0), x(T)) + \varphi(x(\cdot)) + \int_0^T L_\delta(t, x(t), \dot{x}(t)) dt.$$

From this point we can follow the proof of Theorem 2.2 with all α_i equal to zero. In other words, we set

$$\Lambda(a, y(\cdot)) = a + \int_0^t y(\tau) d\tau$$

$\mathbf{x} = (x(\cdot), a, y(\cdot))$ and consider the functionals

$$\begin{aligned} I_1(\mathbf{x}) &= \ell(a, \Lambda(a, y(T))), & I_2(\mathbf{x}) &= \varphi(\Lambda(a, y(\cdot))), \\ I_3(\mathbf{x}) &= \int_0^T L_\delta(t, x(t), y(t)) dt, & I_4(\mathbf{x}) &= \int_0^T k_0(t) \|x(t) - \Lambda(a, y(\cdot))(t)\| dt. \end{aligned}$$

All subsequent arguments in the analysis part of the proof of Theorem 2.2 with some simplifications work well in our case. Basically, we no longer need the last relation in (11) connected with α_i and all arguments at the end of the proof associated with the Weierstrass condition.

5.2. Proof of Theorems 2.3 and 2.4. Theorem 2.3 is an immediate consequence of Theorem 2.1 and Proposition 3.5.

Let us prove Theorem 2.4. Again, for notational simplicity we assume in the proof that $\bar{x}(t) \equiv 0$. Set

$$\hat{J}(x(\cdot)) = \ell(x(0), x(T)) + \varphi(x(\cdot)) + \int_0^T \hat{L}(t, x(t), \dot{x}(t)) dt.$$

Clearly $J(x(\cdot)) \geq \hat{J}(x(\cdot))$ for any $x(\cdot) \in W^{1,1}$.

On the other hand, take an $\varepsilon > 0$, a positive $\delta < \bar{\delta}/2$ and positive-valued $\gamma(t) < \bar{\gamma}(t)/2$, and let $x(\cdot)$ satisfy $\|x(t)\| < \delta$ and $\|\dot{x}(t)\| < \gamma(t)$ a.e. Let further

$Y_\varepsilon(t)$ be the collection of $2(n+1)$ -tuples $(y_1, \dots, y_{n+1}, \alpha_1, \dots, \alpha_{n+1})$ such that $\alpha_i \geq 0$, $\sum \alpha_i = 1$, $\|y_i\| \leq N_{\bar{\delta}, \bar{\gamma}(t)}(t)$ and

$$\sum \alpha_i y_i = \dot{x}(t), \quad \sum \alpha_i L(t, x(t), y_i) \leq \hat{L}(t, x(t), \dot{x}(t)) + \varepsilon.$$

This set is nonempty for any t and the set-valued mapping $t \rightarrow Y_\varepsilon(t)$ is measurable and closed-valued. This is obvious. Let $(y_1(t), \dots, y_{n+1}(t), \alpha_1(t), \dots, \alpha_{n+1}(t))$ be a measurable selection of Y_ε . By the relaxation theorem (Proposition 3.8) there is a sequence of $u_m(\cdot) \in W^{1,1}$ uniformly converging to $x(\cdot)$ and such that

$$\lim \int_0^T L(t, u_m(t), \dot{u}_m(t)) dt \leq \int_0^T \hat{L}(t, x(t), \dot{x}(t)) dt + \varepsilon T$$

and $u_m(0) = x(0)$, $u_m(T) = x(T)$. It follows (as ε can be arbitrarily small) that zero is a local minimum of $\hat{J}$.

We claim next that the functional $\hat{J}$ satisfies (H₁)–(H₅) with ε_0 replaced by δ , $k_0(t)$ replaced by $k(t) = K_{\bar{\delta}}(N_{\bar{\delta}, \bar{\gamma}(t)}(t), t)$ and $Q(t) = B(\bar{x}(t), \bar{\gamma}(t))$. Indeed, ℓ and φ are the same in $\hat{J}$, whence (H₁) and (H₂), and (H₃) is obviously valid as the definition of $\hat{L}$ uses operations that do not violate measurability of set-valued mappings. To verify (H₅), we note first that $\hat{L}(t, \cdot, y)$ is $k(t)$ -Lipschitz in the δ -neighborhood of zero since $L(t, \cdot, y)$ is $k(t)$ -Lipschitz if $\|y\| \leq N_{\bar{\delta}, \bar{\gamma}(t)}(t)$. On the other hand, as $\hat{L}(t, x, \cdot)$ is convex and finite on the δ -ball around zero, it is a continuous function on the interior of the ball. Together with Lipschitzness of $\hat{L}(t, \cdot, y)$, this implies continuity of $\hat{L}(t, \cdot, \cdot)$. And (H₆) is of course trivially valid if $Q(t) = \bar{\gamma}(t)B$.

Applying Theorem 2.1, we shall find a function $p(\cdot)$ of bounded variation, a $q(\cdot) \in L^1$ and a measure $\nu \in \partial\varphi(0)$ such that the inclusion

$$q(t) \in \text{conv}\{q : (q, p(t)) \in \partial\hat{L}(t, \cdot, \cdot)(0, 0)\}, \text{ a.e.};$$

along with the last two relations stated in the theorem. But as $\hat{L}$ is convex in the last argument, it follows from Proposition 3.5 that

$$q(t) \in \text{conv}\{q : (-q, 0) \in \partial H(t, \cdot, \cdot)(0, p(t))\}, \text{ a.e.};$$

as claimed.

REFERENCES

- [1] F. H. CLARKE. The Euler–Lagrange differential inclusion. *J. Diff. Equations* **19**, 1 (1975), 80–90.
- [2] F. H. CLARKE. The generalized problem of Bolza. *SIAM J. Control Optim.* **14**, 4 (1976), 682–699.
- [3] F. H. CLARKE. Optimization and Nonsmooth Analysis. Canad. Math. Soc. Ser. Monogr. Adv. Texts Wiley-Intersci. Publ., New York, John Wiley & Sons, Inc., 1983.
- [4] F. H. CLARKE. Hamiltonian analysis of generalized problem of Bolza. *Trans. Amer. Math. Soc.* **301**, 1 (1987), 385–400.
- [5] F. H. CLARKE. Necessary Conditions in Dynamic Optimization. *Mem. Amer. Math. Soc.* **173**, 816 (2005), 113 pp.
- [6] F. H. CLARKE, YU. S. LEDYAEV, R. J. STERN, P. R. WOLENSKI. Nonsmooth Analysis and Control Theory. Grad. Texts in Math., vol. **178**. New York, Springer-Verlag, 1998.
- [7] F. H. CLARKE, R. B. VINTER. The relationship between the maximum principle and dynamic programming. *SIAM J. Control Optim.* **25**, 5 (1987), 1291–1311.
- [8] H. Frankowska, Necessary conditions for the Bolza problem. *Math. Oper. Res.* **10**, 2 (1985), 361–366.
- [9] A. D. IOFFE. Euler-Lagrange and Hamiltonian formalisms in dynamic optimization. *Trans. Amer. Math. Soc.* **349**, 7 (1997), 2871–2900.
- [10] A. D. IOFFE. Existence and relaxation theorems for unbounded differential inclusions. *J. Convex Anal.* **13**, 2 (2006), 353–362.
- [11] A. D. IOFFE. Variational Analysis of Regular Mappings. Springer Monogr. Math. Cham, Springer, 2017.
- [12] A. D. IOFFE. On generalized Bolza problem and its application to dynamic optimization. *J. Optim. Theory Appl.* **182**, 1 (2019), 285–309.
- [13] A. D. IOFFE, R. T. ROCKAFELLAR. The Euler and Weierstrass conditions for nonsmooth variational problems. *Calc. Var. Partial Differential Equations* **4**, 1 (1996), 59–87.

- [14] A. JOURANI. Lagrangian and Hamiltonian necessary conditions for the generalized Bolza problem and applications. *J. Nonlinear Convex Anal.* **10**, 3 (2009), 437–454.
- [15] M. KRASTANOV, N. RIBARSKA. A functional analytic approach to a Bolza problem. In: *Lecture Notes in Econom. and Math. Systems*, vol. **687**. Cham, Springer, 2018, 97–117.
- [16] P. D. LOEWEN. *Optimal Control via Nonsmooth Analysis*. CRM Proc. Lecture Notes, vol. **2**. Providence, RI, AMS, 1993.
- [17] P. D. LOEWEN, R. T. ROCKAFELLAR. New necessary conditions for the generalized problem of Bolza. *SIAM J. Control Optim.* **34**, 5 (1996), 1496–1511.
- [18] R. T. ROCKAFELLAR. Existence and duality theorems for convex problems of Bolza. *Trans. Amer. Math. Soc.* **159** (1971), 1–40.
- [19] R. T. ROCKAFELLAR, R. J.-B. WETS. *Variational Analysis*. Grundlehren Math. Wiss., **317** [Fundamental Principles of Mathematical Sciences]. Berlin, Springer-Verlag, 1998.
- [20] R. B. VINTER. *Optimal Control*. Systems Control Found. Appl. Boston, MA, Birkhauser Boston, Inc., 2000.
- [21] R. B. VINTER. The Hamiltonian inclusion for nonconvex velocity sets. *SIAM J. Control Optim.* **52**, 2 (2014), 1237–1250.
- [22] R. B. VINTER, H. ZHENG. The extended Euler-Lagrange conditions in non-convex variational problems. *SIAM J. Control Optim.* **35**, 1 (1997), 56–77.

Department of Mathematics
Technion, Haifa 32000, Israel
e-mail: alexander.ioffe38@gmail.com

Received July 11, 2022

Accepted February 21, 2023