PROD. TYPE: FTP
PP: 1-10 (col.fig.: nil)
NA4628 DTD VER: 5.0.1
ED: Jyoti
PAGN: Vish – SCAN: Pradeep

ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT

Nonlinear Analysis ▮▮▮ (▮▮▮) ▮▮▮–▮▮▮

Nonlinear Analysis

www.elsevier.com/locate/na

Saddle critical point analysis and computation☆

Jianxin Zhou*

Texas A&M University, College Station, TX, USA

Abstract

■; ■; ■

Keywords: Multiple saddle points; Numerical computation; Instability analysis

1. Introduction

Let H be a Hilbert space, $\langle , \rangle$ its inner product and $\| \cdot \|$ its norm. Let $J: H \to \mathbb{R}$ be C^1. A point $u^* \in H$ is a *critical point* of J, if u^* solves the Euler–Lagrange equation

$$\nabla J(u) = 0, \tag{1.1}$$

where $\nabla J(u^*)$ is the gradient of J at u^*. A critical point u^* is *nondegenerate* if $J''(u^*)$ is invertible. Otherwise u^* is *degenerate*. The first candidates for a critical point are the local extrema of J, to which the classical critical point theory was devoted in calculus of variation. Traditional numerical algorithms focus on finding such *stable* solutions. Critical points that are not local extrema are called *saddle points*, i.e., for any neighborhood $\mathcal{N}(u^*)$ of u^*, there are $v, w \in \mathcal{N}(u^*)$ s.t. $J(v) < J(u^*) < J(w)$. In physical systems, saddle points appear as *unstable* equilibria or transient excited states, thus very elusive for numerical computation. It presents a great challenge to numerical algorithm design.

Saddle point analysis and computation have many interesting applications. In theoretical chemistry and condensed matter physics, finding a saddle point between two local minima

☆ Supported in part by NSF Grant DMS-0311905.

* Tel.: +1 979 845 2927.
E-mail address: jzhou@math.tamu.edu.

0362-546X/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.na.2005.03.008

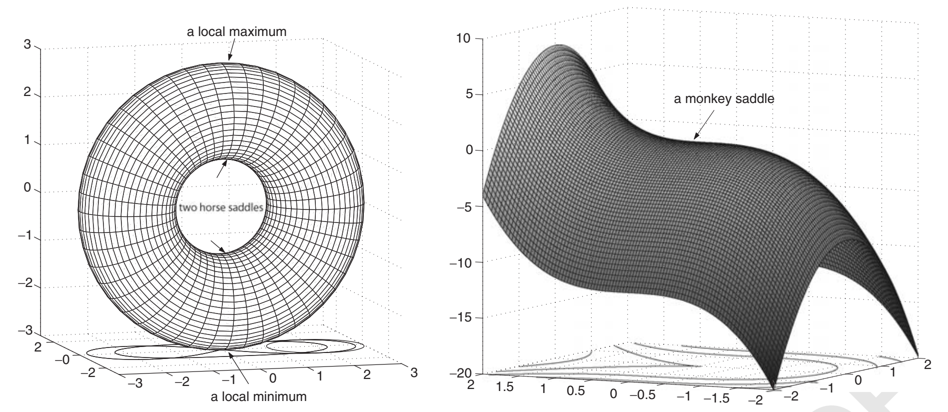

Fig. 1. A local maximum, local minimum and two horse saddle points on the torus(left) and a monkey saddle point (right).

on the potential energy hypersurface is a common and important problem, since it leads to determine the transition state or the minimum energy path between reactant molecules and product molecules [6]. A large literature can be found in this area (Fig. 1).

In computational topology and geometry, people are interested in polygonization of an implicit surface defined by $\{(x, y, z) : F(x, y, z) = c\}$ where $F : \mathbb{R}^3 \to \mathbb{R}$ is a given function and $c \in \mathbb{R}$ is a threshold. From Morse theory, the topology of an implicit surface changes as the critical points are passed. Thus the key point for guaranteeing a topology correct polygonization of an implicit surface is to locate the critical points of its associated function at a given threshold and classify its critical points.

Saddle point analysis and computation have also wide applications in solving nonlinear partial differential equations.

Example 1.1. Consider solving a semilinear elliptic BVP

$$\begin{cases} \Delta u(x) - \lambda u(x) + f(x, u(x)) = 0 & \text{in } \Omega, \\ B_1 u(x) + B_2 \dfrac{\partial u(x)}{\partial n} = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.2}$$

where $\Omega \subset \mathbb{R}^n$ is open and bounded, $\partial\Omega$ its boundary, $\lambda \geqslant 0$, $f(x, u)$ is nonlinear satisfying $f(x, 0) = 0$ and other standard conditions. Let $F(x, t) = \int_0^t f(x, \tau)\, d\tau$. The variational functional is

$$J(u) = \int_\Omega \left\{ -\frac{1}{2}\Delta u(x) \cdot u(x) + \frac{\lambda}{2}|u(x)|^2 - F(x, u(x)) \right\} dx, \tag{1.3}$$

where $H = H_0^1(\Omega)$ if $B_2 = 0$ and $H = H^1(\Omega)$ otherwise. Then it is easy to check that (weak) solutions of (1.2) coincide with critical points of (1.3).

Minimax principle is the most popular approach in critical point theory. However, most minimax results, e.g., the mountain pass, various linking and saddle point theorems [1,2,9–12,14,15], in the literature focus mainly on existence issue. They characterize a saddle point as a solution to a minimax problem

$$\min_{A\in\mathscr{A}}\max_{u\in A} J(u), \tag{1.4}$$

where $\mathscr{A}$ is a collection of compact sets A, both max and min are global. It becomes a two-level global optimization problem and thus is not helpful for algorithm design.

2. A local minimax method

Based on a local minimax characterization, a local minimax algorithm is developed in [7] for computing multiple saddle critical points. Let $L \subset H$ be a closed subspace and $H = L \oplus L^{\perp}$. Denote $S_{L^\perp} = \{u \in L^{\perp}: \|u\| \leqslant 1\}$ and $[L, v] = L \oplus \{v\},\ \forall v \in S_{L^\perp}$.

Definition 2.1. A set-valued mapping $P : S_{L^\perp} \to 2^H$ is called the *peak mapping* of J if $P(v)=\{w \in [L, v]: w=\arg \text{local-max}_{u\in[L,v]} J(u)\},\ \forall v \in S_{L^\perp}$. A mapping $p : S_{L^\perp} \to H$ is a *peak selection* of J if $p(v) \in P(v), \forall v \in S_{L^\perp}$. If p is locally defined, then p is called a local *peak selection* of J.

Remark 2.1. For each $v \in S_{L^\perp}$, $\nabla J(p(v)) \perp [L, v]$ and then $\nabla J(p(v)) \perp p(v)$.

Lemma 2.1 (*Li and Zhou* [7]). *If p is a local peak selection of J near $v_0 \in S_{L^\perp}$ s.t.* (a) *p is continuous at v_0,* (b) *$p(v_0) = t_0 v_0 + v_0^L \notin L$ for some $t_0 \neq 0$, $v_0^L \in L$,* (c) *$\nabla J(p(v_0)) \neq 0$, then there is $s_0 > 0$ s.t. when $0 < s < s_0$,*

$$J(p(v(s))) - J(p(v_0)) < -t_0/2\|\nabla J(p(v_0))\|\,\|v(s) - v_0\|, \tag{2.1}$$

where $v(s) = (v_0 - s\nabla J(p(v_0)))/\|v_0 - s\nabla J(p(v_0))\|$.

Inequality (2.1) defines a stepsize rule in our algorithm and also directly leads to the following local minimax characterization of saddle critical points.

Theorem 2.1 (*Li and Zhou* [7]). *If p is a local peak selection of J near $v_0 \in S_{L^\perp}$ s.t. p is continuous at v_0, $p(v_0) \notin L$ and $v_0 = \arg local \min_{v\in S_{L^\perp}} J(p(v))$, then $u_0 = p(v_0)$ is a saddle point of J.*

Let $\mathscr{M}=\{p(v): v \in S_{L^\perp}\}$. Then local $\min_{u\in\mathscr{M}} J(u)$ yields a saddle point $u^*=p(v^*)$, which can be numerically approximated by, e.g., a steepest descent method. Based on Theorem 2.1, a local minimax algorithm is developed to numerically compute multiple saddle critical points and its convergence results are established in [7,8] where many numerical examples in semilinear elliptic PDE are carried out for multiple solutions.

3. Instability analysis of saddle points

Saddle points are unstable. Can and how do we measure their instability?

A vector $v \in H$ is a *decreasing (increasing) direction* of J at $u^* \in H$ if there is $T > 0$ s.t.

$$J(u^* + tv) < (>) J(u^*) \quad \forall T > t > 0.$$

The maximum dimension of a subspace of decreasing directions of J at u^* is called the *local instability index* (LII) of J at u^*. But such an index lacks of characterization and is too difficult to compute. Let us study other alternatives.

Morse index: If $J''(u^*)$ is a self-adjoint Fredholm operator, then we have $H = H^- \oplus H^0 \oplus H^+$, where H^-, H^+ and H^0 are, respectively, the maximum negative, positive definite, the null subspaces of $J''(u^*)$ in H with $\dim(H^0) < \infty$, and $\dim(H^-)$ is called the *Morse index* (MI) of u^*. If u^* is nondegenerate, i.e., $H^0 = \{0\}$, then $\mathrm{MI}(u^*) = \mathrm{LII}(u^*)$. MI is very expensive to compute, not useful to degenerate cases since many different situations may happen in H^0.

Order of Saddles: Let $u^* \in H$. If $H = H^{\mathrm{I}} \oplus H^{\mathrm{D}}$ for some subspaces H^{I}, H^{D} in H s.t. for each $u_1 \in S_{H^{\mathrm{I}}}$ and $u_2 \in S_{H^{\mathrm{D}}}$, there exist $r_1 > 0$ and $r_2 > 0$ satisfying

$$J(u^* + tu_1) > J(u^*), \quad \forall 0 < |t| \leqslant r_1, \tag{3.1}$$

$$J(u^* + tu_2) < J(u^*), \quad \forall 0 < |t| \leqslant r_2. \tag{3.2}$$

Then u^* is a saddle point of J and $\dim(B^{\mathrm{D}})$ is called the *order* of u^*. This notion is of particular interests in computational chemistry and physics, since the subspace B^{D} corresponds to the reaction coordinates.

Since (3.1) and (3.2) lack of characterization and robustness, we replace (3.2) by

$$J(u^* + tu_2 + o(t)) < J(u^*), \quad \forall 0 < |t| \leqslant r_2, \tag{3.3}$$

where $o(t)$ represents a higher order term. Now if u^* is nondegenerate, we have

$$J(u^* + tu_1) - J(u^*) = \frac{t^2}{2}\langle J''(u^*)u_1, u_1\rangle + o(t^2) > 0,$$

$$J(u^* + tu_2 + o(t)) - J(u^*) = \frac{t^2}{2}\langle J''(u^*)u_2, u_2\rangle + o(t^2) < 0$$

$\forall u_1 \in H^+, u_2 \in H^-$ and $|t|$ small, i.e., $H^{\mathrm{D}} = H^-$, $H^I = H^+$, or $\mathrm{Order}(u^*) = \mathrm{MI}(u^*)$.

But (3.1) and (3.3) do not concern degeneracy, thus more general than MI.

Theorem 3.1 (Zhou [18]). *If p is a local peak selection differentiable at $v^* \in S_{L^\perp}$, $p(v^*) \notin L$ and $v^* = \arg\text{local-min}_{v \in S_{L'}} J(p(v))$, then $u^* = p(v^*)$ is a saddle point with*

$$\dim(L) + 1 = \mathrm{MI}(u^*) + \dim(H^0 \cap \{L, v^*\}). \tag{3.4}$$

Theorem 3.2 (Zhou and Yao [19]). *Let p be a local peak selection differentiable at $v^* \in S_{L'}$, $p(v^*) \notin L$ and $v^* = \arg\text{local-min}_{v \in S_{L'}} J(p(v))$. If the local max with p is strict, then $u^* \equiv p(v^*)$ is a saddle point of order $k = \dim(L) + 1$.*

Remark 3.1. The number $\dim(L)+1$ is known before u^* is computed and called the *local minimax index* (LMI) of u^*. (3.4) implies that LMI is better than MI in measuring local instability of u^*. In the literature, MI or the order of u^* is computed with two steps. First find u^* and then compute the index or order. It is usually very expensive. *Here we reverse the order of the process*, the local minimax method utilizes the geometric and topological structure of LMI to numerically compute a saddle point with such an index.

4. Application to local minimax solutions in control theory

Saddle points and minimax solutions in game and control theory are well-known and have many important applications. Their definitions are actually quite special if compared to their counterparts in critical point theory.

Let X and Y be Hilbert spaces and $H = X \times Y$. A point $u^* = (x^*, y^*) \in H$ is a *saddle point* if $J(x^*, y) \leqslant J(x^*, y^*) \leqslant J(x, y^*)$, $\quad \forall (x, y) \in X \times Y$ or equivalently (x^*, y^*) solves

$$\min_{x\in X} \max_{y\in Y} J(x, y) = \max_{y\in Y} \min_{x\in X} J(x, y). \tag{4.1}$$

It is known that to possess such a saddle point, J needs to satisfy certain convexity in x and concavity in y. Otherwise two solutions to the two sides of (4.1) may not be equal. On the other hand, many applications in control theory require to find a solution to only one side of (4.1). A point $u^* = (x^*, y^*)$ is called a *minimax solution* of $J(x, y)$ if

$$(x^*, y^*) = \arg \min_{x\in X} \max_{y\in Y} J(x, y). \tag{4.2}$$

The significant feature here is that $H = X \times Y$ has a Cartesian product splitting structure. This splitting structure is known and fixed before any solution (x^*, y^*) is found. Interaction between X and Y are not permitted. While such a known splitting structure is in general not available in critical point theory. In order to relate to critical point theory, in this section we assume that J is C^1 and consider *local minimax solutions* of (4.2), i.e., the min and max in (4.2) are in the sense of local neighborhoods. Then the cases become much more complicated. For a given $x \in X$, $J(x, y)$ may have no or multiple local maximum points in y. We need to follow our local minimax approach.

Definition 4.1. A set-valued mapping $P : X \to 2^Y$ is called the *peak selection* of J in Y if for each $x \in X$, $P(x)$ is the set of all local maximum of $J(x, y)$ in y. A selection p of P is called a peak selection of J in Y. If a peak selection p is *locally* defined at $x_0 \in X$, then p is called a local peak selection of J at x_0.

Remark 4.1. By the definition, it is clear that $J_y(x, p(x)) = 0 \quad \forall x \in X$ and $p(x)$ is defined where J_y denotes the partial derivative of J with respect to y.

Theorem 4.1. *Let p be a local peak selection of J at $x_0 \in X$ s.t. p is continuous at x_0. If $J_x(x_0, p(x_0)) \neq 0$, then there exists $s_0 > 0$ s.t. when $0 < s < s_0$*

$$J(x(s), p(x(s))) < J(x_0, p(x_0)) - \frac{s}{2}\|J_x(x_0, p(x_0))\|^2, \tag{4.3}$$

where $x(s) = x_0 - sJ_x(x_0, p(x_0))$.

Proof. Since J is C^1, $x(s) \to x_0$ and $p(x(s)) \to p(x_0)$ as $s \to 0$, if we denote $z(s) = (x(s), p(x(s))) - (x_0, p(x_0))$, we have

$$\begin{aligned} J(x(s), p(x(s))) &= J(x_0, p(x_0)) + \langle \nabla J(x_0, p(x_0)), z(s)\rangle_H + o(\|z(s)\|) \\ &= J(x_0, p(x_0)) + \langle J_x(x_0, p(x_0)), x(s) - x_0\rangle_X + o(\|z(s)\|) \\ &= J(x_0, p(x_0)) - s\|J_x(x_0, p(x_0))\|^2 + o(\|z(s)\|) \\ &< J(x_0, p(x_0)) - \frac{s}{2}\|J_x(x_0, p(x_0))\|^2, \quad \text{as } 0 < s < s_0, \end{aligned}$$

where the last inequality holds because $z(s) \to 0$ as $s \to 0$ and $-s\|J_x(x_0, p(x_0))\|^2 < 0$. □

Then it is direct to relate a local minimax solution to a saddle critical point.

Theorem 4.2. *Let p be a local peak selection of J at $x^* \in X$ s.t. p is continuous at x^*. If $x^* = \arg\min_{x\in X} J(x, p(x))$, then $(x^*, p(x^*))$ is a saddle critical point of J, i.e.,* $\nabla J(x^*, p(x^*)) = 0$.

Now let p be a continuous peak selection of J and denote $G(x) \equiv J(x, p(x)), \quad \forall x \in X$. Can we develop a first order method to find a local minimum point of G? To do so, usually G need be C^1. But under our assumption G is only continuous. However due to its special structure, in this section, we develop a steepest descent method to find a local minimum point of G and prove its convergence. The algorithm is described by the following six steps.

Step 1: Given $\varepsilon > 0$, $\lambda > 0$. Let $x_0 \in X$ be an initial guess for a solution x^* and y_0 be an initial guess for $p(x_0)$. Let $k = 0$;

Step 2: Using the initial guess y_k to solve for $p(x_k) = \arg\max_{y\in Y} J(x_k, y)$;

Step 3: Denote $y_k = p(x_k)$, compute the steepest descent vector $d_k = -J_x(x_k, y_k)$;

Step 4: If $\|d_k\| \leqslant \varepsilon$ then output (x_k, y_k), stop; else go to Step 5;

Step 5: Set $x_k(s) = x_k - sd_k$ and find

$$\begin{aligned} s_k = \max_{m\in\mathbb{N}} \Bigg\{ \frac{\lambda}{2^m} \Bigg| \; & 2^m > \|d_k\|, \\ & J\left(x_k\left(\frac{\lambda}{2^m}\right), p\left(x_k\left(\frac{\lambda}{2^m}\right)\right)\right) - J(x_k, y_k) \leqslant -\frac{1}{2}\|d_k\| \left\| x_k\left(\frac{\lambda}{2^m}\right) - x_k \right\| \Bigg\}, \end{aligned}$$

where initial guess $y = y_k$ is used to find $p(x_k(\lambda/2^m))$.

Step 6: Set $x_{k+1} = x_k(s_k)$ and update $k = k + 1$ then go to Step 2.

Due to multiplicity of solutions, convergence analysis of the algorithm becomes more complicated and it is important to use initial guess stated in the algorithm in order to consistently trace and keep in the same peak selection p and to make p continuous. For $x \in X$, $d = -J_x(x, p(x)) \neq 0$, define $x(s) = x + sd$ and

$$s(x) = \max_{\lambda \geqslant s > 0} \left\{ s | \lambda \geqslant s\|d\|, J(x(s), p(x(s))) - J(x, p(x)) \leqslant -\tfrac{1}{2}\|d\| \, \|x(s) - x\| \right\}.$$

Then it is clear that

$$\tfrac{1}{2}s(x_k) \leqslant s_k \leqslant s(x_k).$$

Theorem 4.3. *If $x_0 \in X$, $J_x(x_0, p(x_0)) \neq 0$, then there are $s_0 > 0$ and a neighborhood $\mathscr{N}(x_0)$ of x_0 s.t. $s(x) > s_0$, $\forall x \in \mathscr{N}(x_0)$.*

Proof. There is $s_1 > 0$ s.t. when $s_1 > s > 0$, we have $\lambda > s\|J_x(x_0(s), p(x_0(s)))\|$ and

$$J(x_0(s), p(x_0(s))) - J(x_0, p(x_0)) < -\tfrac{1}{2}s\|J_x(x_0, p(x_0))\|^2$$

or

$$\frac{J(x_0(s), p(x_0(s))) - J(x_0, p(x_0))}{s\|J_x(x_0, p(x_0))\|^2} < -\frac{1}{2}.$$

Fix $s_0 : s_1 > s_0 > 0$ and let x_0 vary, the term

$$\frac{J(x(s_0), p(x(s_0))) - J(x, p(x))}{s_0\|J_x(x, p(x))\|^2}$$

is continuous in x. Thus there is $\mathscr{N}(x_0)$ s.t. for each $x \in \mathscr{N}(x_0)$, we have

$$\frac{J(x(s_0), p(x(s_0))) - J(x, p(x))}{s_0\|J_x(x, p(x))\|^2} < -\frac{1}{2},$$

i.e., $s(x) > s_0$, $\forall x \in \mathscr{N}(x_0)$. □

Theorem 4.4. *If $\lim_{k\to\infty} J(x_k, p(x_k)) > -\infty$ where $\{x_k\}$ is the sequence generated by the algorithm, then there exists a subsequence $\{x_{k_i}\} \subset \{x_k\}$ s.t. $\nabla J(x_{k_i}, p(x_{k_i})) \to 0$ as $i \to \infty$.*

Proof. Suppose not. Then there is $\delta > 0$ s.t. $\|J_x(x_k, p(x_k))\|_X = \|\nabla J(x_k, p(x_k))\|_H > \delta$ when k is large, say $k > k_0$. By our stepsize rule, we have

$$\begin{aligned} J(x_{k+1}, p(x_{k+1})) - J(x_k, p(x_k)) &\leqslant -\tfrac{1}{2}s(x_k)\|J_x(x_k, p(x_k))\|^2 \\ &= -\tfrac{1}{2}\|x_{k+1} - x_k\|\|J_x(x_k, p(x_k))\|. \end{aligned}$$

Summarizing two sides from k_0 to ∞, we have

$$\begin{aligned} -\infty < \lim_{k\to\infty} J(x_k, p(x_k)) - J(x_{k_0}, p(x_{k_0})) &= \sum_{k=k_0}^{\infty} J(x_{k+1}, p(x_{k+1})) - J(x_k, p(x_k)) \\ &\leqslant -\frac{1}{2}\sum_{k=k_0}^{\infty} \|x_{k+1} - x_k\|\|J_x(x_k, p(x_k))\| \leqslant -\frac{1}{2}\delta\sum_{k=k_0}^{\infty}\|x_{k+1} - x_k\|, \end{aligned}$$

which implies that $\{x_k\}$ is a Cauchy sequence. There is $x^* \in X$ s.t. $x_k \to x^*$ as $k \to \infty$. Since $J_x(x, p(x))$ is continuous in x, we have $\|\nabla J(x^*, p(x^*))\|_H = \|J_x(x^*, p(x^*))\|_X \geqslant \delta$. By Theorem 4.3, there are $s_0 > 0$ and $\mathscr{N}(x^*)$ s.t. $s(x) \geqslant s_0$, $\forall x \in \mathscr{N}(x^*)$. This implies that when k is large, $x_k \in \mathscr{N}(x^*)$ and $\|x_{k+1} - x_k\| = s_k\|J_x(x_k, p(x_k))\| \geqslant \frac{1}{2}s_0\delta$. Then $\{x_k\}$ is not a convergent sequence. It leads to a contradiction. □

Definition 4.2. A C^1 function $J : X \times Y \to \mathbb{R}$ is said to satisfy PS (Palais–Smale) condition if any sequence $\{(x_n, y_n)\} \subset X \times Y$ s.t. $J(x_n, y_n)$ is bounded and $\nabla J(x_n, y_n) \to 0$ has a convergent subsequence.

From now on we assume J satisfies PS condition. Note that $J_y(x, p(x)) = 0$. If $J(x, p(x))$ is bounded from below where p is a continuous peak selection of J in Y, then Ekeland's variational principle guarantees the existence of a solution x^* to

$$\min_{x \in X} G(x) = \min_{x \in X} \max_{y \in Y} J(x, y).$$

Let $K = \{x \in X : J_x(x, p(x)) = 0\}$ and $K_c = \{x \in K : J(x, p(x)) = c\}$. Then by PS condition, the set K_c is always compact. By adopting a similar approach in [16], we prove the following point-to-set convergence result.

Theorem 4.5. *Given two open sets* $\emptyset \neq V_2 \subset V_1 \subset X$. *Assume* (a) $V_1 \cap K \subset K_c$ *where* $c = \inf_{x \in V_1} J(x, p(x)) > -\infty$ *and* (b) *there is* $d > 0$ *s.t.*

$$\inf\{J(x, p(x)) : x \in V_1, \operatorname{dis}(x, \partial V_1) \leqslant d\} = a > b = \sup\{J(x, p(x)) : x \in V_2\} > c.$$

Let $\lambda < d$ *be given in the algorithm and* $\{x_n\}$ *be the sequence generated by the algorithm started from an initial guess* $x_0 \in V_2$. *Then* $\forall \varepsilon > 0$, *there is* $N > 0$ *s.t.*

$$\operatorname{dis}(x_n, K_c) < \varepsilon, \quad \forall n > N.$$

Proof. Let $G(x) \equiv J(x, p(x))$. Since $\lambda < d$, by our stepsize rule, if $x_n \in V_1$ and $\operatorname{dis}(x_n, \partial V_1) > d$, then $x_{n+1} \in V_1$. By the monotone decreasing nature of our algorithm and (b), $x_0 \in V_2$ implies $G(x_1) < b$, $x_1 \in V_1$. Thus $x_2 \in V_1$. Then $G(x_n) < b$ and $x_n \in V_1$ imply $\operatorname{dis}(x_n, \partial V_1) > d$ and $x_{n+1} \in V_1$, i.e. $\{x_n\} \subset V_1$. Theorem 4.4 states that $\{x_n\}$ has a subsequence that converges to $x^* = \arg\min_{x \in X} G(x)$. By (a) and Theorem 4.2, we must have $x^* \in \bar{V}_1 \cap K_c \neq \emptyset$. By the monotone decreasing nature of our algorithm, we have $\lim_{n \to \infty} G(x_n) = c$. Denote

$$\bar{G}(x) = \begin{cases} G(x), & x \in \bar{V}_1, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is clear that $\bar{G}$ is lower semi-continuous on X. By Ekeland's variational principle, for each $\delta_n = (G(x_n) - c)^{1/2}$, there is $\bar{x}_n \in X$ s.t.

$$\bar{G}(\bar{x}_n) - \bar{G}(x) \leqslant \delta_n \|\bar{x}_n - x\| \quad \forall x \in X,$$
$$\bar{G}(\bar{x}_n) - G(x_n) \leqslant -\delta_n \|\bar{x}_n - x_n\|.$$

It is clear that $\bar{x}_n \in \bar{V}_1$. Then we have

$$G(\bar{x}_n) - G(x) \leqslant \delta_n \|\bar{x}_n - x\| \quad \forall x \in \bar{V}_1,$$
$$G(\bar{x}_n) - G(x_n) \leqslant -\delta_n \|\bar{x}_n - x_n\|. \tag{4.4}$$

It follows $c \leqslant G(\bar{x}_n) \leqslant G(x_n) - \delta_n \|\bar{x}_n - x_n\|$ or $\|\bar{x}_n - x_n\| \leqslant \delta_n^{\frac{1}{2}}$. Then $G(x_n) \to c$ implies $\delta_n \to 0$ and $G(\bar{x}_n) \to c$. For those large n with $J_x(x_n, p(x_n)) \neq 0$, by Theorem 4.1, when $s > 0$ is small, we have

$$G(\bar{x}_n(s)) - G(\bar{x}_n) \leqslant -\tfrac{1}{2} \|\bar{x}_n(s) - \bar{x}_n\| \|J_x(\bar{x}, p(\bar{x}_n))\|.$$

Hence by (4.4)

$$\|\nabla J(\bar{x}_n, p(\bar{x}_n))\|_H = \|J_x(\bar{x}_n, p(\bar{x}_n))\|_X \leqslant 2\delta_n \to 0,$$

which also implies

$$\|\nabla J(x_n, p(x_n))\| \to 0. \tag{4.5}$$

Let β be any limit point of $\{\text{dis}(x_n, K_c)\}$ and $\{x_{n_k}\} \subset \{x_n\}$ be a subsequence s.t. $\beta = \lim_{k\to\infty} \text{dis}(x_{n_k}, K_c)$. Since $J(x_{n_k}, p(x_{n_k})) \to c$ and $\nabla J(x_{n_k}, p(x_{n_k})) \to 0$ as $k \to \infty$, by PS condition, $\{x_{n_k}\}$ has a subsequence, denoted by $\{x_{n_k}\}$ again, s.t. $x_{n_k} \to x^* \in \bar{V}_1$. Then (4.5) leads to $x^* \in K_c$ or $\beta = 0$. $\square$

Condition (b) in Theorem 4.5 is similar to those in linking theorems in critical point theory to form a trap. Theorem 4.5 directly leads to a usual point-to-point convergent result if $V_1 \cap K_c$ contains only one point.

We may now ask the question: Let $x_k \to x^*$. Then $y^* = \lim_{k\to\infty} p(x_k)$ need not be a local maximum of $J(x^*, y)$ in y. How can we check if p is continuous? We need to adopt an approach developed in [17].

4.1. A local min-⊥ method

Based on Remark 4.1, we introduce

Definition 4.3. A set-valued mapping $P: X \to 2^Y$ is called the Y-⊥ mapping of *J* if $P(x) = \{y \in Y: J_y(x, y) = 0\} \quad \forall x \in X$. A selection p is called a *Y*-⊥ selection of *J*. If p is locally defined then p is a local *Y*-⊥ selection.

It is clear that any peak selection of *J* is a *Y*-⊥ selection of *J*. So we assume that p is a *Y*-⊥ selection of *J*. Then all previous results remain true. It can be directly checked that

Lemma 4.1. *If J is C^1, then the graph $\mathbb{G} = \{(x, y): x \in X, y \in P(x) \neq \emptyset\} = \{(x, y): J_y(x, y) = 0\}$ is closed.*

We can embed a local min–max problem into a min-Y-⊥ problem and use the following method to numerically check the continuity (smoothness) of a peak selection p.

When $Y = [y_1, y_2, \ldots, y_n]$ and $x^* \in X$. Then the implicit function theorem can be used to check if p is differentiable or not at x^*. By the definition of p, $y^* = p(x^*) = \sum_{i=1}^n t_i y_i$ where $(t_1, \ldots, t_n)$ is solved locally from n orthogonal conditions,

$$F_j(x^*, t_1, \ldots, t_n) \equiv \left\langle J_y\left(x^*, \sum_{k=1}^n t_k y_k\right), y_j\right\rangle = 0, \quad j = 1, \ldots, n.$$

Then we have $n \times n$ terms

$$\frac{\partial F_j}{\partial t_i} = \left\langle J_{yy}\left(x^*, \sum_{k=1}^n t_k y_k\right) y_i, y_j\right\rangle, \quad i, j = 1, \ldots, n.$$

By the implicit function theorem, if the $n \times n$ matrix

$$Q''(x^*) = \left[\left\langle J_{yy}\left(x^*, \sum_{k=1}^{n} t_k y_k\right) y_i, y_j \right\rangle \right], \quad i, j = 1, \ldots, n$$

is invertible or $|Q''(x^*)| \neq 0$, then p is differentiable at and near x^*. This condition can be easily and numerically checked.

5. Uncited references

[3–5,13].

References

[1] A. Ambrosetti, P. Rabinowitz, Dual variational methods in critical point theory and applications, J. Funct. Anal. 14 (1973) 349–381.
[2] K.C. Chang, Infinite Dimensional Morse Theory and Multiple Solution Problems, Birkhäuser, Boston, 1993.
[3] Y.S. Choi, P.J. McKenna, A mountain pass method for the numerical solution of semilinear elliptic problems, Nonlinear Anal. 20 (1993) 417–437.
[4] Z. Ding, D. Costa, G. Chen, A high linking method for sign changing solutions for semilinear elliptic equations, Nonlinear Anal. 38 (1999) 151–172.
[5] W.Y. Ding, W.M. Ni, On the existence of positive entire solutions of a semilinear elliptic equation, Arch. Rat. Mech. Anal. 91 (1986) 238–308.
[6] G. Henkelman, G. Johannesson, H. Jonsson, Methods for finding saddle points and minimum energy paths, in: D. Schwartz (Ed.), Computer Chemistry, vol. 5, Kluwer, Dordrecht, 2000.
[7] Y. Li, J. Zhou, A minimax method for finding multiple critical points and its applications to nonlinear PDEs, SIAM Sci. Comput. 23 (2001) 840–865.
[8] Y. Li, J. Zhou, Convergence results of a local minimax method for finding multiple critical points, SIAM Sci. Comput. 24 (2002) 840–865.
[9] F. Lin, T. Lin, Minimax solutions of the Ginzburg–Landau equations, Slecta Math. (N.S) 3 (1) (1997) 99–123.
[10] J. Mawhin, M. Willem, Critical Point Theory and Hamiltonian Systems, Springer, New York, 1989.
[11] P. Rabinowitz, Minimax method in critical point theory with applications to differential equations, CBMS Regional Conference Series in Mathematics, No. 65, AMS, Providence, RI, 1986.
[12] M. Schechter, Linking Methods in Critical Point Theory, Birkhauser, Boston, 1999.
[13] M. Struwe, Variational Methods, Springer, New York, 1996.
[14] Z.-Q. Wang, On a superlinear elliptic equation, Ann. Inst. Henri Poincare 8 (1991) 43–57.
[15] M. Willem, Minimax Theorems, Birkhauser, Boston, 1996.
[16] X. Yao, J. Zhou, A unified convergence result on a minimax method for finding multiple critical points in Banach spaces, preprint.
[17] J. Zhou, A local min-orthogonal method for finding multiple saddle points, J. Math. Anal. Appl. 291 (2004) 66–81.
[18] J. Zhou, Instability analysis of saddle points by a local minimax method, Math. Comput. electronically published on 7/20/04.
[19] J. Zhou, X. Yao, Stabilization in computing saddle points, Control and Boundary Analysis, Marcel Dekker, New York, 2004.