

A Minimax Method for Finding Multiple Critical Points in Banach Spaces and Its Application to Quasilinear Elliptic PDE*

Xudong Yao[†] and Jianxin Zhou[†]

Abstract

A minimax method for finding multiple critical points in Banach spaces is developed by using a modified pseudo-gradient. The method is implemented successfully to solve several quasilinear elliptic boundary value problems for multiple solutions. Numerical solutions are presented by their profiles to illustrate the theory and method. A unified convergence result of the algorithm is established in a subsequent paper [21].

Keywords. Multiple critical points, Banach space, Pseudo-gradient, p-Laplacian, Quasilinear elliptic PDE

AMS(MOS) subject classifications. 58E05, 58E30,35A40,35J65

Abbreviated titles. A Minimax Method in Banach Spaces

1 Introduction

Many nonlinear boundary value problems (BVP) are equivalent to solving

$$(1.1) \quad Au = 0$$

where $A : X \rightarrow Y$ is a nonlinear operator between two Banach spaces. Denote X^* the topological dual of X and $\langle \cdot \rangle$ the dual relation. When a problem is variational, there exists $J \in C^1(X, R)$ s.t. $A = \nabla J$, where ∇J is the gradient of J , i.e., for each $u \in X$, $\nabla J(u) \in X^*$

*This research project is supported in part by NSF DMS-0311905

[†]Math. Dept., Texas A&M University, College Station, TX 77843 (xdyao@math.tamu.edu, jzhou@math.tamu.edu)

s.t. $\langle \nabla J(u), v \rangle = \frac{d}{dt} J(u + tv)|_{t=0}$, $\forall v \in X$. Then (1.1) is equivalent to solving the Euler-Lagrange equation

$$(1.2) \quad \nabla J(u) = 0 \quad \text{or} \quad \langle \nabla J(u), v \rangle = 0, \quad \forall v \in X.$$

A solution u^* to (1.2) is called a *critical point* of J and its value $c = J(u^*)$ is called a *critical value* of J , and the set $J^{-1}(c)$ is called a *critical level* of J . Multiple critical points with different performance exist in many nonlinear problems in applications [9,15,16,17,20]. The first candidates for critical points are the local extrema to which the classical critical point theory was devoted in calculus of variation. Traditional numerical methods focus on finding such stable solutions. Critical points that are not local extremes are called *saddle points*. In physical systems, saddle points appear as *unstable* equilibria or transient excited states. To theoretical and computational physics and chemistry, saddle points between two stable states on the potential hypersurface are of great interests and are the theme of so called Transition State Theory or Activated Complex Theory, as they correspond to the transition states and lead to the minimum energy paths between reactant molecules and product molecules [8]. Although a vast literature on algorithms for computing transition states can be found in computational physics and chemistry, etc., most of them lacked of proper mathematical justification and were just for certain special cases in low finite-dimension spaces, e.g., a transition state between two local minima [8].

How to numerically find saddle points in a stable way is interesting to both theory and applications. However it is a challenging problem due to the unstable nature, saddle points are very elusive to capture numerically.

Minimax principle, which characterizes a saddle point as a solution to $\min_{A \in \mathcal{A}} \max_{v \in A} J(v)$ for some collection \mathcal{A} of subsets A in X , is one of the most popular approaches in critical point theory. Since the mountain pass lemma proved in 1973 [1] set a milestone for modern nonlinear analysis, many minimax theorems, such as various linking and saddle point theorems, have been successfully established to prove the existence of multiple solutions to numerous nonlinear problems [1,2,5,12,13,14,17,18,20, etc.]. But most of them focus mainly on the existence issue and require one to solve a two-level *global* minimax problem, and therefore are not useful for algorithm implementation.

The first numerical minimax algorithm for finding critical points basically with $MI=1$,¹ was developed by Choi-McKenna in 1993. Ding-Costa-Chen proposed a numerical minimax method in 1999 to capture critical points basically with $MI=2$. But no mathematical justification or convergence of the algorithms was established.

A numerical local minimax algorithm together with its mathematical justification and convergence was successfully developed by Li-Zhou in [10,11], to find multiple saddle points of $MI=1,2,\dots,n$. All those three algorithms are formulated in Hilbert spaces, where the gradient and orthogonality played important roles. As a matter of fact, the gradient is used as a search direction to update an approximation point and the orthogonality is used to prevent a search from degenerating to a lower critical level. In terms of minimax approach, at the first level, A is a 1D simplex in Choi-McKenna's method, a 2D simplex in Ding-Costa-Chen's method and an n D subspace in Li-Zhou's method.

On the other hand, many nonlinear problems in application, such as the wellknown nonlinear p -Laplacian equation in the study of non-Newtonian fluid flows [4,9,17], have to be formulated in Banach spaces. How to find multiple critical points in Banach spaces? So far no such numerical methods are available in the literature. In this paper, we develop a numerical minimax method for this purpose.

Since the local minimax characterization of saddle points in a Hilbert space in [10,11] serves as a mathematical justification of the local minimax method of Li-Zhou, it provides us with a theoretical foundation to work on. Let us closely examine Li-Zhou's method. We can see that the maximization at the first level is to reach a desired orthogonality, i.e., $\nabla J(u)$ is orthogonal to the subspace generated by previously found solutions, and the negative gradient is used to search for a minimum at the second level. When mathematical justification of a minimax method in a Banach space is concerned, there are two basic theoretical difficulties to overcome. The first is related to the gradient. Since, in general, a Banach space X is different from its topological dual X^* , the gradient $\nabla J(u)$ is in X^* not X . Thus it cannot be used at the second level as a search direction to update a point in X to approximate a local minimum. To overcome this difficulty, a *pseudo-gradient* can be

¹When $X = H$ is a Hilbert space and J'' exists at a critical point u^* , the Morse index (MI) of u^* is the maximum dimension of a subspace H^- of H on which $J''(u^*)$ is negative definite. MI is not defined in Banach spaces.

used to replace the gradient as a search direction for a local minimum at the second level. Then the second basic difficulty at the first level takes place, i.e., the orthogonality between a search direction and the subspace generated by previously found solutions is broken. We propose to use a modified pseudo-gradient to overcome this difficulty. Those are done in Section 2. Once a local minimax characterization of saddle points in a Banach space is established, its numerical implementation will be developed in Section 3. In Section 4, we discuss implementation of computing a pseudo-gradient. In Section 5, we carry out several numerical examples to illustrate the new theory and its applications.

2 Local Min-Max Theorems in Banach Spaces

For a subspace $X' \subseteq X$, denote $S_{X'} = \{v | v \in X', \|v\| = 1\}$ -the unit sphere in X' . Assume that $X = L \oplus L'$, where L (called a support) and L' are closed subspaces of X , and $\mathcal{P} : X \rightarrow L'$ is the corresponding linear projection with a bound $M \geq 1$ (see Theorem 5.6 in [7]).

DEFINITION 2.1. *A set-valued mapping $P : S_{L'} \rightarrow 2^X$ is the peak mapping of J w.r.t. L if $\forall v \in S_{L'}, P(v)$ is the set of all local maximum points of J in the subspace $[L, v] = \{tv + w | w \in L, t \in \mathbb{R}\}$. A single-valued mapping $p : S_{L'} \rightarrow X$ is a peak selection of J w.r.t. L if*

$$p(v) \in P(v), \quad \forall v \in S_{L'}.$$

For a given $v \in S_{L'}$, p is said to be a local peak selection of J at v w.r.t. L if p is a peak selection of J locally defined near v .

DEFINITION 2.2. *Let $u \in X$ be a point s.t. $\nabla J(u) \neq 0$. For given $\theta \in (0, 1]$, a point $\Psi(u) \in X$ is a pseudo-gradient of J at u w.r.t. θ if*

$$(2.1) \quad \|\Psi(u)\| \leq 1, \quad \langle \nabla J(u), \Psi(u) \rangle \geq \theta \|\nabla J(u)\|.$$

Denote $\hat{X} = \{u \in X : \nabla J(u) \neq 0\}$. A pseudo-gradient flow of J with a constant θ is a continuous mapping $\Psi : \hat{X} \rightarrow X$ s.t. $\forall u \in \hat{X}$, $\Psi(u)$ satisfies (2.1).

Lemma 2.10.1 in [16] guarantees the existence of a pseudo-gradient flow.

REMARK 2.1. Note that the number 1 in (2.1) can be replaced by any number $m \geq 1$, since it can be absorbed by the constant θ to become $0 < \frac{\theta}{m} \leq 1$.

Pseudo-gradients have been used in the literature to search for a local minimum of a C^1 functional in a Banach space. However, when a saddle point is concerned, such a pseudo-gradient does not work, since it always leads to a local minimum. To prevent this degeneracy from a saddle point (higher critical level) to a local minimum (lower critical level), we need to modify the notion of a pseudo-gradient. Due to the definition of a peak selection, we are able to construct a special pseudo-gradient whose role in a local minimax method is similar to the gradient in a Hilbert space. Such a pseudo-gradient will be used to replace the gradient to design a local minimax method for finding a saddle point in a Banach space.

LEMMA 2.1. *Let $0 < \theta < 1$ be given. For $v_0 \in S_{L'}$, if p is a local peak selection of J w.r.t. L at v_0 s.t. $\nabla J(p(v_0)) \neq 0$ and $\Psi(p(v_0)) \in X$ is a pseudo-gradient of J at $p(v_0)$ w.r.t. the constant θ , then there exists a (modified) pseudo-gradient $G(p(v_0))$ of J at $p(v_0)$ w.r.t. the constant θ s.t.*

(a) $G(p(v_0)) \in L'$, $0 < \|G(p(v_0))\| \leq M$ where $M \geq 1$ is the bound of the linear projection \mathcal{P} from X to L' ;

(b) $\langle \nabla J(p(v_0)), G(p(v_0)) \rangle = \langle \nabla J(p(v_0)), \Psi(p(v_0)) \rangle$;

(c) *If $\Psi(p(v_0))$ is the value of a pseudo-gradient flow $\Psi(\cdot)$ of J at $p(v_0)$, then $G(\cdot)$ is continuous and $G(p(v_0))$ is called the value of a modified pseudo-gradient flow of J at $p(v_0)$.*

Proof. Let $G(p(v_0)) = \mathcal{P}(\Psi(p(v_0))) \in L'$. Then $\|G(p(v_0))\| \leq M\|\Psi(p(v_0))\| \leq M$. Denote $\Psi(p(v_0)) = \Psi_L(p(v_0)) + G(p(v_0))$ for some vector $\Psi_L(p(v_0)) \in L$. By the definition of a peak selection p , we have $\langle \nabla J(p(v_0)), \Psi_L(p(v_0)) \rangle = 0$. Thus

$$\langle \nabla J(p(v_0)), G(p(v_0)) \rangle = \langle \nabla J(p(v_0)), \Psi(p(v_0)) \rangle \geq \theta \|\nabla J(p(v_0))\| > 0.$$

Therefore $G(p(v_0)) \neq 0$ is a pseudo-gradient of J at $p(v_0)$ w.r.t. θ . The results follow. ■

The following inequality will be used in later development.

LEMMA 2.2. *For each $v \in X$ with $\|v\| = 1$, it holds*

$$\left\| v - \frac{v - w}{\|v - w\|} \right\| \leq \frac{2\|w\|}{\|v - w\|}, \quad \forall w \in X.$$

Proof. In fact,

$$\begin{aligned} \left\| v - \frac{v-w}{\|v-w\|} \right\| &= \frac{\|v(\|v-w\| - 1) + w\|}{\|v-w\|} \leq \frac{\|v\| |\|v-w\| - 1| + \|w\|}{\|v-w\|} \\ &= \frac{|\|v-w\| - \|v\|| + \|w\|}{\|v-w\|} \leq \frac{2\|w\|}{\|v-w\|}. \end{aligned}$$

The next lemma is crucial in this paper, which shows the relation between the gradient of J and the variation of a peak selection. It will be used to establish a local minimax characterization of saddle points and to design a stepsize rule in a local minimax algorithm.

LEMMA 2.3. *For $v_0 \in S_{L'}$, if there is a local peak selection p of J w.r.t. L at v_0 satisfying (1) p is continuous at v_0 , (2) $d(p(v_0), L) > 0$ and (3) $\nabla J(p(v_0)) \neq 0$, then there exists $s_0 > 0$ s.t. as $0 < s < s_0$*

$$(2.2) \quad J(p(v_s)) - J(p(v_0)) < -\frac{s\theta|t_s|\|\nabla J(p(v_0))\|}{4} \leq -\frac{\theta d(p(v_0), L)\|\nabla J(p(v_0))\|}{8M}\|v_s - v_0\|$$

where $p(v_0) = t_0 v_0 + w_0$, $p(v_s) = t_s v_s + w_s$ with $t_0, t_s \neq 0$ and $w_0, w_s \in L$,

$$v_s = \frac{v_0 - \text{sign}(t_0)sG(p(v_0))}{\|v_0 - \text{sign}(t_0)sG(p(v_0))\|}$$

and $G(p(v_0))$ is a modified pseudo-gradient of J at $p(v_0)$ as defined in Lemma 2.1.

Proof. Since $J \in C^1(X, R)$, we have

$$(2.3) \quad J(p(v_s)) = J(p(v_0)) + \langle \nabla J(p(v_0)), p(v_s) - p(v_0) \rangle + o(\|p(v_s) - p(v_0)\|).$$

Since p is a peak selection, we have $\langle \nabla J(p(v_0)), v_0 \rangle = \langle \nabla J(p(v_0)), v \rangle = 0$, $\forall v \in L$. Thus

$$\begin{aligned} \langle \nabla J(p(v_0)), p(v_s) - p(v_0) \rangle &= t_s \langle \nabla J(p(v_0)), v_s \rangle \\ &= -\frac{\text{sign}(t_0)t_s s \langle \nabla J(p(v_0)), G(p(v_0)) \rangle}{\|v_0 - \text{sign}(t_0)sG(p(v_0))\|} = -\frac{\text{sign}(t_0)t_s s \langle \nabla J(p(v_0)), \Psi(p(v_0)) \rangle}{\|v_0 - \text{sign}(t_0)sG(p(v_0))\|} \end{aligned}$$

by Lemma 2.1. When p is continuous at v_0 and $X = L \oplus L'$, we have $t_s \rightarrow t_0$ and $w_s \rightarrow w_0$ as $s \rightarrow 0$. Then, by the definition of a pseudo-gradient, as $s > 0$ is small

$$(2.4) \quad \langle \nabla J(p(v_0)), p(v_s) - p(v_0) \rangle \leq -\frac{s\theta|t_s|}{\|v_0 - \text{sign}(t_0)sG(p(v_0))\|}\|\nabla J(p(v_0))\|.$$

Hence, by (2.3) and (2.4), there is $s_0 > 0$ s.t. as $0 < s < s_0$,

$$(2.5) \quad J(p(v_s)) - J(p(v_0)) < -\frac{s\theta|t_s|\|\nabla J(p(v_0))\|}{2\|v_0 - \text{sign}(t_0)sG(p(v_0))\|} < -\frac{s\theta d(p(v_0), L)\|\nabla J(p(v_0))\|}{4\|v_0 - \text{sign}(t_0)sG(p(v_0))\|}.$$

Choose $s > 0$ small such that $\|v_0 - \text{sign}(t_0)sG(p(v_0))\| \leq 2$. Then

$$(2.6) \quad J(p(v_s)) - J(p(v_0)) < -\frac{s\theta|t_s|\|\nabla J(p(v_0))\|}{4} < -\frac{s\theta d(p(v_0), L)\|\nabla J(p(v_0))\|}{8}.$$

On the other hand, by Lemma 2.2

$$\begin{aligned} \|v_s - v_0\| &= \left\| \frac{v_0 - \text{sign}(t_0)sG(p(v_0))}{\|v_0 - \text{sign}(t_0)sG(p(v_0))\|} - v_0 \right\| \\ &\leq \frac{2\|\text{sign}(t_0)sG(p(v_0))\|}{\|v_0 - \text{sign}(t_0)sG(p(v_0))\|} \leq \frac{2sM}{\|v_0 - \text{sign}(t_0)sG(p(v_0))\|}, \end{aligned}$$

where the last inequality holds due to the fact that

$$\|G(p(v_0))\| = \|\mathcal{P}(\Psi(p(v_0)))\| \leq M\|\Psi(p(v_0))\| \leq M.$$

We have

$$-s \leq -\frac{\|v_s - v_0\|}{M}.$$

Then (2.6) becomes (2.2). \blacksquare

The following theorem characterizes saddle points as local minimax solutions.

THEOREM 2.1. *Let $v_0 \in S_L$. Suppose that J has a local peak selection p w.r.t. L at v_0 satisfying (1) p is continuous at v_0 , (2) $d(p(v_0), L) > 0$ and (3) v_0 is a local minimum point of $J(p(\cdot))$. Then $p(v_0)$ is a critical point of J .*

Proof. Suppose $p(v_0)$ is not a critical point, then, by Lemma 2.3, there is $s_0 > 0$ s.t.

$$J(p(v_s)) < J(p(v_0)) - \frac{\theta d(p(v_0), L)\|\nabla J(p(v_0))\|}{8M}\|v_s - v_0\|, \quad \forall s \in (0, s_0)$$

where $p(v_0) = t_0v_0 + w_0$ ($t_0 \neq 0$ and $w_0 \in L$), $v_s = \frac{v_0 - \text{sign}(t_0)sG(p(v_0))}{\|v_0 - \text{sign}(t_0)sG(p(v_0))\|}$ and $G(p(v_0))$ is a modified pseudo-gradient of J with the constant θ at $p(v_0)$ as defined in Lemma 2.1. This contradicts the assumption that v_0 is a local minimum point of $J(p(v))$. \blacksquare

To establish an existence result, we need the following PS condition.

DEFINITION 2.3. *A functional $J \in C^1(X, R)$ is said to satisfy the Palais-Smale (PS) condition if any sequence $\{u_i\} \subseteq X$ s.t. $J(u_i)$ is bounded and $\nabla J(u_i) \rightarrow 0$ possesses a convergent subsequence.*

THEOREM 2.2. *Let $J \in C^1(X, R)$ satisfy the PS condition. If there is a peak selection p of J w.r.t. L satisfying (1) p is continuous, (2) $d(p(v), L) \geq \alpha, \forall v \in S_L$ for some $\alpha > 0$ and*

(3) $\inf_{v \in S_{L'}} J(p(v)) > -\infty$, then there is $v_0 \in S_{L'}$ s.t. $p(v_0)$ is a critical point of J , and

$$J(p(v_0)) = \min_{v \in S_{L'}} J(p(v)).$$

Proof. Since $S_{L'}$ is a closed subset and $J(p(\cdot))$ is a continuous function on $S_{L'}$, bounded from below, by Ekeland's variational principle [20], for any integer n , there is $v_n \in S_{L'}$ s.t.

$$(2.7) \quad J(p(v_n)) \leq \inf_{v \in S_{L'}} J(p(v)) + \frac{1}{n} \text{ and } J(p(v)) - J(p(v_n)) \geq -\frac{1}{n} \|v - v_n\|, \quad \forall v \in S_{L'}, v \neq v_n.$$

By Lemma 2.3, for some $v \in S_{L'}$ and close to v_n ,

$$J(p(v)) - J(p(v_n)) < -\frac{\theta d(p(v_n), L) \|\nabla J(p(v_n))\|}{8M} \|v - v_n\|.$$

Thus

$$(2.8) \quad \|\nabla J(p(v_n))\| < \frac{8M}{n\theta d(p(v_n), L)} \leq \frac{8M}{n\theta\alpha}.$$

By the PS condition, $\{p(v_n)\}$ has a subsequence, denoted again by $\{p(v_n)\}$, converging to a point $u_0 \in X$. If denote $p(v_n) = t_n v_n + x_n$ where $t_n \in R$ and $x_n \in L$, then, $\{t_n v_n\}$ is convergent since $X = L \oplus L'$. Hence, $\{t_n\}$ is convergent. Assume $\{t_n\}$ is a convergent subsequence. Denote $t_0 = \lim_{n \rightarrow \infty} t_n$. Then, by our assumption (2), $|t_0| \geq \alpha > 0$. Thus, $v_n \rightarrow v_0 \in S_{L'}$. Since p is continuous, by (2.8), $p(v_0)$ is a critical point of J and by (2.7), $J(p(v_0)) = \min_{v \in S_{L'}} J(p(v))$. ■

3 A Local Minimax Algorithm in Banach Spaces

Let u_1, u_2, \dots, u_{n-1} be $n - 1$ previously found critical points of J , $L = [u_1, u_2, \dots, u_{n-1}]$, $X = L \oplus L'$. Given $\varepsilon, \lambda > 0$ and $\theta \in (0, 1)$. A flow chart of the algorithm reads:

Step 1: Let $v_n^1 \in S_{L'}$ be an increasing-decreasing direction at u_{n-1} .

Step 2: Set $k = 1$ and solve for

$$\begin{aligned} u_n^k &= p(v_n^k) = t_0^k v_n^k + t_1^k u_1 + \dots + t_{n-1}^k u_{n-1} \\ &= \arg \text{local-max} \{J(t_0 v_n^k + t_1 u_1 + \dots + t_{n-1} u_{n-1}) \mid t_i \in R, i = 0, 1, \dots, n-1\}. \end{aligned}$$

Step 3: If $\|\nabla J(p(v_n^k))\| < \varepsilon$, then output u_n^k , stop. Otherwise, do Step 4.

Step 4: Find a descent direction $w_n^k = -\text{sign}(t_0^k)G_n^k$ of J at u_n^k , where $G_n^k \in L'$ is a modified pseudo-gradient of J at $u_n^k = p(v_n^k)$ with the constant θ as defined in Lemma 2.1.

Step 5: Denote $v_n^k(s) = \frac{v_n^k + sw_n^k}{\|v_n^k + sw_n^k\|}$ and

$$p(v_n^k(s)) = \arg \text{ local-max} \left\{ J(t_0 v_n^k(s) + \sum_{i=1}^{n-1} t_i u_i) \mid t_i \in R, i = 0, 1, \dots, n-1 \right\},$$

where initial point $(t_0^k, t_1^k, \dots, t_{n-1}^k)$ is used. Then solve

$$s_n^k = \max \left\{ s = \frac{\lambda}{2^m} \mid m \in N, 2^m > \|w_n^k\|, J(p(v_n^k(s))) - J(p(v_n^k)) \leq -\frac{\theta}{4} |t_0^k| s \|\nabla J(p(v_n^k))\| \right\}$$

and set $u_n^{k+1} = p(v_n^{k+1}) = p(v_n^k(s_n^k))$.

Step 6: Update $k = k + 1$ and go to Step 3.

REMARK 3.1. It is worthwhile making some remarks on the algorithm:

- (a) $v_n^1 \in S_{L'}$ is an increasing-decreasing direction at u_{n-1} if $J(u_{n-1} + tv_n^1) > (<) J(u_{n-1})$ when $|t|$ is small (large). It reflects a local mountain pass structure. If X is a Hilbert space, by taking $L' = L^\perp$ and $G_n^k = \nabla J(u_n^k)$, it becomes Li-Zhou's algorithm.
- (b) Step 5 will not stop until $\|\nabla J(u_n^k)\| < \varepsilon$ since by Lemma 2.1, $\nabla J(u_n^k) \neq 0$ implies $G(u_n^k) \neq 0$.
- (c) There are two key steps: (1) computation of a modified pseudo-gradient, (2) finite-dimensional unconstrained optimization. (2) can be done by some standard optimization method. The implementation of (1) will be addressed later.
- (d) In Step 3, theoretically we can either follow a modified pseudo-gradient flow given by Lemma 2.1, i.e., to keep the continuity of G_n^k in u_n^k or just find a modified pseudo-gradient.
- (e) The following theorem indicates that the algorithm is stable.

THEOREM 3.1. *In the algorithm, if $u_n^k = p(v_n^k) \notin L$, $\nabla J(u_n^k) \neq 0$ and p is continuous at $v_n^k \in S_{L'}$, then $s_n^k > 0$ and $u_n^{k+1} = p(v_n^k(s_n^k))$ is well defined. Consequently $J(u_n^{k+1}) < J(u_n^k)$.*

Proof. By the setsize rule and Lemma 2.3. ■

4 Computation of Pseudo-Gradients

In this section, we present some formulas to compute a pseudo-gradient and a pseudo-gradient flow in $L^p(\Omega)$ ($p > 1$). Their modified versions follow from a projection to a subspace. Assume that Ω is a measurable space with measure μ and $\|\cdot\|_p$ represents the norm in $L^p(\Omega)$. Let us recall some wellknown results.

LEMMA 4.1. ([10]) Let $f, \{f_n\}$ be in $L^p(\Omega)$, $1 \leq p < \infty$,

(a) if $f_n \rightarrow f$ in $L^p(\Omega)$, then $\{f_n\}$ has a subsequence that converges to f pointwise a.e.;

(b) if $f_n \xrightarrow{a.e.} f$ and $\|f_n\|_p \rightarrow \|f\|_p$, then $f_n \rightarrow f$ in $L^p(\Omega)$. ■

LEMMA 4.2. Let $p, q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$ and $f, f_n \in L^q(\Omega)$ s.t. $f_n \rightarrow f$. Then $sign(f_n)|f_n|^{\frac{1}{p-1}} \rightarrow sign(f)|f|^{\frac{1}{p-1}}$ in $L^p(\Omega)$, where

$$sign(g)(x) = \begin{cases} 1 & \text{if } g(x) \geq 0, \\ -1 & \text{if } g(x) < 0, \end{cases} \quad g \in L^q(\Omega).$$

Proof. It suffices to show that any subsequence, denoted always by $\{sign(f_n)|f_n|^{\frac{1}{p-1}}\}$, has a subsequence that converges to $sign(f)|f|^{\frac{1}{p-1}}$ in $L^p(\Omega)$. Since $f_n \rightarrow f$ in $L^q(\Omega)$, by Lemma 4.1, we have $|f_n|^{\frac{1}{p-1}} \xrightarrow{a.e.} |f|^{\frac{1}{p-1}}$. It follows,

$$sign(f_n)(x)|f_n(x)|^{\frac{1}{p-1}} \xrightarrow{a.e.} sign(f)(x)|f(x)|^{\frac{1}{p-1}}.$$

Since $\frac{1}{p} + \frac{1}{q} = 1$ and $f_n \rightarrow f$ in $L^q(\Omega)$, it leads to

$$\|sign(f_n)|f_n|^{\frac{1}{p-1}}\|_p^p = \|f_n\|_q^q \rightarrow \|sign(f)|f|^{\frac{1}{p-1}}\|_p^p = \|f\|_q^q.$$

By Lemma 4.1, the proof is complete. ■

THEOREM 4.1. Let $p \geq 2$ and $\frac{1}{p} + \frac{1}{q} = 1$. Assume that $J : L^p(\Omega) \rightarrow R$ is Fréchet differentiable at $f \in L^p(\Omega)$ s.t. $\nabla J(f) \neq 0$. Let $G(f) = sign(\nabla J(f))|\nabla J(f)|^{\frac{1}{p-1}}$. Then

$$\Psi(f) = \frac{G(f)}{\|\nabla J(f)\|_q^{q-1}}$$

is a pseudo-gradient of J at f with the constant 1. If in addition, J is C^1 , then Ψ is a pseudo-gradient flow of J with the constant 1.

Proof. $\|\Psi(f)\|_p = 1$ can be seen from

$$\|G(f)\|_p = \left(\int_{\Omega} |\nabla J(f)|^{\frac{p}{p-1}} d\mu \right)^{\frac{1}{p}} = \|\nabla J(f)\|_q^{\frac{q}{p}} = \|\nabla J(f)\|_q^{q-1}.$$

On the other hand,

$$\langle \nabla J(f), G(f) \rangle = \int_{\Omega} \nabla J(f)(x) G(f)(x) d\mu = \int_{\Omega} |\nabla J(f)(x)|^{\frac{p}{p-1}} d\mu = \|\nabla J(f)\|_q^q.$$

Hence $\langle \nabla J(f), \Psi(f) \rangle = \|\nabla J(f)\|_q$ and $\Psi(f)$ is a pseudo-gradient at f with the constant 1.

To show Ψ is continuous. Let $f_0 \in L^p(\Omega)$ with $\nabla J(f_0) \neq 0$ and $\{f_n\} \subseteq L^p(\Omega)$ s.t. $f_n \rightarrow f_0$. Since $J \in C^1(L^p(\Omega), R)$, we have $\nabla J(f_n) \rightarrow \nabla J(f_0)$ in $L^q(\Omega)$ and $\|\nabla J(f_n)\|_q \rightarrow \|\nabla J(f_0)\|_q$. Then Lemma 4.2 leads to

$$G(f_n) \rightarrow G(f_0) \text{ in } L^p(\Omega), \text{ i.e. } \Psi(f_n) \rightarrow \Psi(f_0) \text{ in } L^p(\Omega). \quad \blacksquare$$

THEOREM 4.2. *Let $1 < p < 2$, $\frac{1}{p} + \frac{1}{q} = 1$, $\mu(\Omega) < \infty$ and $\theta = \max(1, (\mu(\Omega))^{\frac{1}{p}-\frac{1}{2}})$. If $J : L^p(\Omega) \rightarrow R$ is Fréchet differentiable at f with $\nabla J(f) \neq 0$. Then $\Psi(f) = \frac{\nabla J(f)}{\theta \|\nabla J(f)\|_2}$ is a pseudo-gradient of J with the constant θ^{-2} at f . If in addition, J is C^1 , then Ψ is a pseudo-gradient of J with the constant θ^{-2} .*

Proof. By the Hölder inequality, we have

$$(4.1) \quad \|\nabla J(f)\|_p \leq \|\nabla J(f)\|_2 (\mu(\Omega))^{\frac{1}{p}-\frac{1}{2}} \quad \text{or} \quad \|\Psi(f)\|_p \leq 1.$$

It follows

$$\langle \nabla J(f), \Psi(f) \rangle = \int_{\Omega} \nabla J(f)(x) \frac{\nabla J(f)(x)}{\theta \|\nabla J(f)\|_2} d\mu = \frac{\|\nabla J(f)\|_2}{\theta} \geq \frac{\|\nabla J(f)\|_p}{\theta^2}.$$

Hence $\Psi(f) = \frac{\nabla J(f)}{\theta \|\nabla J(f)\|_2}$ is a pseudo-gradient of J at f with the constant θ^{-2} .

To show Ψ is continuous, let $\{f_n\} \subset L^p(\Omega)$ s.t. $f_n \rightarrow f$ in $L^p(\Omega)$. Since $J \in C^1(L^p(\Omega), R)$, $\nabla J(f_n) \rightarrow \nabla J(f)$ in $L^q(\Omega)$. It follows $\nabla J(f_n) \rightarrow \nabla J(f)$ in $L^p(\Omega)$ and $\|\nabla J(f_n)\|_2 \rightarrow \|\nabla J(f)\|_2$, since $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p < 2 < q$ and $\mu(\Omega) < \infty$. Hence $\Psi(f_n) \rightarrow \Psi(f)$ in $L^p(\Omega)$, i.e., Ψ is a pseudo-gradient flow of J with the constant θ^{-2} . \blacksquare

In a general Banach space X , when $\nabla J(f) \neq 0$ is computed in X^* at some $f \in X$, a pseudo-gradient of J at f corresponding to a constant $0 < \theta < 1$ can be computed through

$$\sup_{\psi \in X, \|\psi\|_X=1} \left\langle \frac{\nabla J(f)}{\theta \|\nabla J(f)\|_{X^*}}, \psi \right\rangle,$$

which has an upper bound $\frac{1}{\theta}$. It seems to us that it is extremely difficult, in this case, to derive an explicit formula for computing a pseudo-gradient for a functional $J : W^{1,p}(\Omega) \rightarrow R$. Instead we develop some numerical techniques to do the job in the next section.

5 Application to Quasilinear Elliptic PDE

Consider solving the following quasilinear elliptic BVP for multiple solutions:

$$(5.1) \quad \Delta_p u(x) + f(x, u(x)) = 0, x \in \Omega, \quad u \in X \equiv W_0^{1,p}(\Omega), \quad p > 1,$$

where Ω is an open bounded domain in R^n and $\Delta_p u(x) = \text{div}(|\nabla u(x)|^{p-2} \nabla u(x))$ is the nonlinear p-Laplacian differential operator, which has a variety of applications in physical fields, such as in fluid dynamics when the shear stress $\vec{\tau}$ and the velocity gradient ∇u of the fluid are related in the manner $\vec{\tau}(x) = r(x)|\nabla u|^{p-2} \nabla u$, where $p = 2, p < 2, p > 2$ if the fluid is Newtonian, pseudoplastic, dilatant, respectively. The p-Laplacian operator also appears in the study of flow in a porous media ($p = \frac{3}{2}$), nonlinear elasticity ($p > 2$) and glaciology ($p \in (1, \frac{4}{3})$) [4]. So far people's knowledge about solutions to (5.1) is still very limited. We hope to examine the qualitative behavior of solutions and find new phenomena through numerical investigation. We have $X^* \equiv W_0^{-1,q}(\Omega)$ where $\frac{1}{p} + \frac{1}{q} = 1$. Under certain standard conditions on f , weak solutions of (5.1) coincide with critical points of the functional

$$(5.2) \quad J(u) = \frac{1}{p} \int_{\Omega} |\nabla u(x)|^p dx - \int_{\Omega} F(x, u(x)) dx \quad \text{where} \quad F(x, t) = \int_0^t f(x, s) ds.$$

For $u \in X$, to find the gradient $d = \nabla J(u) \in X^*$, for each $v \in X$, we have

$$\begin{aligned} \langle d, v \rangle &= \int_{\Omega} \nabla d(x) \nabla v(x) dx = \int_{\Omega} -\Delta d(x) v(x) dx = \frac{d}{dt} \Big|_{t=0} J(u + tv) \\ &= \int_{\Omega} \left(|\nabla u(x)|^{p-2} \nabla u(x) \nabla v(x) - f(x, u(x)) v(x) \right) dx = \int_{\Omega} (-\Delta_p u(x) - f(x, u(x))) v(x) dx. \end{aligned}$$

Thus $d = \nabla J(u)$ can be computed through solving the linear elliptic equation

$$(5.3) \quad \begin{cases} \Delta d(x) = \Delta_p u(x) + f(x, u(x)), & x \in \Omega, \\ d(x) = 0, & x \in \partial\Omega. \end{cases}$$

Where since $\Delta_p u(x) + f(x, u(x)) \in W_0^{-1,q}(\Omega)$, we have $d \in W_0^{1,q}(\Omega)$. When $u = p(v)$ for some $v \in S_{L'}$, by the definition of a peak selection, $d = \nabla J(u)$ satisfies

$$\langle d, w \rangle = \int_{\Omega} \nabla d(x) \nabla w(x) dx = 0, \quad \forall w \in L,$$

i.e., $d = \nabla J(u) \perp L$. In our numerical examples, we check the ratio

$$(5.4) \quad \gamma = \frac{\|d\|_2^2}{\|d\|_p \cdot \|d\|_q},$$

where $\|\cdot\|_r$ is the norm in $W_0^{1,r}(\Omega)$. $\gamma \leq 1$ by the Hölder inequality. If $\gamma > \alpha > 0$, then $G(u) = \frac{d}{\|d\|_p} \in L'$ is a modified pseudo-gradient of J at u as in Lemma 2.1. It is interesting to point out that although we have not been able to analytically prove $\gamma > \alpha > 0$, we can numerically check this ratio in each computation. All our numerical examples show that the ratio γ is a way above 0. For $p > 2$, since $X \subset X^*$, we define $L' = L^\perp = \{v \in X : \langle u, v \rangle = 0, \forall u \in L\}$. For $p < 2$, $\nabla J(u) \in X^* \subset X$, it can be used directly in the algorithm. Set $G_n^k = \nabla J(u_n^k)$ in Step 3 of the algorithm and modify the stepsize rule in Step 5 as

$$s_k = \max \left\{ s = \frac{\lambda}{2^m} \mid m \in N, 2^m > \|w_k\|, J(p(v_k(s))) - J(p(v_k)) \leq \frac{|t_0^k|s}{-4} \|\nabla J(p(v_k))\|_2^2 \right\}.$$

where $\|\cdot\|_2$ is the norm in $W_0^{1,2}(\Omega)$. The convergence will be handled with extra care [21].

Next, the local minimax algorithm is applied to find multiple solutions for the p-Emden-Flower Equation:

$$(5.5) \quad \Delta_p u(x) + |u(x)|^{q-1} u(x) = 0, \quad x \in \Omega, \quad u \in W_0^{1,p}(\Omega)$$

and the p-Henon Equation:

$$(5.6) \quad \Delta_p u(x) + |x - \vec{1}|^r |u(x)|^{q-1} u(x) = 0, \quad x \in \Omega, \quad u \in W_0^{1,p}(\Omega)$$

where $|\cdot|$ is the Euclidean norm, $\vec{1} = (1, \dots, 1)$, $1 < p < q + 1 < p^*$ with $p^* = \frac{np}{n-p}$ for $p < n$ and $p^* = \infty$ for $n \leq p$, and n is the dimension of the domain space. In our computation, $\Omega = [0, 2] \times [0, 2] \subset R^2$.

Note that the right-hand-side of (5.3) involves an evaluation of a higher-order derivative of a numerical solution u , i.e., $\Delta_p u(x)$, which causes difficulty for using linear finite elements. To solve the problem², we utilize a weak form of (5.3)

$$(5.7) \quad \int_{\Omega} \Delta d(x) v(x) dx = \int_{\Omega} (\Delta_p u(x) + f(x, u(x))) v(x) dx \quad \forall v \in W_0^{1,p}$$

²The authors would like to thank Dr. Jun Zhao for providing us with this technique suggestion

and the identity

$$(5.8) \quad \int_{\Omega} \Delta_p u(x) v(x) dx = - \int_{\Omega} |\nabla u(x)|^{p-2} \nabla u(x) \nabla v(x) dx \quad \forall v \in W_0^{1,p}$$

to replace the higher-order derivative term by a first-order derivative term. Thus linear finite elements can be applied. Since $p < 2$ and $p > 2$ represent two distinct mathematical and physical applications, and the parameter r is related to the symmetry breaking phenomenon, we will select some values for p and r to examine their solution profiles. In all figures, Itn , if listed, is the number of iterations, e , if listed, is the the norm of the gradient at the ending point and γ_{min} , if listed, is the minimum ratio of (5.4) in the last 10 iterations of the computation for each solution. L is the support used in the algorithm and $L = [0]$ if it is not listed. In our numerical computations, $\lambda = 1$ and $\varepsilon = 10^{-3}$, either 400×400 or 800×800 linear square elements are used. Thus our equations and finite-element meshes possess certain symmetries. For computational efficiency, some solutions are captured by using symmetry invariance, refer to [19] for details. In this case, the symmetry is listed in the caption for reference.

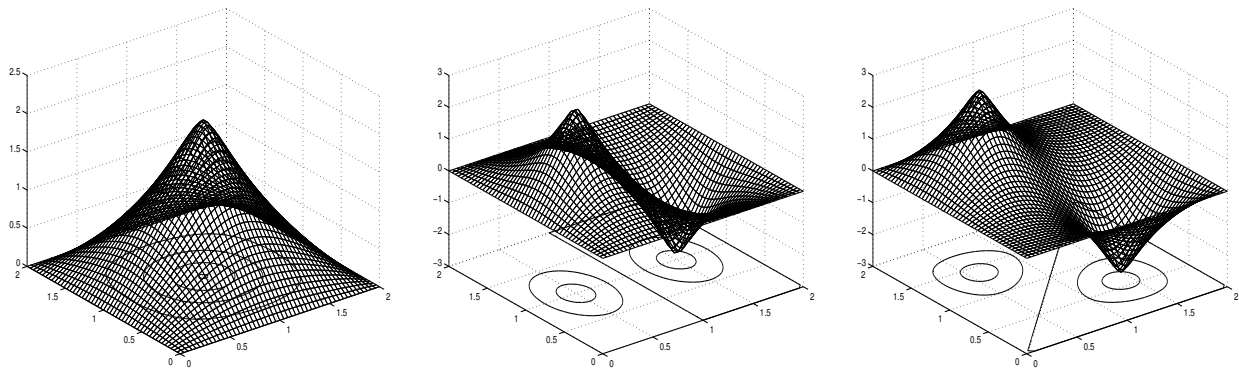


FIG. 1. Equation (5.5) with $p=3$, $q=7$. The ground state u_1 with $J = 4.4829$, $\gamma_{min} = 0.88$ (left). Two solutions with $J = 40.9568$, $\gamma_{min} = 0.93$, $L = [u_1]$ (middle) and $J = 34.4457$, $\gamma_{min} = 0.65$, $L = [u_1]$ (right).

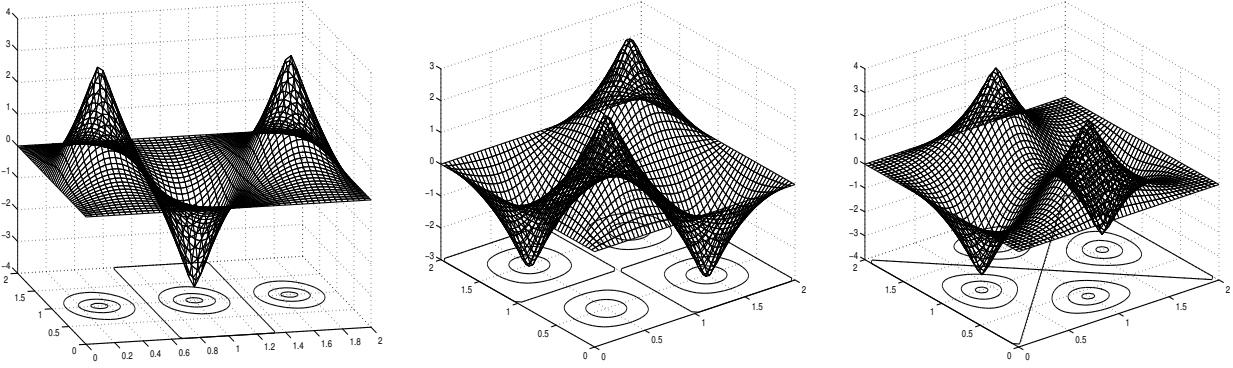


FIG. 2. Equation (5.5) with $p=3$, $q=7$. Three solutions with $J = 188.5327$, $\gamma_{min} = 0.73$ (left, locally odd symmetric about the lines $x = \frac{2}{3}$, $x = \frac{4}{3}$), $J = 124.875$, $\gamma_{min} = 0.83$ (middle, odd symmetric about the lines $y = 1$ and $x = 1$) and $J = 181.7966$, $\gamma_{min} = 0.88$ (right, odd symmetric about the lines $y = x$, $x + y = 2$).

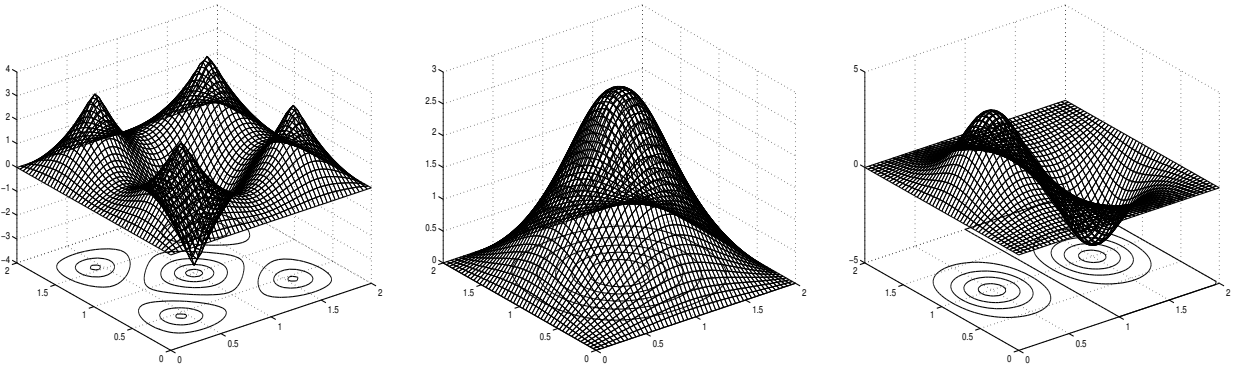


FIG. 3. Equation (5.5) with $p=3$, $q=7$. A solution with $J = 228.2925$, $\gamma_{min} = 0.74$ and $L = [u_1]$ (left, 4-rotation symmetry). Equation (5.5) with $p = 1.75$, $q = 3$. The ground state u_1 with $J = 7.0745$, $\gamma_{min} = 0.9$, $Itn = 13$, $e = 0.000247$ (middle) and a solution u_2 with $J = 25.4653$, $\gamma_{min} = 0.96$, $Itn = 16$, $e = 0.000375$, $L = [u_1]$ (right).

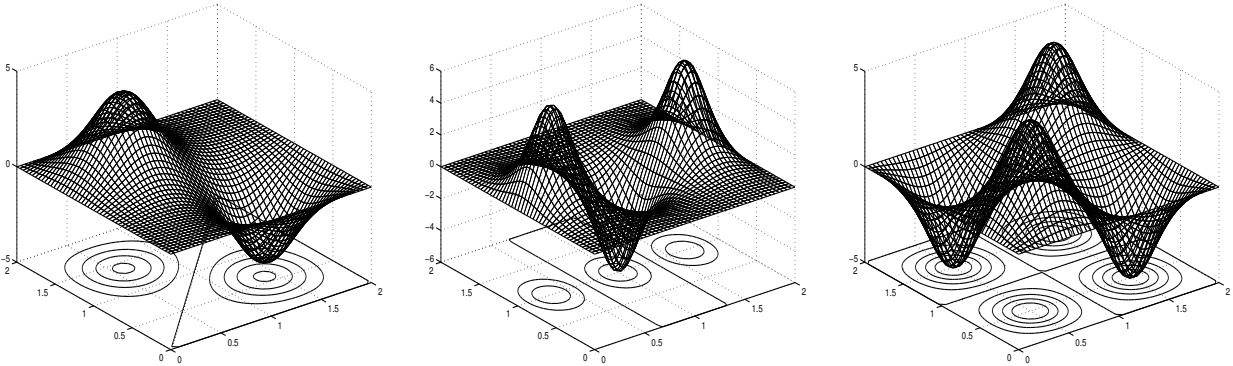


FIG. 4. Equation (5.5) with $p = 1.75$, $q = 3$. Three solutions with $J = 24.0274$, $\gamma_{min} = 0.91$, $Itn = 37$, $e = 0.000643$, $L = [u_1]$ (left), $J = 59.4209$, $Itn = 31$, $\gamma_{min} = 0.93$, $e = 0.000238$ (middle, locally odd symmetric about the lines $x = \frac{2}{3}$, $x = \frac{4}{3}$) and $J = 61.1246$, $\gamma_{min} = 0.98$, $Itn = 33$, $e = 0.000197$, $L = [u_1, u_2(x, y), u_2(y, x)]$ (right).

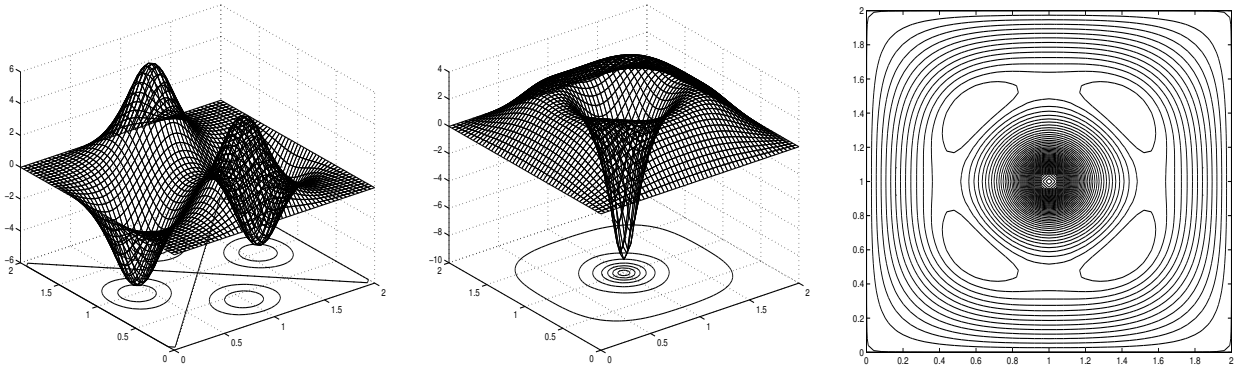


FIG. 5. Equation (5.5) with $p = 1.75$, $q = 3$. Two solutions with $J = 70.6261$, $\gamma_{min} = 0.94$, $I_{tn} = 61$, $e = 0.000205$ (left, odd symmetric about the lines $y = x, x + y = 2$) and $J = 77.2337$, $\gamma_{min} = 0.91$, $I_{tn} < 160$, $e = 0.000805$, $L = [u_1]$ (middle, 4-rotation symmetry) and its contours (right).

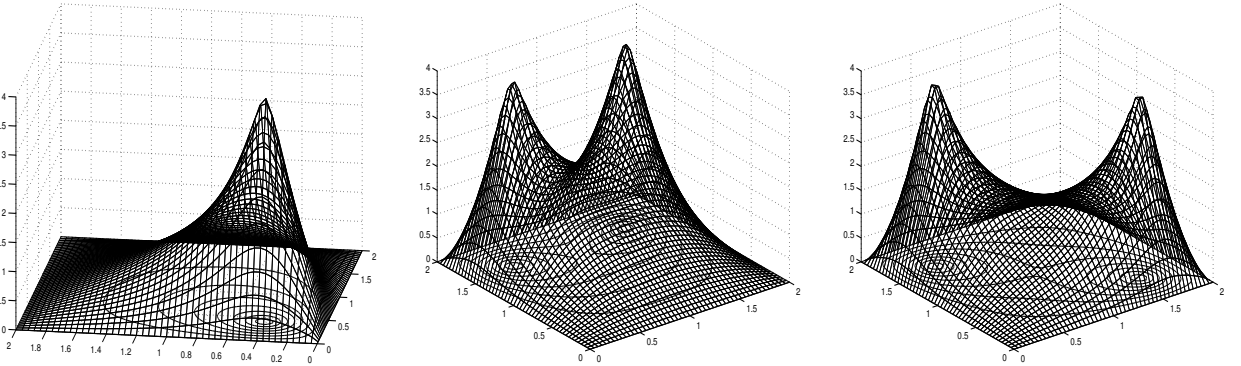


FIG. 6. Equation (5.6) with $p = 3, q = 7, r = 7$. The ground state with $J = 60.46$ (left). A 2-peak solution with $J = 116.231$ (middle, symmetric about the line $x = 1$). Another 2-peak solution with $J = 118.906$ (right, symmetric about the lines $y = x, x + y = 2$).

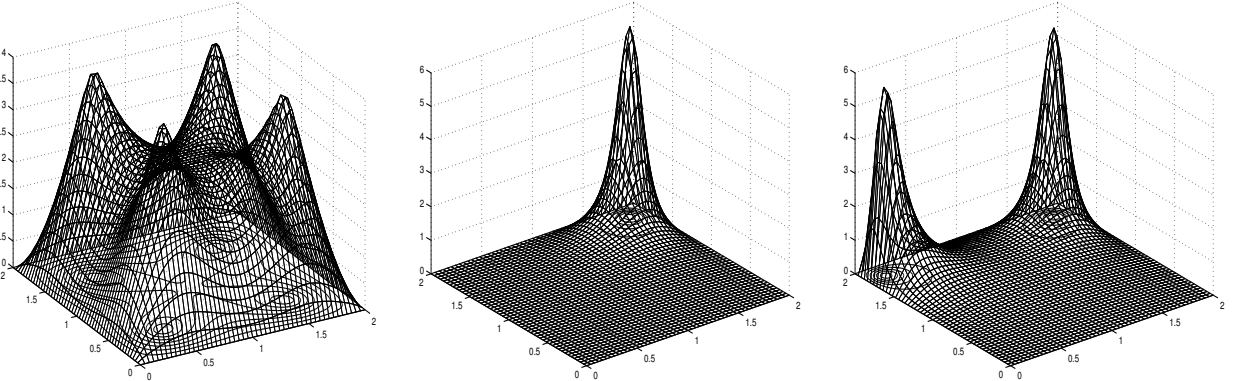


FIG. 7. Equation (5.6) with $p = 3, q = 7, r = 7$. A 4-peak solution with $J = 219.8671$ (left, 4-rotation symmetry). Equation (5.6) with $p = 1.75, q = 3, r = 7$. The ground state u_1 with $J = 15.7588$ (middle). A 2-peak solution with $J = 31.3832$ (right, symmetric about the line $x = 1$).

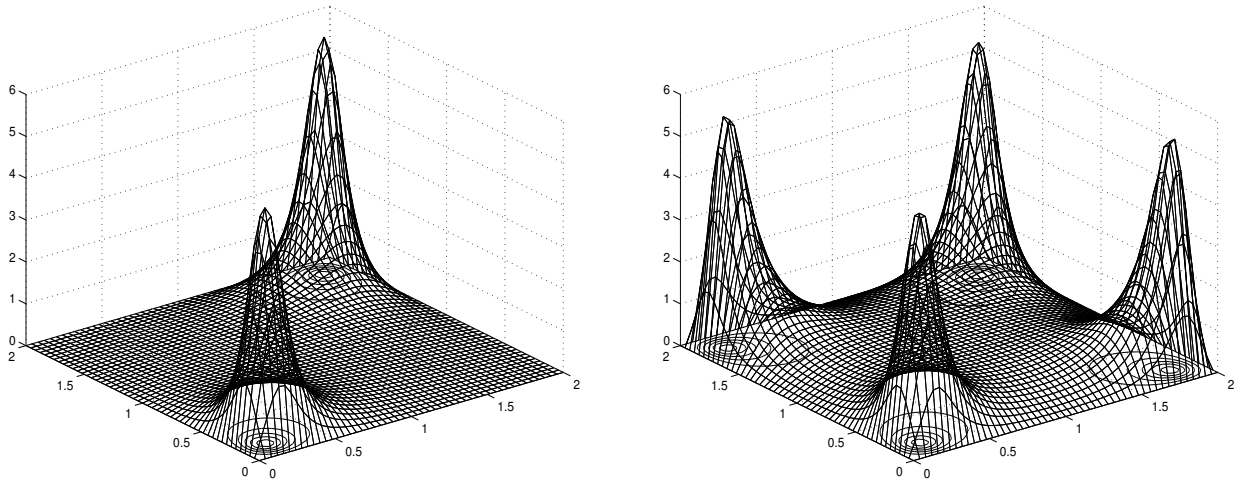


FIG. 8. Equation (5.6) with $p = 1.75, q = 3, r = 7$. Another 2-peak solution with $J = 31.4178$ (left, symmetric about the lines $y = x, x + y = 2$) and a 4-peak solution with $J = 62.2163, L = [u_1]$ (right, 4-rotation symmetry).

REMARK 5.1. It is to the best of our knowledge that the above solutions are the first time to be computed and visualized. Solution profiles to (5.5) are listed in Fig. 1(left)-Fig. 3(left) for $p < 2$ and in Fig. 3(middle)-Fig. 5(right) for $p > 2$. It deserves to notice the contour, the right of Fig.5, which shows the solution, the middle of Fig.5, has 5 peaks. Fig. 3(left) with Fig. 5(middle). For (5.6), if r is small, the solution profiles are similar to that of (5.5). When r is large, as shown by solution profiles in Fig. 6(left)-Fig. 7(left) for $p < 2$ and in Fig. 7(middle)-Fig. 8(right) for $p > 2$, the ground states in Fig. 6(left) and Fig. 7(middle) become asymmetric and multi-peak positive solutions take place. Such cases are called symmetry breaking phenomena and are impossible to (5.5).

As a final remark concerning convergence of the algorithm, we point out that due to multiplicity of solutions and Banach space setting, convergence analysis of the algorithm is very complicated. In a subsequent paper [21], we established a unified point-to-set convergence result (Theorem 2.1 in [21]) for a theoretically generalized version of the algorithm by first proving an abstract convergence result with a general condition and then verifying that the pseudo-gradient flow or pseudo-gradients computed in several different ways all satisfy the condition.

Acknowledgement: The authors would like to thank two anonymous referees for their interesting comments.

References

- [1] A. Ambrosetti and 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 and P.J. McKenna, A mountain pass method for the numerical solution of semilinear elliptic problems, *Nonlinear Analysis*, 20(1993), 417-437.
- [4] J.I. Diaz, *Nonlinear Partial Differential Equations and Free Boundaries*, vol. I, Elliptic Equations, Research Notes in Mathematics 106, Pitman Publishing, Boston, 1985.
- [5] W.Y. Ding and W.M. Ni, On the existence of positive entire solutions of a semilinear elliptic equation, *Arch. Rational Mech. Anal.*, 91(1986) 238-308.
- [6] Z. Ding, D. Costa and G. Chen, A high linking method for sign changing solutions for semilinear elliptic equations, *Nonlinear Analysis*, 38(1999), 151-172.
- [7] M. Fabian, P. Habala, P. Hajek, V.M. Santalucia, J. Pelant and V. Zizler, *Functional Analysis and Infinite-Dimensional Geometry*, Springer, New York, 2001.
- [8] G. Henkelman, G. Johannesson and H. Jonsson, Methods for finding saddle points and minimum energy paths, *Comput. Chemistry*, D. Schwartz, ed., vol. 5, Kluwer, 2000.
- [9] Prem K. Kythe, *Differential Operators and applications*, Birkhauser, Boston, 1996.
- [10] Y. Li and J. Zhou, A minimax method for finding multiple critical points and its applications to nonlinear PDEs, *SIAM Sci. Comp.*, 23(2001), 840-865.
- [11] Y. Li and J. Zhou, Convergence results of a local minimax method for finding multiple critical points, *SIAM Sci. Comp.*, 24(2002), 865-885.
- [12] F. Lin and T. Lin, Minimax solutions of the Ginzburg-Landau equations, *Slecta Math.(N.S.)*, 3(1997) no.1, 99-123.
- [13] J. Mawhin and M. Willem, *Critical Point Theory and Hamiltonian Systems*, Springer-Verlag, New York, 1989.
- [14] Z.H. Musslimani, M. Segev, D.N. Christodoulides and M. Soljacic, Composite Multihump vector solitons carrying topological charge, *Phy. Rev. Lett.* 84(2000), 1164-1167.
- [15] P. Rabinowitz, *Minimax Method in Critical Point Theory with Applications to Differential Equations*, CBMS Regional Conf. Series in Math., No.65, AMS, Providence, 1986.
- [16] M. Schechter, *Linking Methods in Critical Point Theory*, Birkhauser, Boston, 1999.

- [17] M.Struwe, *Variational Methods*, Springer, 1996.
- [18] Z.-Q. Wang, On a superlinear elliptic equation, *Ann. Inst. Henri Poincare*, 8(1991), 43-57.
- [19] Z.-Q. Wang and J. Zhou, An Efficient and Stable Method for Computing Multiple Saddle Points with Symmetries. *SIAM Num. Anal.*, to appear.
- [20] M.Willem, *Minimax Theorems*, Birkhauser, Boston, 1996.
- [21] X.Yao and J.Zhou, A unified convergence result on a minimax method for finding multiple critical points in Banach spaces, preprint.