TAIWANESE JOURNAL OF MATHEMATICS
Vol. 4, No. 4, pp. 627-634, December 2000

ON THE PALAIS-SMALE CONDITION FOR NONDIFFERENTIABLE FUNCTIONALS

Hong-Kun Xu

Abstract. Two kinds of Palais-Smale condition, $(PS)_c$ and $(PS)_c^*$, for nondifferentiable functionals are studied. It is shown that $(PS)_c$ implies $(PS)_c^*$ and that they are equivalent for convex functionals. This points out a gap in the proof of Costa and Goncalves [5, Proposition 3]. Some other nonsmooth versions of known smooth results are also obtained.

1. Introduction

We are concerned in this paper with the Palais-Smale condition for nondifferentiable functionals which was first introduced by Chang [4] in 1981. Later, in 1990, Costa and Goncalves [6] proposed another kind of Palais-Smale condition for nondifferentiable functionals and they claimed that their condition, denoted $(PS)_c^*$, is equivalent to Chang's condition, denoted $(PS)_c$. However, there is a gap in their proof. (In fact, their proof works only when the involving functional f satisfies the property that the Clarke differential $\partial f(x)$ is a singleton for all x, i.e., f is strictly differentiable in the sense of Clarke [5].) In this note, we first show that Chang's condition $(PS)_c$ implies Costa and Goncalves' condition $(PS)_c^*$ and that the two conditions coincide for convex functionals. (It is however not clear yet whether this remains true in general.) Then we show that for an important and useful class of functionals, Chang's condition $(PS)_c$ is equivalent to its weak version $(PS)_{c,w}$. The third result of the paper states that Chang's condition (PS) is versus the coercivity for nondifferentiable functionals, which presents a nonsmooth version of the smooth result due to Li [12] (see also Costa and Silva [7] and Brezis and Nirenberg [2]).

[0] Received April 27, 1999.
Communicated by P. Y. Wu.
2000 *Mathematics Subject Classification*: Primary 58E05; Secondary 58E30, 58C20.
Key words and phrases: Palais-Smale condition, nondifferentiable functional, generalized derivative, Clarke's differential, critical point, Ekeland's Principle.
This work was supported in part by NRF (South Africa).

2. The Palais-Smale Condition

Let X be a real Banach space with norm $\|\cdot\|$ and X^* be its dual. By $\langle\cdot,\cdot\rangle$ we denote the pairing between X and X^*. Let $f: X \to \mathbb{R}$ be a locally Lipschitzian functional on X (notation: $f \in \mathrm{Lip}_{\mathrm{loc}}(X,\mathbb{R})$), i.e., for each $x \in X$, there is a neighbourhood $N(x)$ of x and a constant k depending on $N(x)$ such that

$$|f(y) - f(z)| \le k\|y - z\| \quad \forall y, z \in N(x).$$

The generalized directional derivative of f at $x \in X$ in the direction $v \in X$ is defined as the number

$$f^\circ(x; v) := \limsup_{\substack{h\to 0\\ \lambda\downarrow 0}} \frac{1}{\lambda}[f(x + h + \lambda v) - f(x + h)].$$

The Clarke subdifferential of f at x is defined as the set (cf. [5])

$$\partial f(x) := \{x^* \in X^* : f^\circ(x; v) \ge \langle v, x^*\rangle \quad \forall v \in X\},$$

i.e., $\partial f(x)$ is the subdifferential of the convex functional $f^\circ(x;\cdot)$ in the sense of convex analysis. A point $x \in X$ is called a critical point of f if $0 \in \partial f(x)$, i.e., $f^\circ(x; v) \ge 0$ for all $v \in X$. A real number c is said to be a critical value of f if there exists a critical point x of f for which $f(x) = c$. For properties of the generalized derivatives and Clarke's differentials, the reader is referred to [5] and [4].

With an $f \in \mathrm{Lip}_{\mathrm{loc}}(X,\mathbb{R})$, we associate a function λ on X defined by

$$\lambda(x) = \min\{\|x^*\| : x^* \in \partial f(x)\}.$$

It is known [4] that λ is lower semicontinuous.

Definition 1 [4]. A functional $f \in \mathrm{Lip}_{\mathrm{loc}}(X,\mathbb{R})$ is said to satisfy the Palais-Smale condition at level c in Chang's sense, denoted $(PS)_c$, if any sequence $(x_n) \subset X$ such that $f(x_n) \to c$ and $\lambda(x_n) \to 0$ possesses a convergent subsequence.

Inspired by Ekeland's Principle [8], Costa and Goncalves [6] proposed another kind of Palais-Smale condition for nondifferentiable functionals.

Definition 2 [6]. A functional $f \in \mathrm{Lip}_{\mathrm{loc}}(X,\mathbb{R})$ is said to satisfy the Palais-Smale condition at level c in the sense of Costa and Goncalves, denoted $(PS)^*_c$, if it satisfies both $(PS)^*_{c,+}$ and $(PS)^*_{c,-}$, where

$(PS)^*_{c,+}$: Whenever $(x_n) \subset X$, (ε_n), $(\delta_n) \subset \mathbb{R}^+$ are sequences with $\varepsilon_n \to 0$, $\delta_n \to 0$, and such that

$$f(x_n) \to c,$$

$$f(x_n) \le f(x) + \varepsilon_n \|x_n - x\| \quad \text{if} \quad \|x_n - x\| \le \delta_n,$$

then (x_n) possesses a convergent subsequence. $(PS)^*_{c,-}$ is similarly defined by interchanging x and x_n in the above inequality.

In their Proposition 3 [6], Costa and Goncalves claimed that in a reflexive Banach space X, their condition $(PS)^*_c$ is equivalent to Chang's condition $(PS)_c$ for every real number c. Unfortunately, there is a gap in their proof (see [6, p. 483]). As a matter of fact, their proof works only in case the functional $f \in \mathrm{Lip}_{\mathrm{loc}}(X, \mathbb{R})$ is strictly differentiable on X in the sense of Clarke [5, Proposition 2.2.4], i.e., $\partial f(x)$ is a singleton for all $x \in X$, which is clearly not the general case. So, it is unclear if $(PS)_c$ is equivalent to $(PS)^*_c$. However, it is fortunate that all the other results in [6] remain true since all the involving functionals indeed do satisfy Chang's condition $(PS)_c$. We now present the following result.

Theorem 1. *Suppose X is a real Banach space and $f \in \mathrm{Lip}_{\mathrm{loc}}(X, \mathbb{R})$. Then for any $c \in \mathbb{R}$, $(PS)_c$ implies $(PS)^*_c$. If, in addition, f is convex, then $(PS)_c$ and $(PS)^*_c$ are equivalent.*

Proof. Assume that f satisfies $(PS)_c$ and $(x_n) \subset X$, (ε_n), $(\delta_n) \subset \mathbb{R}^+$ are sequences such that $\varepsilon_n \to 0$, $\delta_n \to 0$, and

$$f(x_n) \to c, \tag{2.1}$$

$$f(x_n) \le f(x) + \varepsilon \|x_n - x\| \quad \text{if} \quad \|x_n - x\| \le \delta_n. \tag{2.2}$$

Set

$$f_n(x) = f(x) + \varepsilon_n \|x_n - x\| \quad \text{and} \quad M_n = \{x \in X : \|x_n - x\| \le \delta_n\}.$$

It is easily seen that $x_n \in \operatorname{int} M_n$ is a minimum point of f_n over M_n. Hence $0 \in \partial f(x_n)$ by [5, Proposition 2.3.2], which implies that

$$0 \in \partial f(x_n) + \varepsilon_n B_{X^*} \tag{2.3}$$

for $\partial \|x_n - \cdot\| \subset B_{X^*}$, the closed unit ball of X^*. It follows from (2.3) that

$$\lambda(x_n) = \min\{\|x^*\| : x^* \in \partial f(x_n)\} \le \varepsilon_n \to 0. \tag{2.4}$$

Combining (2.1) and (2.4) we get by $(PS)_c$ that (x_n) admits a convergent subsequence. This verifies that f satisfies $(PS)^*_{c,+}$. $(PS)^*_{c,-}$ is similarly verified.

To conclude the proof, we now assume, in addition, that f is also convex. We obviously need only to show the implication: $(PS)^*_c \Longrightarrow (PS)_c$. Towards this end, we let $(x_n) \subset X$ be given so that

$$f(x_n) \to c \quad \text{and} \quad \lambda(x_n) \to 0.$$

Since $\partial f(x_n)$ is w^*-compact convex, there is $x_n^* \in \partial f(x_n)$ such that $\|x_n^*\| = \lambda(x_n)$. Set

$$\varepsilon_n = \begin{cases} \|x_n^*\|, & \text{if } \|x_n^*\| > 0, \\ \frac{1}{n}, & \text{otherwise.} \end{cases}$$

By the subdifferential inequality for a convex function, we obtain that for all $x \in X$,

$$\begin{aligned} f(x_n) &\le f(x) + \langle x_n - x, x_n^* \rangle \\ &\le f(x) + \|x_n - x\| \, \|x_n^*\| \\ &\le f(x) + \varepsilon_n \|x_n - x\|. \end{aligned}$$

It then follows from the $(PS)_{c,+}^*$ that (x_n) has a convergent subsequence. ∎

Definition 3. A functional $f \in \mathrm{Lip}_{\mathrm{loc}}(X, \mathbb{R})$ is said to satisfy $[PS]_c$ whenever $(x_n) \subset X$ fulfils the property:

$$f(x_n) \to c \quad \text{and} \quad \lambda(x_n) \to 0,$$

then c is a critical value of f.

Definition 4 [6]. A functional $f \in \mathrm{Lip}_{\mathbf{loc}}(X, \mathbb{R})$ is said to satisfy $[PS]_c^*$ if it satisfies both $[PS]_{c,+}^*$ and $[PS]_{c,-}^*$, where

$[PS]_{c,+}^*$: Whenever $(x_n) \subset X$, (ε_n), $(\delta_n) \subset \mathbb{R}^+$ are sequences with $\varepsilon_n \to 0$, $\delta_n \to 0$, and such that

$$f(x_n) \to c,$$

$$f(x_n) \le f(x) + \varepsilon_n \|x_n - x\| \quad \text{if} \quad \|x_n - x\| \le \delta_n,$$

then c is a critical value of f. $[PS]_{c,-}^*$ is defined similarly by interchaning x and x_n in the above inequality.

By the same argument as above, we have the following result.

Theorem 2. *For every $c \in \mathbb{R}$ and $f \in \mathrm{Lip}_{\mathrm{loc}}(X, \mathbb{R})$, $[PS]_c$ implies $[PS]_c^*$. If, in addition, f is convex, then $[PS]_c$ and $[PS]_c^*$ are equivalent.*

Remark. If the statement "(x_n) has a convergent subsequence" in $(PS)_c$, $(PS)_c^*$, $[PS]_c$, and $[PS]_c^*$ is replaced by the statement "(x_n) has a weakly convergent subsequence", then we have the corresponding concepts of $(PS)_{c,w}$, $(PS)_{c,w}^*$, $[PS]_{c,w}$, and $[PS]_{c,w}^*$, respectively. It is clear that there hold the implications: $(PS)_c \Rightarrow (PS)_{c,w}$, $[PS]_c \Rightarrow [PS]_{c,w}$, $(PS)_c^* \Rightarrow (PS)_{c,w}^*$, and $[PS]_c^* \Rightarrow [PS]_{c,w}^*$. It is also easy to verify the implications: $(PS)_{c,w} \Rightarrow (PS)_{c,w}^*$ and $[PS]_{c,w} \Rightarrow [PS]_{c,w}^*$ and the inverse implications: $(PS)_{c,w}^* \Rightarrow (PS)_{c,w}$ and $[PS]_{c,w}^* \Rightarrow [PS]_{c,w}$ for convex $f \in \mathrm{Lip}_{\mathrm{loc}}(X, \mathbb{R})$. We shall show below in some important cases which are often met in applications the equivalences: $(PS)_c \Leftrightarrow (PS)_{c,w}$, $[PS]_c \Leftrightarrow [PS]_{c,w}$, $(PS)^*_{c,w} \Leftrightarrow (PS)^*_c$, and $[PS]^*_c \Leftrightarrow [PS]^*_{c,w}$.

Theorem 3. *Suppose X is a reflexive Banach space whose norm and dual norm are Frechet differentiable, Y is a reflexive Banach space such that X is compactly embedded in Y and is dense in Y (as a subspace of Y). Suppose also f is a locally Lipschitzian functional defined on Y and $\hat{f}$ is the restriction of f on X, i.e., $\hat{f} = f\mid_X$. Let*

$$F(x) = \frac{1}{p}\|x\|_X^p - \hat{f}(x), \quad x \in X,$$

where $1 < p < \infty$ is a constant. Then
(i) *F satisfies $(PS)_c$ if and only if F satisfies $(PS)_{c,w}$;*
(ii) *F satisfies $(PS)^*_c$ if and only if F satisfies $(PS)^*_{c,w}$.*

Proof. Denote by $J_p(x)$ the subdifferential of the convex function $\frac{1}{p}\|\cdot\|^p$ at $x \in X$. Since X is Frechet differentiable, $J_p : X \to X^*$ is single-valued and norm-to-norm continuous. By Theorem 2.2 of Chang [4], we have $\partial\hat{f}(x) \subset \partial f(x)$ for all $x \in X$. Hence

$$\partial F(x) = J_p(x) - \partial\hat{f}(x) \subseteq J_p(x) - \partial f(x), \quad x \in X.$$

To prove (i), it suffices to demonstrate the implication: $(PS)_{c,w} \Rightarrow (PS)_c$. So assume $(x_n) \subset X$ satisfies the properties: $F(x_n) \to c$ and $\lambda(x_n) = \min\{\|x_n^*\| : x_n^* \in \partial F(x_n)\} \to 0$. Suppose $x_n^* \in \partial F(x_n)$ and $y_n^* \in \partial\hat{f}(x_n)$ are chosen so that $\|x_n^*\| = \lambda(x_n)$ and $x_n^* = J_p(x_n) - y_n^*$. In view of the $(PS)_{c,w}$, we may assume that (x_n) converges weakly to some $z \in X$. Noting that X is compactly embedded in Y, we may further assume that the convergence of (x_n) to z is strong. It follows that (y_n^*) is bounded in Y^*. By reflexivity of Y, we may assume that (y_n^*) converges weakly to a $y^* \in Y^*$. The compact embedding of Y^* into X^* then ensures the strong convergence of (y_n^*) to y^* in X^*. Therefore, $x_n = J_p^{-1}(x_n^* + y_n^*)$ converges strongly to $J_p^{-1}(y^*) \in X^*$ because J_p^{-1} is the duality mapping from X^* to X which is norm-to-norm continuous as X^* is Frechet differentiable. (i) is thus verified.

Next we show $(PS)^*_{c,+,w} \Rightarrow (PS)^*_{c,+}$. Assume $(x_n) \subset X$ satisfies the properties:

$$F(x_n) \to c, \tag{2.5}$$

$$F(x_n) \le F(x) + \varepsilon_n\|x_n - x\| \quad \text{if} \quad \|x_n - x\| \le \delta_n, \tag{2.6}$$

where $\varepsilon_n \downarrow 0$ and $\delta_n \downarrow 0$. Applying the $(PS)^*_{c,w}$, we have a subsequence of (x_n) (denoted (x_n) again) weakly converging to some $z \in X$. Now repeating the same proof of (i), we conclude that the convergence of (x_n) to z is actually in norm. The implication: $(PS)^*_{c,-,w} \Rightarrow (PS)^*_{c,-}$ can be proved similarly. ∎

3. (PS) versus Coercivity

By using the method of gradient flows, Li [12] (see also [13]) first observed that the (PS) condition implies the coercivity for C^1 functionals bounded from below. Using Ekeland's Principle, Caklovic, Li and Willem [3] proved the same result for a Gateaux differentiable functional which is lower semicontinuous. While Goeleven [11] extended the result to the case in which the functional is the sum of a Gateaux differentiable lower semicontinuous function with a convex proper lower semicontinuous function. The same conclusion was also proved by Costa and Silva [7] and Brezis and Nirenberg [2] for C^1 functionals by also employing Ekeland's Principle. In this section we shall show that this is also valid for nondifferentiable functionals. We begin by restating Ekeland's Principle which has been shown to be a powerful tool in solving various nonlinear problems (cf. [9, 1, 10]).

Ekeland's Principle ([8]; cf. also [2]). Let (M, d) be a complete metric space. Let $g : M \to (-\infty, +\infty]$, $\not\equiv +\infty$, be a lower semicontinuous function bounded from below. Then, given $\varepsilon > 0$ and $z_0 \in M$, there exists a point $z \in M$ such that

$$g(z) \le g(x) + \varepsilon d(z, x) \quad \forall x \in M,$$
$$g(z) \le g(z_0) - \varepsilon d(z, z_0).$$

The following is the nonsmooth version of Proposition 1 of Brezis and Nirenberg [2].

Theorem 4. *Suppose $f \in \mathrm{Lip}_{\mathrm{loc}}(X, \mathbb{R})$ and*

$$\alpha := \liminf_{\|x\|\to\infty} f(x)$$

is finite. Then there exists a sequence $(x_n) \subset X$ such that $\|x_n\| \to \infty$, $f(x_n) \to \alpha$, and $\lambda(x_n) \to 0$.

Proof. We follow the ideas of [2]. Set

$$m(r) = \inf_{\|x\|=r} f(x).$$

Then m is a nondecreasing function and $\lim_{r\to\infty} m(r) = \alpha$. For $0 < \varepsilon < 1/2$, choose first $\overline{r} > 1/\varepsilon$ such that

$$\alpha - \varepsilon^2 < m(r) \quad \text{for all } r \ge \overline{r},$$

and then z_0 with $\|z_0\| \geq 2\overline{r}$ such that

$$f(z_0) < m(2\overline{r}) + \varepsilon^2 \leq \alpha + \varepsilon^2.$$

Let $M = \{x \in X : \|x\| \geq 2\overline{r}\}$. By Ekeland's Principle, we get a $z \in M$ such that

$$f(z) \leq f(x) + \varepsilon\|z - x\| \quad \forall x \in M, \tag{3.1}$$

$$f(z) \leq f(z_0) - \varepsilon\|z - z_0\|. \tag{3.2}$$

Since $f(z) \geq m(2\overline{r}) > f(z_0) - \varepsilon^2$, it follows from (3.2) that $\|z - z_0\| < \varepsilon$ and $\|z\| > \overline{r}$. Hence $z \in \operatorname{int} M$. From (3.1), it is easily seen that the function

$$g(x) := f(x) + \varepsilon\|z - x\|, \quad x \in M,$$

assumes its minimum on M at $z \in \operatorname{int} M$; so $0 \in \partial g(z) \subseteq \partial f(z) + \varepsilon B_{X^*}$, which implies $\lambda(z) \leq \varepsilon$. Letting $\varepsilon = \varepsilon_n \downarrow 0$ concludes the proof. ■

Corollary 1. *Suppose $f \in \operatorname{Lip}_{\mathrm{loc}}(X, \mathbb{R})$ is bounded from below and satisfies Chang's condition $(PS)_c$ for all $c \in \mathbb{R}$. Then f is coercive, i.e., $f(x) \to \infty$ as $\|x\| \to \infty$.*

Proof. Suppose the contrary; then $\alpha := \liminf_{\|x\| \to \infty} f(x)$ is finite. By Theorem 4, there exists a sequence (x_n) in X such that $\|x_n\| \to \infty$, $f(x_n) \to \alpha$, and $\lambda(x_n) \to 0$. Then the $(PS)_\alpha$ implies that (x_n) has a convergent subsequence, which clearly leads to a contradiction. ■

We conclude the paper by presenting the nonsmooth version of Proposition 2 of Brezis and Nirenberg [2].

Theorem 5. *Suppose $f \in \operatorname{Lip}_{\mathrm{loc}}(X, \mathbb{R})$ is bounded from below and satisfies Chang's condition $(PS)_c$ for all $c \in \mathbb{R}$. Then every minimizing sequence of f has a convergent subsequence.*

Proof. Suppose (x_n) is a minimizing sequence of f. We may assume that for all integers $n \geq 1$,

$$f(x_n) < \inf_{x \in X} f(x) + \frac{1}{n^2}.$$

Using Ekeland's Principle, there exists $v_n \in X$ such that

$$f(v_n) \leq f(v) + \frac{1}{n}\|v_n - v\|, \quad v \in X,$$

$$f(v_n) \leq f(x_n) - \frac{1}{n}\|v_n - x_n\|.$$

It then follows that $\|v_n - x_n\| < 1/n$ and $0 \in \partial f(x_n) + (1/n)B_{X^*}$. Thus $\lambda(v_n) \leq 1/n$. Now the $(PS)_c$ with $c = \inf_{x\in X} f(x)$ implies that (v_n) (and hence (x_n)) has a convergent subsequence. ∎

References

1. J. P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley-Interscience, New York, 1984.
2. H. Brezis and L. Nirenberg, Remarks on finding critical points, *Comm. Pure Appl. Math.* **44** (1991), 939-963.
3. L. Caklovic, S. Li and M. Willem, A note on Palais-Smale condition and coercivity, *Differential Integral Equations* **3** (1990), 799-800.
4. K. C. Chang, Variational methods for non-differentiable functionals and their applications to partial differential equations, *J. Math. Anal. Appl.* **80** (1981), 102-129.
5. F. H. Clarke, *Optimization and Nonsmooth Analysis*, SIAM, Philadelphia, 1990.
6. D. G. Costa and J. V. A. Goncalves, Critical point theory for nondifferentiable functionals and applications, *J. Math. Anal. Appl.* **153** (1990), 470-485.
7. D. G. Costa and E. A. de B. e Silva, The Palais-Smale condition versus coercivity, *Nonlinear Anal.* **16** (1991), 371-381.
8. I. Ekeland, On the variational principle, *J. Math. Anal. Appl.* **47** (1974), 324-353.
9. I. Ekeland, Nonconvex minimization problems, *Bull. Amer. Math. Soc.* (*N.S.*) **1** (1979), 443-474.
10. D. G. de Figueiredo, *Lectures on the Ekeland Variational Principle with Applications and Detours*, Springer-Verlag, Berlin, 1989.
11. D. Goeleven, A note on Palais-Smale condition in the sense of Szulkin, *Differential Integral Equations* **6** (1993), 1041-1043.
12. S. Li, An existence theorem on multiple critical points and its applications, *Acta Math. Sci.* **4** (1984).
13. M. Willem, *An Introduction to Critical Point Theory*, Lecture Notes, SMR 281/5211, I.C.T.P., Trieste, 1988.

Department of Mathematics, University of Durban-Westville
Private Bag X54001, Durban 4000
South Africa
E-mail: hkxu@pixie.udw.ac.za