PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 67, Number 1, November 1977

A REMARK ON THE DIRECT METHOD OF THE CALCULUS OF VARIATIONS

J. P. PENOT

ABSTRACT. This note deals with the problem of minimizing a real-valued function f on a weakly closed subset of a reflexive Banach space. We use a mild monotonicity assumption introduced by P. Hess [11] on the derivative f' of f to get the weak lower semicontinuity of f. We show that one can dispense with any continuity assumption on f', so that we get a true generalization of F. E. Browder's results [4]. The relevance of the monotonicity property to the calculus of variations is shown by an example.

0. **Introduction.** In [4] F. E. Browder brought forward the rôle of monotonicity assumptions on the derivative f' of a function f: $X \to \mathbf{R}$, X a reflexive Banach space, to ensure its weak lower semicontinuity. P. Hess showed [11] that the monotonicity assumption of f' can be relaxed provided f' is continuous. In this note we prove that this restriction is unnecessary, so that F. E. Browder's result is encompassed within the more general frame of [11].

1. **Preliminaries.** Let (X, τ) be a locally convex topological vector space (l.c.s.) with dual X'. We denote by σ the weak topology $\sigma(X, X')$ on X and by σ_s the sequential topology associated with σ: a subset A of X is σ_s-closed if and only if A contains the limit of any σ-convergent sequence of A. It is easily seen that σ and σ_s have the same convergent sequences which are called weakly convergent sequences. The lower semicontinuity (l.s.c.) of a map f: $X \to \mathbf{R}$ with respect to σ_s amounts to its sequential weak lower semicontinuity.

The following definition is an extension to multivalued mappings of a definition of P. Hess [11].

DEFINITION 1. A multivalued mapping F: $X \to X'$ is called pertinent if it satisfies:

(P) If (x_n) is a sequence of dom F weakly converging to $x \in$ dom F, if (x_n') is a sequence of X' with $x_n' \in F(x_n)$ for every n, then $\limsup\langle x_n', x_n - x\rangle \geqslant 0$.

As observed by P. Hess, condition (P) is an extremely mild monotonicity assumption. Multivalued monotone mappings, multivalued pseudo-monotone

Received by the editors March 26, 1976 and, in revised form, January 5, 1977.

AMS (MOS) subject classifications (1970). Primary 47H05, 49A25, 49A50.

Key words and phrases. Variational problem, direct method, weak lower semicontinuity, operator of monotone type, pseudomonotone operator, pertinent operator, nonlinear eigenvalue problem, multiple integral functional.

mappings [2], [6], [7], and demicontinuous bounded operators of type (S_+) [5], [6] are examples of pertinent mappings. In particular, if X is finite dimensional, any continuous mapping from X to X' is pertinent. In [5] F. E. Browder has given analytical conditions on concrete operators of Euler-Lagrange type arising from multiple integral functionals which can be used to verify condition (P). A simple case is described in the last part of the present paper. The class of pertinent mappings is obviously stable under addition.

Let us say that a multivalued mapping $F\colon X \to X'$ is *sequentially bounded* on $D \subset \operatorname{dom} F$ if for each converging sequence (x_n) of D with limit in D, and each $x_n' \in F(x_n)$, the sequence (x_n') is (strongly) bounded in X'. It has been shown by P. M. Fitzpatrick, P. Hess and T. Kato [10], generalizing a result of F. E. Browder [3] and T. Rockafellar [18], that if X is a Fréchet space, if D is convex and algebraically open (i.e. for every $x \in D, y \in X$ there exists $\varepsilon > 0$ with $x + ty \in D$ for every $t \in [0, \varepsilon]$), and if F is monotone, then F is sequentially bounded.

Let D be an arbitrary subset of X. The *radial tangent cone* $T_a^r D$ to D at $a \in \overline{D}$ is the set of $v \in X$ such that 0 is an accumulation point of the set $D_{a,v}$ of $t \in \mathbf{R}_+$ such that $a + tv \in D$. It is easily seen that if D is convex $T_a^r D$ is the cone generated by $D - a = \{x - a | x \in D\}$. We define a linear form $x' \in X'$ to be a G-differential of $f\colon D \to \mathbf{R}$ at $a \in D$ if for every $v \in T_a^r D$ we have

$$\lim_{t \to 0_+;\, t \in D_{a,v}} t^{-1}\left[f(a + tv) - f(a)\right] = \langle x', v\rangle.$$

If $T_a^r D$ is total in X, in particular, if D is algebraically open, then the set $f'(a)$ of G-differentials of f at a is at most one point. The map f is said to be G-differentiable on D if $f'(a) \neq \varnothing$ for every $a \in D$. Let us observe that if f is G-differentiable on D, for every $a \in D$ and every $v \in T_a^r D$ the value $f'(a) \cdot v$ is unambiguously defined by $f'(a) \cdot v = \langle x', v\rangle$ with $x' \in f'(a)$ arbitrary.

2. **Main result.**

THEOREM 1. *Let (X, τ) be a l.c.s., let D be a convex subset of X or a σ_s-open subset of X. Let $f\colon D \to \mathbf{R}$ be a G-differentiable function whose derivative f' is pertinent. If f' is sequentially bounded, or if f is l.s.c. for τ, then f is l.s.c. for σ_s.*

PROOF. Let us show that if (x_n) is a sequence of D which weakly converges to $x \in D$, then the inequality $\liminf f(x_n) < f(x)$ is impossible. The assumptions on D allow us to suppose that for every n the intervals

$$[x, x_n] = \{x + t(x_n - x) | t \in [0, 1]\}$$

are contained in D. Taking a subsequence if necessary, we may also suppose that $(f(x_n))$ converges to a limit $r \in [-\infty, f(x) - c[$, with $c > 0$, and that $f(x_n) < f(x) - c$ for every $n \in \mathbf{N}$. The following lemma is the first step of the proof.

LEMMA. *There exists $\varepsilon > 0$ so that for every $n \in \mathbf{N}$ the following inequality holds*:

$$f(x + \varepsilon(x_n - x)) - f(x) \geqslant -c/2. \tag{1}$$

PROOF OF THE LEMMA. The result is obvious if f is τ-sequentially l.s.c., for $(x_n - x)$ is bounded; hence for any increasing sequence (n_k) and any sequence (t_k) of $\mathbf{R}_+$ converging to 0, the sequence $(t_k(x_{n_k} - x))$ converges to 0. Let us now suppose that f' is sequentially bounded. We claim that there exist $d \in]0, 1[$ and $M > 0$ such that

$$\langle f'(x + t(x_n - x)), x_n - x\rangle \geqslant -M \tag{2}$$

for every $n \in \mathbf{N}$ and every $t \in]0, d[$. Let us suppose on the contrary that there exists a sequence (t_k) of $[0, 1]$ converging to 0, and a sequence (n_k) of $\mathbf{N}$ with

$$\langle f'(x + t_k(x_{n_k} - x)), x_{n_k} - x\rangle \leqslant -k.$$

As $b_k = x_{n_k} - x$ defines a bounded sequence, and as $a_k = x + t_k b_k$ converges to x, we obtain a contradiction to the fact that f' is sequentially bounded. Hence our claim is justified.

We choose $\varepsilon \in]0, d[$ with $\varepsilon M \leqslant c/2$ and we apply the classical mean value theorem (J. Dieudonné [9, 8.5.3]) to the functions $t \to f(x + t(x_n - x))$ between 0 and ε to get

$$f(x + \varepsilon(x_n - x)) - f(x) \geqslant -\varepsilon M \geqslant -c/2. \quad \square \tag{1}$$

PROOF OF THE THEOREM. We now apply the mean value theorem to the same functions, but between ε and 1:

$$\begin{aligned} f(x_n) - f(x + \varepsilon(x_n - x)) \\ \geqslant (1 - \varepsilon) \inf_{t \in]\varepsilon, 1[} \langle f'(x + t(x_n - x)), x_n - x\rangle. \end{aligned} \tag{3}$$

Let us show the existence of an $m \in \mathbf{N}$ such that

$$(1 - \varepsilon) \inf_{t \in]\varepsilon, 1[} \langle f'(x + t(x_n - x)), x_n - x\rangle \geqslant -c/2 \tag{4}$$

for every $n \geqslant m$. If no such m exists we can find sequences $(n_k) \subset \mathbf{N}$ with $\lim n_k = +\infty$, $(t_k) \subset]\varepsilon, 1[$ such that

$$(1 - \varepsilon)\langle f'(a_k), x_{n_k} - x\rangle < -c/2$$

with $a_k = x + t_k(x_{n_k} - x)$. Hence,

$$\langle f'(a_k), a_k - x\rangle = t_k\langle f'(a_k), x_{n_k} - x\rangle < -c\varepsilon/2(1 - \varepsilon),$$

and we get a contradiction to (P) as (a_k) weakly converges to x. Gathering (1), (3) and (4) we get $f(x_n) - f(x) \geqslant -c$ for $n \geqslant m$, a contradiction to what we assumed. $\square$

3. **An extension.** The above result still rules out continuous convex functions which are not differentiable. To cover such a fundamental case of lower semicontinuity, we remark that the use of the derivative of f was limited to the application of the mean value theorem. Various subdifferentials or paradifferentials of f may play this rôle. Let us state the needed hypothesis in

sufficiently general terms so that our formulation is not tied to a particular choice of such a notion of generalized differential.

DEFINITION 2. A multivalued function F: $D \to X'$ is a generalized mean for f: $D \to \mathbf{R}$ if its values are nonvoid and if for every a, b in D the following inequality holds:

$$\text{(MV)} \qquad f(b) - f(a) \geqslant \inf\{\langle y, b - a\rangle | y \in F(x), x \in [a, b]\}.$$

It is shown in [17] that if one of the subdifferentials

$$\underline{\partial} f(x) \subset \underline{\partial}_s f(x) \subset \underline{\partial}_r f(x) \subset \bar{\partial}_r f(x)$$

introduced in [15] and [16] is always nonvoid for $x \in D$, then it is a generalized mean for f. In fact one has then the stronger inequality

$$\text{(MV}'\text{)} \quad f(b) - f(a) \geqslant \inf\{\sup(\langle y, b - a\rangle | y \in F(x)) | x \in [a, b]\}.$$

Under infinitesimal convexity assumptions on f one can also use the subdifferentials $\bar{\partial}_s f$ and $\bar{\partial} f$ of [15], [16]. G. Lebourg shows in [13] that when f is locally lipschitz the generalized gradient introduced by F. Clarke [8] (see also L. Thibault [19]) is a generalized mean for f. Another instance is described in R. Janin [12].

THEOREM 2. *Let X, D be as in Theorem* 1. *Suppose f: $D \to \mathbf{R}$ has a pertinent generalized mean F. If F is sequentially bounded or if f is l.s.c. for τ then f is l.s.c. for σ_s.*

The proof is a simple adjustment of the proof of Theorem 1. We replace assertion (2) by the claim of the existence of $d \in]0, 1[$ and $M > 0$ such that for every $t \in [0, d]$, every $n \in \mathbf{N}$ and every $y \in F(x + t(x_n - x))$, one has

$$(2') \qquad \langle y, x_n - x\rangle \geqslant -M.$$

Analogously, we replace assertion (4) by the claim of the existence of an $m \in \mathbf{N}$ such that for every $n \geqslant m$, every $t \in [\varepsilon, 1]$ and every $y \in F(x + t(x_n - x))$, one has

$$(4') \qquad (1 - \varepsilon)\langle y, x_n - x\rangle \geqslant -c/2. \quad \square$$

REMARK. If we suppose that F verifies (MV′) instead of (MV) we may replace (P) by the weaker condition:

(P′) For every sequence (x_n) of D which weakly converges to $x \in D$, there exists a sequence (x_n') in X' with $x_n' \in F(x_n)$ for every n and $\limsup\langle x_n', x_n - x\rangle \geqslant 0$.

COROLLARY 1. *With the assumptions of Theorems* 1 *or* 2 *suppose that A is a subset of D. If $f|A$ is bounded below and if a minimizing sequence of $f|A$ converges weakly to a point $a \in A$, then $f|A$ achieves its minimum at a.*

PROOF. If $(a_n) \subset A$ converges weakly to $a \in A$ and if $(f(a_n)) \to \inf f(A)$, we have $f(a) \leqslant \liminf f(a_n) = \inf f(A)$; hence $f(a) = \inf f(A)$. $\square$

COROLLARY 2. *Suppose X is a reflexive Banach space, A is a closed subset of*

(X, σ_s) (i.e. a weakly sequentially closed subset of X) and f satisfies the assumptions of Theorems 1 *or* 2 *with $D \supseteq A$, for instance $D = X$. If*

$$\text{(CO)} \qquad \liminf_{\|x\| \to \infty,\, x \in A} f(x) > \inf_{x \in A} f(x)$$

then $f|A$ achieves its minimum value and any minimizing sequence has a weak limit point at which f assumes its minimum on A.

The coercivity condition (CO) is equivalent to the boundedness of all the minimizing sequences of $f|A$. It is obviously implied by the more usual coercivity condition $\lim_{\|x\| \to \infty; x \in A} f(x) = +\infty$.

COROLLARY 3. *Suppose X is a reflexive Banach space, C is a closed convex subset of X and A: $C \to X'$ is the G-derivative of a continuous function f: $C \to \mathbf{R}$. If A is pertinent and if for some $x_0 \in C$ one has $\langle A(x), x - x_0\rangle / \|x\| \to +\infty$ as $\|x\| \to +\infty$, $x \in C$, then for every $y \in X'$ there exists a solution z of the variational inequality*

$$\text{(VI)} \qquad \langle A(z), x - z\rangle \geqslant \langle y, x - z\rangle \qquad \forall x \in C.$$

In particular, for $C = X$ we obtain the surjectivity of A.

PROOF. Let g: $C \to \mathbf{R}$ be defined by $g(x) = f(x) - \langle y, x\rangle$. As g is continuous, pertinent and coercive, Corollary 2 implies that g reaches its minimum on C, say at $z \in C$. Then we have $g'(z)v \geqslant 0$ for every $v \in T_z^r C$, which is exactly (VI) as $g'(z) = A(z) - y$ and $T_z^r C = \mathbf{R}_+(C - z)$. $\square$

REMARK. An inspection of the proof of Theorem 1 shows that the results of Corollaries 1 and 2 hold if (P) is replaced by the following weaker assumption:

(Q) If (x_n) is a minimizing sequence which converges weakly to x and if (t_n) is a sequence of $[0, 1]$, then

$$\limsup \langle f'(x + t_n(x_n - x)), x_n - x\rangle \geqslant 0.$$

This condition is verified if f is quasi-convex, continuous and G-differentiable.

Using Corollary 2, applications of the preceding results to nonlinear eigenvalue problems or to problems of the calculus of variations can be given along the lines of [5] and [11]. The next section deals with the lower semicontinuity question for a multiple integral functional; an existence result follows from coercivity properties on the integrand.

4. **An application.** Let T be the closure of a regular bounded open subset of $\mathbf{R}^d$ (or, more generally, a compact manifold of dimension d) and let E be a finite dimensional vector space (or, more generally, a vector bundle over T with finite dimensional fibers). We denote by X the Sobolev space $W^{1,p}(T, E)$ of p-integrable functions from T to E with distributional derivatives in $L_p(T, E)$ with $p \in [1, \infty)$. Let $q \in [1, \infty)$ be arbitrary if $p \geqslant d$ and strictly less than $pd(d-p)^{-1}$ if $p < d$. Then we have a compact embedding of $W^{1,p}(T, E)$ into $L_q(T, E)$. Let L: $E \times E^d \times T \to \mathbf{R}$ be a

continuous map with continuous partial derivatives with respect to E and E^d satisfying the following conditions (a), (b), (c).

(a)

$$|D_1L(e, v, t)| \leqslant a_1(t) + b_1|e|^{q/s} + b_1|v|^{p/s} \quad \text{for } (e, v, t) \in E \times E^d \times T;$$

$$|D_2L(e, v, t)| \leqslant a_2(t) + b_2|e|^{q/r} + b_2|v|^{p/r} \quad \text{for } (e, v, t) \in E \times E^d \times T$$

with $a_1 \in L_s(T, \mathbf{R})$, $a_2 \in L_r(T, \mathbf{R})$, $b_1 \in \mathbf{R}_+$, $b_2 \in \mathbf{R}_+$, $r^{-1} + p^{-1} = 1$, $s^{-1} + q^{-1} = 1$.

(b) For every $(e, t) \in E \times (T \setminus N)$ where $N \subset T$ has zero measure, the function $v \mapsto L(e, v, t)$ is convex;

(c) There exists $\bar{x} \in W^{1,p}(T, E)$ such that $t \to L(\bar{x}(t), D\bar{x}(t), t)$ is integrable. Then it is easily seen that for every $x \in W^{1,p}(T, E)$ the functional

$$J(x) = \int_T L(x(t), Dx(t), t)\, dt$$

is well defined and is of class C^1 with derivative given by

$$J'(x) \cdot y = \int_T \left[D_1L(x(t), Dx(t), t) \cdot y(t) + D_2L(x(t), Dx(t), t) \cdot Dy(t) \right] dt.$$

THEOREM 3. *Under the preceding conditions on L the functional J is sequentially weakly lower semicontinuous on $X = W^{1,p}(T, E)$.*

PROOF. It suffices to show that J' is pertinent. Let (x_n) be a sequence of $W^{1,p}(T, E)$ weakly converging to $x \in W^{1,p}(T, E)$. Without loss of generality we may suppose that (x_n) converges to x in $L_q(T, E)$ as we have to prove that

$$\limsup J'(x_n)(x_n - x) \geqslant 0.$$

We can write $J'(x_n)(x_n - x)$ as the sum of the five following integrals in which expressions like $D_1L_t(x, Dx)$ stand for $D_1L(x(t), Dx(t), t)$:

$$I_1(x_n) = \int_T \left[D_1L_t(x_n, Dx_n) - D_1L_t(x, Dx) \right] \cdot (x_n(t) - x(t))\, dt,$$

$$I_2(x_n) = \int_T D_1L_t(x, Dx) \cdot (x_n(t) - x(t))\, dt,$$

$$I_3(x_n) = \int_T \left[D_2L_t(x_n, Dx_n) - D_2L_t(x_n, Dx) \right] \cdot (Dx_n(t) - Dx(t))\, dt,$$

$$I_4(x_n) = \int_T \left[D_2L_t(x_n, Dx) - D_2L_t(x, Dx) \right] \cdot (Dx_n(t) - Dx(t))\, dt,$$

$$I_5(x_n) = \int_T D_2L_t(x, Dx) \cdot (Dx_n(t) - Dx(t))\, dt.$$

Weak convergence of (Dx_n) to Dx in $L_p(T, E^d)$ implies that $I_5(x_n) \to 0$; similarly $(I_2(x_n)) \to 0$. As (x_n) and (Dx_n) are bounded sequences in $L_q(T, E)$ and $L_p(T, E^d)$, respectively, the estimate on D_1L in (a) and the convergence of (x_n) to x in $L_q(T, E)$ show that $(I_1(x_n)) \to 0$. Similarly, $(I_4(x_n)) \to 0$ as $|Dx_n - Dx|_p$ is bounded, and

$$\left|D_2L(x_n(\cdot), Dx(\cdot), \cdot) - D_2L(x(\cdot), Dx(\cdot), \cdot)\right|_r \to 0$$

by general results on Nemytskii's operator satisfying estimate similar to the second one in (a). Condition (b) implies that $I_3(x_n)$ is nonnegative. The result follows. □

Higher order integrands can be handled similarly. Using covariant derivatives and Finsler structures the same result holds for vector bundle sections with the same proof. Let us observe that slightly more general results have been obtained by L. D. Berkovitz [1], C. Olech [14] and others by different methods which avoid differentiability of the integrand.

References

1. L. D. Berkovitz, *Lower semi-continuity of integral functionals*, Trans. Amer. Math. Soc. **192** (1974), 51–57.

2. H. Brezis, *Equations et inequations non-linéaires dans les espaces vectoriels en dualité*, Ann. Inst. Fourier (Grenoble) **18** (1968), 115–175.

3. F. E. Browder, *Nonlinear monotone and accretive operators in Banach spaces*, Proc. Nat. Acad. Sci. U. S. A. **61** (1968), 388–393.

4. _____, *Pseudo-monotone operators and the direct method of the calculus of variations*, Arch. Rational Mech. Anal. **38** (1970), 268–277.

5. _____, *Existence theorems for nonlinear partial differential equations*, Proc. Sympos. Pure Math., vol. 16, Amer. Math. Soc., Providence, R. I., 1970, pp. 1–60. MR **42** #4855.

6. _____, *Nonlinear operators and nonlinear equations of evolution in Banach spaces*, Proc. Sympos. Pure Math., vol. 18, part 2, Amer. Math. Soc., Providence, R. I., 1976.

7. F. E. Browder and P. Hess, *Nonlinear mappings of monotone type in Banach spaces*, J. Functional Analysis **11** (1972), 251–294.

8. F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.

9. J. Dieudonné, *Foundations of modern analysis*, Academic Press, New York and London, 1960.

10. P. M. Fitzpatrick, P. Hess and T. Kato, *Local boundedness of monotone-type operators*, Proc. Japan Acad. **48** (1972), 275–277.

11. P. Hess, *A variational approach to a class of nonlinear eigenvalue problems*, Proc. Amer. Math. Soc. **29** (1971), 272–276.

12. R. Janin, *Sur la dualité et la sensibilité dans les problèmes de programmation mathématique*, Thèse, Paris, 1974.

13. G. Lebourg, *Valeur moyenne pour gradient généralisé*, C. R. Acad. Sci. Paris Sér. A **281** (1975), A795–A797.

14. C. Olech, *Weak lower semi-continuity of integral functionals*, J. Optimization Theory and Appl. **19** (1976), 3–16.

15. J. P. Penot, *Calcul sous-différentiel et optimisation*, J. Functional Analysis (to appear).

16. _____, *Sous-différentiels de fonctions numériques non convexes*, C. R. Acad. Sci. Paris **278** (1974), 1553–1555.

17. _____, *Règles de calcul sous-différentiel* (to appear).

18. T. Rockafellar, *Local boundedness of nonlinear maximal monotone operators*, Michigan Math. J. **16** (1969), 397–407. MR **40** #6229.

19. L. Thibault, *Quelques propriétés des sous-différentiels de fonctions réelles localement lipschitziennes définies sur un espace de Banach séparable*, C. R. Acad. Sci. Paris Sér. A **282** (1976), A507–A510.

Département de Mathématiques, BP 523, Université de Pau, Pau, France 64010