Solving Multiple Solution Problems: Computational Methods and Theory Revisited

Jianxin Zhou*

Abstract

This paper first gives a simple survey on methods and theory for numerically solving multiple solution problems existed in computational mathematics, physics, chemistry, biology, etc. and then presents a more detailed revisit to the results in this subject obtained by the author’s research group and associates. A long list of references is compiled and cited to provide readers with more sufficient information and better systematic understanding on application background, theoretical existence results, computational methods and theory for solving multiple solution problems.

Keywords Multiple solution problem, Saddles, Minimax method, Instability analysis, Min-⊥-method, Modified Newton method.

AMS(MOS) subject classifications. 65N20,58E05,35A15,49M05

1 Introduction

Let $B$ and $V$ be Banach spaces and $F: \mathbb{R} \times B \rightarrow V$ be a nonlinear operator. Mathematically at least the following classes of nonlinear multiple solution problems can be explored [3,9,10,16,17,95]. Multiple solutions due to a singular linear operator will not considered here.

1. Multiple solution problem: Given $\lambda \in \mathbb{R}$, find all $u \in B$ s.t. $F(\lambda, u) = 0$;

2. Multiple fixed point problem: Given $\lambda \in \mathbb{R}, V \subset B$, find all $u \in V$ s.t. $F(\lambda, u) = u$;

3. Nonlinear eigen-problem: Find all eigen-solutions $(\lambda, u) \in \mathbb{R} \times B$ s.t. $F(\lambda, u) = 0$ where $\lambda$ is called an eigen-value and $u$ is its eigen-function. Since some eigen-values may have multiple eigen-functions, it is natural to consider

*Math. Department, Texas A&M University, College Station, TX 77843 (jzhou@math.tamu.edu).
4. Bifurcation problem: For a nonlinear eigen-problem $F(\lambda, u) = 0$, find $\lambda$ across which its multiplicity changes;

5. Nonlinear eigen-function problem: Find $(\lambda, u) \in \mathbb{R} \times B$ s.t. $F(u) = \lambda G(u)$ where $F$ and $G$ are some given operators from $B$ into $V$;

6. Nonlinear eigen-value problem: Find $(\lambda, u) \in \mathbb{R} \times (B \setminus \{0\})$ s.t. $F(\lambda)u = 0$ where $F(\lambda)$ defines a bounded linear operator on $B$, e.g., $F(\lambda) = \lambda^2 A_1 + \lambda A_2 + A_3$ where $B = \mathbb{R}^n$ and $A_1, A_2, A_3 \in \mathbb{R}^{n \times n}$ [6];

7. Viable problem: For given $F, G : B \to \mathbb{R}$ and a threshold $\lambda_0 > 0$, find a largest subspace $S \subset B$ s.t. $F(u) \geq \lambda_0 G(u) \forall u \in S$ [35].

Real world applications provide a rich source of multiple solution problems. As unstable local equilibria, excited states with various configurations and performance indices at different energy levels have been physically observed and mathematically verified to exist in many excitation/reaction/transition processes in physics, chemistry, biology, etc. Nowadays new technologies (synchrotronic, laser, etc.) can be developed to induce, reach or control some of those excited states so that they become long-lived as to be stable for practical purposes. Such excited states are called metastable. Thus search and study of those metastable states for new applications are of great interests in modern science and advanced engineering [2,24,26,32,36,41,42,43,44,45,49,50,58,60,65,67,69,70,73,75,77,79,83,85,87,89,90,97] and call for the development of numerical algorithms to compute such solutions [12,13,14,24,30,32,33,34,40,44,46,52,56,57,64,66,71].

Multiple solution problems can be divided into two classes: variational vs. nonvariational. The above multiple solution problems 1-5 are called variational, if $F(\lambda, u) = f'(\lambda, u)$ and $G(u) = g'(u)$ for some derivatives of functions $f : \mathbb{R} \times B \to \mathbb{R}$ and $g : B \to \mathbb{R}$. Otherwise the problems are called nonvariational. In this survey we mainly focus on variational multiple solution problems except in Section 5 where some non-variational results are presented.

Let $H$ be a Hilbert space with its inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Let $J \in C^1(H, \mathbb{R})$. A point $u^* \in H$ is called a critical point of $J$ if its Frechet derivative $J'(u^*) = 0$. The most well-studied critical points are the local extrema of $J$ on which the classical calculus of variations and numerical methods focus. Critical points $u^*$ that are not local extrema of $J$ are called saddle points, or saddles, i.e., in any neighborhood $\mathcal{N}(u^*)$ of $u^*$, there are points $v, w \in \mathcal{N}(u^*)$ s.t. $J(v) < J(u^*) < J(w)$. An order-k saddle point or k-saddle is a critical point that is a local maximum of $J$ in a k-dimensional subspace and a local minimum of $J$ in the corresponding k-co-dimensional subspace. Such saddles are sometimes called minimax type. Critical points correspond to local equilibrium states in a physical process. Thus mathematically, the saddle order can be used to measure their instabilities, e.g., a ground state as a stable local equilibrium is a local minimum point or 0-saddle;
excited states, as unstable local equilibria, correspond to saddles and metastable states are among the first few saddles. When \( J \) is a \( C^2 \)-functional and \( J''(u^*) \) is a self-adjoint Fredholm operator, there is a spectrum decomposition \( H = H^- \oplus H^0 \oplus H^+ \) where \( H^- \), \( H^0 \) and \( H^+ \) are respectively the maximum negative definite, the null and the maximum positive definite subspaces of \( J''(u^*) \) in \( H \) with \( \dim(H^0) < \infty \). The Morse index \([16]\) of a critical point \( u^* \) is \( \text{MI}(u^*) = \dim(H^-) \). A critical point \( u^* \) is nondegenerate if \( H^0 = \{0\} \), otherwise \( u^* \) is degenerate. If \( u^* \) is nondegenerate, then \( \text{MI}(u^*) \) can be used as the saddle order to measure its local instability \([78]\). Typically a monkey saddle is degenerate and non-minimax type.

The above definition of a critical point can be generalized to the following: Let \( B \) be a Banach space, \( M \subset B \) and \( D(F) \) be the domain of a functional \( F : D(F) \subset B \to \mathbb{R} \). A point \( u^* \in M \) is a critical point of \( F \) in \( M \), if \( D(F) \) contains a neighborhood \( N(u^*) \) of \( u^* \) s.t.

\[
\frac{d}{dt}F(u(t))|_{t=0} = 0, \ \forall u(t) \in M, \ t \in (-\varepsilon, \varepsilon), \ u(0) = u^*, \ u'(0) \text{ exists.} \quad (1.1)
\]

If \( M \) has a tangent space \( TM_{u^*} \) at \( u^* \), then (1.1) leads to \( F'(u^*)h = 0 \ \forall h \in TM_{u^*} \). If \( M = \{u \in B|G(u) = 0\} \) where \( G \in C^1(B, \mathbb{R}) \), then there exists \( \lambda \in \mathbb{R} \) s.t.

\[
F'(u^*) = \lambda G'(u^*) \quad \text{or} \quad \mathcal{L}'(u^*) = F'(u^*) - \lambda G'(u^*) = 0. \quad (1.2)
\]

If \( u^* \in \text{int}(M) \) then \( u^* \) is called a (free) critical point of \( F \), i.e.,

\[
F'(u^*) = 0. \quad (1.3)
\]

When the Frechet derivative \( F' \) is replaced by the generalized derivative \( \partial F \) in the sense of Clarke \([25]\), then (1.3) is replaced by \( 0 \in \partial F(u^*) \) \([15]\). Other derivatives such as the geometric, topological, shape derivatives can also be explored.

Comparing to finding a local minimum, numerical computation for saddles is much more challenging due to their instability and multiplicity involved. Multi-level optimization skills have to be developed.
Many algorithms are proposed in computational physics/chemistry/biology to find 1-saddles in various applications [12,13,14,24,32,36,34,40,44,46,60,64,66,74,96,97]. Most of them are based on physical intuition of the problems and have no proper mathematical justifications such as analysis on algorithm stability and convergence.

A huge literature exists on mathematical analysis of multiple solution problems [3,8-11,15-17,31,68,72,80,84]. Most of them focus on the existence issue. Typically a minimax type critical point is characterized by the Ljusternik-Schnirelman principle (LSP) [95]

$$\min_{A \in A} \max_{u \in A} J(u)$$  \hspace{1cm} (1.4)

where $A$ is a collection of certain compact sets $A$, e.g., a k-D simplex, max and min are all in the global sense. A function $J \in C^1(H, \mathbb{R})$ is said to satisfy the Palais-Smale (PS) condition, if any sequence $\{u_k\} \subset H$ with $\{J(u_k)\}$ bounded and $J'(u_k) \to 0$ has a convergent subsequence. As a milestone in critical point theory, the mountain pass lemma proved by Ambrosetti-Rabinowitz [3] in 1973 states

**Theorem 1.1.** Let $J \in C^2(H, \mathbb{R})$, $e \in H$ and $r > 0$ be s.t. $\|e\| > r$ and

$$b = \inf_{u \in H, \|u\| = r} J(u) > J(0) \geq J(e).$$ \hspace{1cm} (1.5)

If $J$ satisfies the PS condition, then $c = \inf_{\gamma \in \Gamma} \max_{t \in [0,1]} J(\gamma(t))$ is a critical value, i.e., there is $u^* \in H$ s.t. $J(u^*) = c$, $J'(u^*) = 0$, where $\Gamma = \{\gamma \in C([0,1], H) : \gamma(0) = 0, \gamma(1) = e\}$, and $\inf$ and $\max$ are in the global sense.

It is clear that within LSP, the mountain pass lemma uses $A = \Gamma$. Other linking theorems have also been established within LSP framework [68,72,80,84]. Unfortunately LSP is a two-level global optimization problem that is too expensive to solve. One needs new approaches.

To motivate our study, let us consider finding soliton solutions to a semilinear Schrodinger equation, which leads to solving a non-autonomous semilinear elliptic PDE:

$$-\Delta u(x) + \lambda u(x) + \kappa |x|^r |u(x)|^{p-1} u(x) = 0,$$  \hspace{1cm} (1.6)

in $H_0^1(\Omega) = W_0^{1,2}(\Omega)$ where $\Omega$ is an open bounded domain in $\mathbb{R}^n$ and $p > 1$, $\lambda, \kappa, r$ are prescribed parameters. Its energy functional is

$$J(u) = \int_{\Omega} \left[ \frac{1}{2} (|\nabla u(x)|^2 + \lambda u^2(x)) + \frac{\kappa}{p+1} |x|^r |u(x)|^{p+1} \right] dx.$$ \hspace{1cm} (1.7)

Thus solutions to (1.6) correspond to critical points of $J$ in (1.7). Let $0 < \mu_1 < \mu_2 < \cdots$ be the eigen-values of $-\Delta$ satisfying zero Dirichlet boundary condition (B.C.) and $\{v_1, v_2, \ldots\}$ be their corresponding eigen-functions. The system (1.6) is called focusing (M-type) if $\kappa < 0$ and $-\mu_{k+1} < \lambda < -\mu_k$ ($k = 0, 1, 2, \ldots, \mu_0 = -\infty$) and defocusing (W-type) if $\kappa > 0$ and
\( \mu_k < -\lambda < \mu_{k+1} \ (k = 1, 2, \ldots) \). Those two cases are very different in physical nature and mathematical structures, refer Figure 2. For both types, 0 is the only k-saddle. All non-trivial saddles have index \( k (\geq k) \) for M-type (W-type). In particular, for the M-type with \( \lambda > -\mu_1 \), \( J \) has a mountain pass structure and 0 is the only local minimum so the mountain pass/linking type approaches can be applied to prove the existence of multiple even infinitely many solutions. However when \( \lambda < -\mu_1 \), \( J \) does not have a mountain pass structure, the mountain pass type approach cannot be applied; for the W-type with \( k \geq 1 \), \( J \) has two local minima without a mountain pass structure, thus the mountain pass/linking type approach are not applicable. Therefore those two cases have to be treated very differently.

![Figure 2](image)

Figure 2: Typical function profiles of M-type (left) where \( \cap \)-shape in \([v_1, \ldots, v_k]\) vs. M-shape in \([v_1, \ldots, v_k]^\perp\) and W-type (right) where \( \cup \)-shape in \([v_1, \ldots, v_k]\) vs. W-shape in \([v_1, \ldots, v_k]^\perp\). They are not upside-down to each other due to the significant difference in space dimensions.

As numerical computation is concerned, since the min and max in LSP are all in the global sense, it prevents people from designing a feasible numerical algorithm. Thus other approaches must be developed to guide numerical algorithm design. Due to the high nonlinearity/multiplicity of the problems in infinite dimensional spaces, the complexity of obtaining an initial guess sufficiently close to an unknown saddle is commonly viewed as almost the same as that of finding the solution. Thus Newton-type or other local convergence based methods (using \( J'' \)) alone even they are very fast are not suitable for solving such multiple solution problems. On the other hand, since a stable algorithm can always be followed by a Newton or other locally fast convergent method to accelerate its convergence, here we focus on developing a stable algorithm for finding unconstrained k-saddles using only \( J' \).

### 2 A Local Minimax Method

In 1993, Choi-McKenna devised a mountain pass algorithm [24] for finding 1-saddles by using \( A = 1 \)-D simplices, global max and a local min in LSP. In 1999, Ding-Costa-Chen, proposed a high-linking algorithm [30] for finding 2-saddles by using \( A = 2 \)-D simplices, global max and a local min in LSP. Those two methods are successfully applied to solve M-type problem (1.6) for 1-2-saddles. Since the algorithms have gaps to LSP and no stepsize rules, there are no convergence results available. However they used a local min thus actually surpassed the
LSP framework. In 1960, Nehari [62] proved that a \textit{global min} of $J$ on the Nehari manifold
\begin{equation}
N = \{t_u u : u \in H, \|u\| = 1, t_u > 0, \langle J'(t_u u), u \rangle = 0\} \tag{2.1}
\end{equation}
is a saddle. In 1986, Ding-Ni [31] proved that a local min of $J$ on $N$ is also a saddle.

Motivated by their works and the Morse theory, Li-Zhou [52] established a characterization of $n$-saddles by a \textit{local minimax principle (LMMP)}: Let $L = [u_1, ..., u_{n-1}]$ be the space spanned by previously found solutions $u_1, ..., u_{n-1}$ and $S_L \perp = \{u \in L \perp, \|u\| = 1\}$,
\begin{equation}
\min_{v \in S_L \perp} \max_{u \in [v, L]} J(u), \tag{2.2}
\end{equation}
where min and max are all in the local sense. Also a collection of subspaces $[v, L]$, not compact sets, are used in LMMP, thus it is very different from LSP.

**Definition 2.1.** The set-valued peak mapping $P : S_L \perp \rightarrow 2^H$ is defined by
\[ P(v) = \{v^* = \text{arg} \max_{u \in [L, v]} J(u)\}, \forall v \in S_L \perp. \]

A peak selection $p : S_L \perp \rightarrow H$ is a function s.t. $p(v) \in P(v), \forall v \in S_L \perp$. If $p$ is locally defined, then $p$ is a local peak selection.

Thus $p(v) = t_v v + v_L, t_v \neq 0, v_L \in L$ if $p(v) \notin L$, and $J'(p(v)) \perp [L, v], \forall v \in S_L \perp$.

It is known that the energy dissipation law is an important property for an algorithm to be stable. However it is also known that the energy dissipation law alone is not enough for an algorithm to converge, in particular in infinite-dimensional spaces. It needs a stepsize rule, a stronger energy dissipation law. Thus we establish

**Lemma 2.1.** (Stepsize Rule)[52] If $p$ is a local peak selection of $J$ near $v_0 \in S_L \perp$ s.t. $J'(p(v_0)) \neq 0$, $p$ is weak-continuous at $v_0$ and $p(v_0) = t_0 v_0 + v_L^0 \notin L$, then there is $s_0 > 0$ s.t. when $0 < s < s_0$,
\begin{equation}
J(p(v(s))) - J(p(v_0)) < -\frac{t_0 s \|J'(p(v_0))\|^2}{4C_0} \tag{2.3}
\end{equation}
where $v(s) = \frac{v_0 - sd_0}{\|v_0 - sd_0\|} \in S_L \perp, d_0 = \frac{J'(p(v_0))}{C_0}, C_0 = \max\{1, \|J'(p(v_0))\|\}.$

**Theorem 2.1.** (LMMP)[52] If $p$ is a local peak selection of $J$ near $v_0 \in S_L \perp$ s.t. $p$ is weak-continuous at $v_0, p(v_0) \notin L$ and
\begin{equation}
v_0 = \arg \min_{v \in S_L \perp} J(p(v)) = \arg \min_{v \in S_L \perp} \max_{u \in [L, v]} J(u), \tag{2.4}
\end{equation}
then $u_0 = p(v_0)$ is a saddle point of $J.$
In the above results the continuity of \( p \) has been weakened to weak-continuity. LMMP (2.4) characterizes a k-saddle by a local minimization problem \( \min_{v \in S_L} J(p(v)) \). It can be numerically approximated by, e.g., a steepest descent method and leads to a local minimax method (LMM) where a stepsize rule is defined by (2.3). This is important to prove the algorithm stability and convergence. Note that the composite function \( J(p(v)) \) is not assumed to be differentiable in \( v \). However we actually designed a negative gradient method.

2.1 A Local Minimax Algorithm (LMM)[52]

Let \( w_1, ..., w_{n-1} \) be \( n-1 \) previously found critical points, \( L = [w_1, ..., w_{n-1}] \). Given \( \varepsilon > 0, \lambda > 0 \) and \( v^0 \in S_L \) be an ascent-descent direction at \( w_{n-1} \).

Step 1: Let \( t_0^0 = 1, v_L^0 = 0 \) and set \( k = 0 \);

Step 2: Using the initial guess \( w = t_0^kv^k + v_L^k \), solve for

\[
  w^k \equiv p(v^k) = \arg \max_{u \in [L,v^k]} J(u), \text{ denote } t_0^kv^k + v_L^k = w^k \equiv p(v^k);
\]

Step 3: Compute \( d^k = -J'(w^k)/C_k \) where \( C_k = \max\{1, \|J'(w^k)\|\} \);

Step 4: If \( \|d^k\| \leq \varepsilon \) then output \( w_n = w^k \), stop; else goto Step 5;

Step 5: Set \( v^k(s) = \frac{v^k + sd^k}{\|v^k + sd^k\|} \in S_{L^\perp} \) and find

\[
  s^k := \max_{m \in \mathbb{N}} \left\{ \frac{\lambda}{2m} : J(p(v^k(\frac{\lambda}{2m}))) - J(w^k) \leq -\frac{t_0^k \lambda\|d^k\|^2C_k}{2m+1} \right\};
\]

Initial guess \( u = t_0^kv^k(\frac{\lambda}{2m}) + v_L^k \) is used to find \( p(v^k(\frac{\lambda}{2m})) \) where \( t_0^k \) and \( v_L^k \) are found in Step 2. (trace a peak selection)

Step 6: Set \( v^{k+1} := v^k(s^k) \) and update \( k = k + 1 \) then goto Step 2.

Remark 2.1. (1) LMM starts with \( n = 1, L = \{0\} \) to find \( w_1 \), then \( n = 2, L = [w_1] \) to find \( w_2 \), etc.; (2) In many cases such a peak selection \( p \) is unique. If it is not, then it is important to use the given initial guess in Step 5 in order to consistently trace a peak selection; (3) A Matlab package of LMM can be downloaded at http://www.math.tamu.edu/~jzhou/minimax.tar to use for research/education purposes under a proper acknowledgment.
2.2 Some M-type Numerical Examples

Consider solving the Henon equation [47], i.e., $\kappa = -1, \lambda = 0$ and $p = 3$ in (1.6) for several different $r$-values and domains $\Omega$.

The construction of an initial guess $v^0$ for a desired profile is flexible, one may use $\sin / \cos$ functions or solve the equation $-\Delta v^0(x) = cv(x) \ x \in \Omega$ in $H_0^1(\Omega)$ where $cv(x) = +1, -1$ depending on if we want $v^0$ to be concave up or down at $x$ and $cv(x) = 0$ if its concavity at $x$ is not of concern. In Step 3 of LMM, the negative gradient direction $d^k = -\nabla J(w^k)/C_k$ is obtained by numerically solving a linear elliptic PDE of the form $-\Delta d^k + d^k = -J'(w^k)$ in $H_0^1(\Omega)$ and then dividing it by a scalar $C_k$. This is where a numerical approximation is used, such as a finite element method, a finite difference method, a boundary element method, etc. Many numerical solvers can be applied here.

**Example 2.1.** Let $r = 0$ and $\Omega$ be the dumbbell-shaped domain as shown in Figure 3 (a). We focus on finding positive solutions. In Figure 3, (b) the ground state with $J = 10.90, u_{\max}^{(1)} = 3.652$; (c) The second one-peak positive solution with $J = 42.22, u_{\max}^{(2)} = 7.037$; (d) The third one-peak positive solution with $J = 159.0, u_{\max}^{(3)} = 13.63$. Its existence was mathematically verified after it was numerically computed for the first time; (e) A two-peak positive solution with $J = 53.12, u_{\max}^{(4)} = 7.037$; (f) A three-peak positive solution with $J = 212.5, u_{\max}^{(5)} = 13.78$.

![Figure 3](image-url)

Figure 3: A dumbbell-shaped domain (a) and 5 positive solutions (b)-(f) in Example 2.1.

**Example 2.2.** Let $r = 0$ and $\Omega = (-0.5, 0.5)^2$. 7 solutions are computed and shown in Figure 4, where the corner effects are observed if we compare to the solutions in Figure 5.
Example 2.3. Let $r = 0$ and $\Omega$ be the unit disk. 10 solutions are computed and shown in Figure 5. Without the corner effects, solution patterns are easy to classify. Except the radial solutions, all other solutions are degenerate.

Example 2.4. Let $\Omega$ be the unit disk and $r = 0.5, 3, 4, 6, 9$. Bifurcation takes place after $r > 0.5$, then positive solutions with more peaks appear and are computed, see Figure 6. Such a numerical bifurcation phenomenon was later mathematically verified [76].

Example 2.5. (A Steklov nonlinear boundary condition problem) [54]. Consider solve an elliptic PDE with nonlinear boundary condition of the form

$$(-\Delta + aI)u(x) = 0, x \in \Omega \text{ and } \frac{\partial u(x)}{\partial \nu} = g(x, u(x)), x \in \partial \Omega. \quad (2.5)$$
Its energy functional is $J(u) = \frac{1}{2} \int_\Omega (|\nabla u(x)|^2 + au^2(x))dx - \int_{\partial \Omega} G(x,u(x))d\sigma_x$, $u \in H^1(\Omega)$, where $\frac{\partial}{\partial u} G(x,u) = g(x,u)$. A Matlab package of LMM for numerically solving this type of Steklov problems can be downloaded at http://www.math.tamu.edu/~zhou/Steklov-eig.tar to use for research/education purposes under a proper acknowledgment.

Here we set $g(x,u(x)) = u^3(x), a = 1$ and $\Omega = \{(x_1, x_2) : \frac{x_1^2}{4} + \frac{x_2^2}{1} < 1\}$ an ellipse. 9 solutions are computed and shown in Figures 7-8 where $(J(u_1), ..., J(u_9)) = (0.5750, 0.7025, 1.7940, 2.7011, 4.0194, 6.0398, 7.1387, 10.7212, 11.4944)$. 

Figure 7: Profiles of solutions $u_1 \sim u_9$ on the boundary $\partial \Omega$ in Example 2.5.
2.3 Convergence Results

Let \( \{w^k\} = \{p(v^k)\} \) be the sequence generated by LMM with \( \varepsilon = 0 \) and \( J \) satisfy PS condition.

**Theorem 2.2.** If \( p \) is weak-continuous at \( v^k \) and \( w^k \not\in L \), then LMM is strictly decreasing.

**Theorem 2.3.** [53](Uniform Stepsize Rule) If \( p \) is a continuous peak selection of \( J \) near \( v_0 \in S_{L^\perp} \) s.t. \( J'(p(v_0)) \neq 0 \) and \( p(v_0) = t_0v_0 + v_0^L \not\in L \), then there are \( \delta > 0 \), \( s_0 > 0 \) s.t. when \( \|v - v_0\| < \delta, 0 < s < s_0 \), we have \( J(p(v(s))) - J(p(v)) < -\frac{t_0s\|J'(p(v))\|^2}{4C} \) where \( v(s) = \frac{v - sd}{\|v - sd\|} \in S_{L^\perp}, d = J'(p(v))/C, C = \max\{1, \|J'(p(v))\|\} \).

Next we present an improved convergence result [100].

**Theorem 2.4.** If \( p \) is continuous, \( d(L, w^k) > \alpha > 0 \) and \( \inf_{v \in S_{L^\perp}} J(p(v)) > -\infty \), then

(a) \( s_k d_k \to 0 \); (b) there is \( \{v^{k_i}\} \subset \{v^k\} \) s.t. \( v^{k_i} \to v^*, w^* = p(v^*), J'(w^*) = 0 \);
(c) if \( w^* \) is isolated then \( v^k \to v^* \).

**Proof.** By the stepsize rule (2.3), \( \frac{\sqrt{1+(s^k\|d^k\|)^2}}{\sqrt{2}} > \frac{1}{2} \) and the inequality

\[
\frac{s^k\|d^k\|}{\sqrt{1+(s^k\|d^k\|)^2}} \leq \|v^{k+1} - v^k\| \leq \frac{\sqrt{2}s^k\|d^k\|}{\sqrt{1+(s^k\|d^k\|)^2}},
\]

(2.6)
we get
\[ J(w^{k+1}) - J(w^k) \leq -\frac{1}{4}|t_0^k|s^k\|d^k\|^2C_k \leq -\frac{1}{8}|t_0^k||J'(w^k)||\|v^{k+1} - v^k\||. \]

Adding it up for all \(k = 1, 2,...\) and noting \(|t_0^k| > \alpha > 0, J(w^k) > -M > -\infty\), we obtain
\[
-\infty < \lim_{k \to \infty} J(w^k) - J(w^1) \leq -\frac{1}{4} \sum_{k=1}^{\infty} |t_0^k|s^k\|d^k\|^2C_k \leq -\frac{\alpha}{8} \sum_{k=1}^{\infty} \|J'(w^k)||\|v^{k+1} - v^k\||. \quad (2.7)
\]

Thus \(s^k\|d^k\|^2C_k = s^k\|J'(w^k)\|^2/C_k \to 0\). Then \(C_k = \max\{1, \|J'(w^k)\|\}\) and \(0 < s^k < \lambda\) lead to \(s^k\|d^k\| \to 0\) as in (a).

Next to prove (b), there are totally two cases, either (1) \(\|d^k\| > \eta > 0, \forall k = 1, 2,...\) for some \(\frac{1}{2} > \eta > 0\) or (2) there is a subsequence \(\{d^k_i\} \subset \{d^k\}\) s.t. \(d^k_i \to 0\).

In case of (1), \(\eta < \|d^k\| \leq 1\), then (2.7) becomes
\[
-\infty < \lim_{k \to \infty} J(w^k) - J(w^1) \leq -\frac{\alpha}{8} \sum_{k=1}^{\infty} \|J'(w^k)||\|v^{k+1} - v^k\| \leq -\frac{\alpha\eta}{8} \sum_{k=1}^{\infty} \|v^{k+1} - v^k\|, \quad (2.8)
\]
i.e., \(\{v^k\} \subset S_{L^\perp}\) is a Cauchy sequence in the Hilbert space \(H\). Thus there exists \(v^* \in S_{L^\perp}\) s.t. \(v^k \to v^*\) as \(k \to \infty\). Since \(p\) is continuous and \(J\) is \(C^1\), we have \(w^* = p(v^*) = \lim_{k \to \infty} p(v^k)\) and \(\|J'(v^*)\| \geq \eta\). Hence \(w^*\) is not a critical point. Then \(s^kd^k \to 0\) and \(\|d^k\| > \eta > 0\) imply \(s^k \to 0\), which contradicts the uniform stepsize rule, Theorem 2.3. Thus case (2) must hold, i.e., there is a subsequence \(\{d^{k_i}\} \subset \{d^k\}\) s.t. \(d^{k_i} \to 0\). Since \(\{J(w^{k_i})\}\) is bounded, by the PS condition, there is a subsequence denoted by \(w^{k_i} = p(v^{k_i}) = t_0^{k_i}v^{k_i} + u^{k_i}\) again for some \(u^{k_i} \in L, \) s.t. \(w^{k_i} \to v^* = w_1^* + w_2^*\) with \(w_1^* \in L^\perp, w_2^* \in L\). By the condition \(|t_0^{k_i}| > \alpha > 0\), we must have \(t_0^{k_i}v^{k_i} \to w_1^*, u^{k_i} \to w_2^*\) and \(v^{k_i} \to v^*\) for some \(v^* \in S_{L^\perp}\). Thus \(t_0^{k_i} \to t_0^*\) for some \(|t_0^*| \geq \alpha > 0\). The continuity of \(p\) then leads to \(w^* = p(v^*)\). We have proved that \(p\) is a homeomorphism and \(J'(w^k) = \lim_{i \to \infty} J'(w^{k_i}) = \lim_{i \to \infty} d^{k_i} = 0\). So (b) is proved.

Finally to prove (c) \(v^k \to v^*\) if \(w^* = p(v^*)\) is an isolated saddle, since \(p\) is a homeomorphism, this means that there is \(\delta > 0\) s.t. there is no point \(v' \in S_{L^\perp}\) satisfying \(\|v^* - v'\| < \delta\) and \(J'(p(v')) = 0\). Let \(I \subset N = \{1, 2,...\}\). We call \(\sum_{i \in I} \|v^{i+1} - v^i\|\) the total distance traveled by the subsequence \(\{v^i\}_{i \in I}\). For any \(\eta > 0\), let \(i \in I \subset N\) denote the whole index set in \(N\) with \(\|d^k\| > \eta\). Then (2.7) leads to
\[
-\infty < \lim_{i \to \infty} J(v^i) - J(w^i) \leq -\frac{\alpha}{8} \sum_{i \in I} \|J'(w^i)||\|v^{i+1} - v^i\| \leq -\frac{\alpha\eta}{8} \sum_{i \in I} \|v^{i+1} - v^i\|, \quad (2.9)
\]
i.e., the total distance traveled by \(\{v^i\}_{i \in I}\) is finite. Note that by (a), we have
\[
\|v^{k+1} - v^k\| \leq \frac{\sqrt{2}s^k\|d^k\|}{\sqrt{1 + (s^k\|d^k\|)^2}} \to 0. \quad (2.10)
\]
Suppose there is $\delta_3 > 0$ s.t. there are infinitely many points $v$ in $\{v_k\}$ with $\|v - v^*\| > \delta_3$. By the inequality (2.10), for any $0 < \delta_1 < \delta_2 < \delta_3$, there is $K > 0$ s.t. when $k > K$, $\|v^{k+1} - v^k\| < \frac{1}{2}(\delta_2 - \delta_1)$. Since $v^*$ is a limit point of $\{v^k\}$, there are infinitely many points $v \in \{v^k\}$ s.t. $0 < \|v - v^*\| < \delta_1$ and there are also infinitely many points $\{v^{k_i}\} \subset \{v^k\}$ s.t. $0 < \delta_1 < \|v^{k_i} - v^*\| < \delta_2$, i.e., the sequence $\{v^k\}$ enters the three ring regions centered at $v^*$ and defined by $0 < \delta_1 < \delta_2 < \delta_3$ infinitely many times. Thus the total distance traveled by such $\{v^k\}$ has to be infinite. However by (2.9), for any $\eta > 0$, the total distance traveled by all the points $v^{k_i} \in \{v^k\}$ with $\|J'(p(v^{k_i}))\| > \eta$ is finite. Thus there must be infinitely many points $\{v^{k_i}\} \subset \{v^k\}$ s.t. $J'(p(v^{k_i})) \to 0$. By the PS condition, there is a subsequence, denoted by $\{w^{k_i}\} = \{p(v^{k_i})\}$ again, s.t. $w^{k_i} \to w' = t'v' + w'_L$ for some $t' > 0, v' \in S_{L^1}, w'_L \in L$ with $J'(w') = 0$ and $\delta_1 \leq \|v^* - v'\| \leq \delta_2$. Since $0 < \delta_1 < \delta_2$ can be any numbers less than $\delta_3$, this contradicts to the assumption that $w^* = p(v^*)$ is an isolated critical point. Thus for any $\delta_3 > 0$, there can be at most a finite number of points $v$ in $\{v^k\}$ with $\|v^* - v\| > \delta_3$, i.e., $v^k \to v^*$. (c) is proved. □

2.4 Instability Analysis of Saddles by LMM [98]

Classical calculus of variations and numerical methods focus on finding a unique stable solution. Saddles correspond to unstable local equilibria in physical process. When multiple saddles are concerned, can we measure their instabilities in certain order? The answer is yes, it can be measured by the saddle order. However the saddle order lacks of characterization and is difficult to obtain. Thus people use the Morse index (MI) as the saddle order to do solution instability analysis. But MI is still very expensive to obtain. Then many researchers tried to find ways to obtain a bound estimation of MI for a solution satisfying certain conditions [4,5,7,16]. Here we present

**Theorem 2.5.** [98] If $p$ is a local peak selection differentiable at $v^* \in S_{L^1}$, $u^* = p(v^*) \notin L$ and $v^* = \arg\min_{v \in S_{L^1}} J(p(v))$, then $u^*$ is a critical point with

$$\dim(L) + 1 = \text{MI}(u^*) + \dim(H^0 \cap [L, v^*]).$$

(2.11)

**Remark 2.2.** When $u^*$ is nondegenerate, we have $\text{MI}(u^*) = \dim(L) + 1$. This number is known in LMM before $u^*$ is computed. Usually $\text{MI}(u^*)$ is computed by two steps: (1) find $u^*$ and (2) compute $\text{MI}(u^*) = \dim(H^-)$ where $H^-$ is the maximum negative definite subspace of $J''(u^*)$ in a Hilbert space $H$, thus very expensive to compute. Here we reverse the process, LMM utilizes the mathematical properties of a solution $u^*$ with $\text{MI}(u^*) = \dim(L) + 1$ to numerically find a saddle $u^*$ with such MI, i.e., LMM finds multiple solutions following the order of $\dim(L) + 1$.  

13
2.5 On Symmetry Invariance

It is proved in [81] that LMM is invariant to symmetry. But such an invariance is sensitive to numerical errors, i.e., when numerical errors dominate the magnitude of the gradient $J'(\cdot)$, a symmetry may be broken depending on if or not the support $L$ is sufficient. On the other hand, a Newton method is fast but strongly depends on an initial guess. Since a Newton method does not assume any variational structure, it is blind to the saddle order. It is also proved in [21,82] that a Newton method is invariant to symmetry and such an invariance is insensitive to numerical errors. Consequently it traps a symmetry possessed by an initial guess. Hence the symmetry of an initial guess must match that of an unknown targeting solution. Otherwise a Newton method fails to find the targeting solution. In summary, when dealing with a multiple solution problem, a Newton method is fast but strongly depends on an initial guess, it is easy to find a known solution but it alone is difficult to find an unexpected solution; while LMM is slower, its dependence on an initial guess is loose, thus it is easy to find an unexpected solution. Thus the best way is to apply LMM first to get close to a solution following the saddle order then it is followed by a Newton method to speed up local convergence. A Matlab package of LMM-Newton method can be downloaded at http://www.math.tamu.edu/~jzhou/minimax.tar to use for research/education purposes under a proper acknowledgment. To better understand the symmetry invariance of LMM and a Newton method, let us solve the Henon equation in Example 2.4 with $r = 2$. Since it is a nonautonomous PDE, the wellknown Gidas-Ni-Nirenberg theorem for symmetry does not apply. Note that since $r > 0.5$, bifurcation takes place. We use a positive symmetric initial guess and then apply LMM for $N_{LMM} = 4, 14, 20$ times before it is followed by a Newton method to speed up its local convergence. We obtain three different positive solutions as shown in Figure 9, where the symmetric solution in (a) generates all other asymmetric positive solutions in (b) and (c) with lower MI due to the bifurcation. Positive solutions with more peaks may appear if $r$ further increases. This example deserves a special attention for all the people using a Newton method alone to solve a multiple solution problem.

![Figure 9](image_url)

Figure 9: Positive solutions to Example 2.4 with $r = 2$. (a) $N_{LMM} = 4$; (b) $N_{LMM} = 14$ and (c) $N_{LMM} = 20$, a ground state.
3 A Local Min-⊥ Method [99]

In the above results, \( p \) is assumed to be continuous. However since \( u^* = \lim_{k \to \infty} p(v^k) \) need not be a local maximum, it is actually improper to discuss its continuity. On the other hand, for the W-type problems, see Figure 2, the current LMM is not working. We need a more general mathematical framework. Note that when \( p \) is a peak selection, we have \( J'(p(v)) \perp [L, v] \), \( \forall v \in S_{L\perp} \). We define

**Definition 3.1.** \( p: S_{L\perp} \to 2^H \) is the \( L \perp \) mapping if \( P(v) := \{ u \in [L, v] : J'(u) \perp [L, v] \} \) for each \( v \in S_{L\perp} \). \( p: S_{L\perp} \to H \) is an \( L \perp \) selection if \( p(v) \in P(v) \forall v \in S_{L\perp} \). If \( p \) is locally defined then \( p \) is a local \( L \perp \) selection.

It is clear that \( p \) is a peak selection implies \( p \) is also an \( L \perp \) selection. If \( J \) is \( C^1 \), \( v^k \to v^* \) and \( u^k = p(v^k) \in [L, v^k] \to u^* \), then \( J'(u^k) \perp [L, v^k] \) implies \( J'(u^*) \perp [L, v^*] \) and \( u^* \in [L, v^*] \), i.e., \( u^* = p(v^*) \). The set \( \{ P(v) : v \in S_{L\perp} \} \) is later called the generalized Nehari manifold in the literature.

**Lemma 3.1.** [99] If \( J \) is \( C^1 \), then \( G = \{ (u, v) : v \in S_{L\perp}, u \in P(v) \} \) is closed.

Thus when \( p \) is a local \( L \perp \) selection, its continuity or differentiability can be discussed and the previous results on stepsize rule and convergence can be similarly established. But one can only get a low bound estimate of \( \text{MI}(u^*) \) [22]. We focus on the algorithm aspect.

**Theorem 3.1.** [99] If \( p \) is continuous at \( v^* \in S_{L\perp} \) and \( p(v^*) \not\in L \), then \( u^* = p(v^*) \) is a critical point of \( J \) if and only if there is \( N(v^*) \) s.t.

\[
J'(p(v^*)) \perp p(v) - p(v^*), \quad \forall v \in N(v^*) \cap S_{L\perp}. \tag{3.12}
\]

**Remark 3.1.** In the solution characterization (3.12), there is no \( J(\cdot) \)-value involved except \( A(\cdot) = J'(\cdot) \). It implies a useful result to solve non-variational problems \( A(u) = 0 \) for multiple solutions, see Section 5. Here is a methodology for finding multiple solutions with a composite function \( J(v) = J(p(v)) \), we design a two-level method: the outer-loop: find \( v^* \) s.t. \( J'(p(v^*)) \perp L^\perp \), e.g., by local-min in LMM; while the inner-loop: \( p(v) \in [L, v] \setminus L \), \( J'(p(v)) \perp [L, v] \), which shows its flexibility to provide a separation of a new solution from previously found solutions contained in \( L \). By using local min/max, many different ways can be explored to find multiple solutions to different types, e.g., the W-type, problems.

### 3.1 Differentiability of an \( L \perp \) Selection \( p \)

To check if \( p \) is differentiable, let \( L = [w_1, w_2, ... , w_n] \) and \( v \in S_{L\perp} \). By the definition, \( p(v) = w = t_0v + t_1w_1 + ... + t_nw_n \) is solved from \( (n+1) \) orthogonal conditions

\[
p(v) = w = t_0v + t_1w_1 + ... + t_nw_n
\]
\langle J'(t_0v + t_1w_1 + \ldots + t_nw_n), v \rangle = 0 \text{ and } \langle J'(t_0v + t_1w_1 + \ldots + t_nw_n), w_j \rangle = 0 \text{ for } j = 1, \ldots, n.

By the implicit function theorem, if the 
\begin{equation}
Q'' = \begin{bmatrix}
\langle J''(w)v, v \rangle, & \langle J''(w)w_1, v \rangle, & \ldots & \langle J''(w)w_n, v \rangle \\
\langle J''(w)v, w_1 \rangle, & \langle J''(w)w_1, w_1 \rangle, & \ldots & \langle J''(w)w_n, w_1 \rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle J''(w)v, w_n \rangle, & \langle J''(w)w_1, w_n \rangle, & \ldots & \langle J''(w)w_n, w_n \rangle
\end{bmatrix}
\end{equation}
is nonsingular or \(|Q''| \neq 0\), then \(p\) is differentiable near \(v\). This can be easily checked.

\section{Several Important Directions}

\subsection{Find multiple saddles in Banach spaces [91,92,93,94]}

When non Newtonian/Darcian fluids/materials are considered, Darcian’s law is replaced by other laws, e.g., the Laplacian \(\Delta\) is replaced by the \(p\)-Laplacian \(\Delta_p u(x) = \nabla \cdot |\nabla u(x)|^{p-2} \nabla u(x)\) \((p > 1)\) where \(p = 2, p < 2, p > 2\) if the fluid/material is Newtonian/Darcian, pseudoplastic, dilatant, respectively. \(\Delta_p\) also appears in the study of flow in porous media \((p = \frac{3}{2})\), non-linear elasticity \((p > 2)\) and glaciology \((p \in (1, \frac{4}{3}))\) [29]. Then solitary wave solutions to a "Schroedinger flow" lead to solve a quasilinear elliptic (Eigen) PDE in \(W^{1,p}(\Omega)\)

\begin{equation}
\lambda u(x) = -\Delta_p u(x) + \kappa f(x, |u(x)|)u(x),
\end{equation}

whose energy functional is

\begin{equation}
J(u) = \frac{1}{p} \int_{\Omega} |\nabla u(x)|^p - \frac{1}{2} \lambda u^2(x) + \kappa F(x, u(x))\}dx
\end{equation}

\end{equation}

where \(F_p^\ast(x, u) = f(x, |u|)u\). Since \(W^{1,p}(\Omega)\) for \(p \neq 2\) is a Banach space but not a Hilbert space, \(J'(u) \notin W^{1,p}(\Omega)\) cannot be used as an algorithm search direction. One needs a pseudo-gradient to replace the gradient. Let \(B\) be a Banach space and \(J \in C^1(B, \mathbb{R})\).

\begin{definition}
Let \(u \in B\) be a point s.t. \(J'(u) \neq 0\). For given \(\theta \in (0,1]\), a point \(\Psi(u) \in B\) is a pseudo-gradient of \(J\) at \(u\) w.r.t. \(\theta\) if \(\|\Psi(u)\| \leq 1\) and \(\langle J'(u), \Psi(u) \rangle \geq \theta \|J'(u)\|\).
\end{definition}

Pseudo-gradients are commonly used in numerical optimization to find a local minimum of a functional \(J\) on a Banach space. However, when saddles are concerned, such pseudo-gradients do not work, since they will lead to a local minimum. To prevent such a degeneracy, we need a new notion called \textit{projected pseudo-gradient} (PPG).

\begin{lemma}
Let \(B = L \oplus L'\), \(0 < \theta < 1\), \(v_0 \in S_{L'}\) be given. If \(p\) is a local peak selection of \(J\) w.r.t. \(L\) at \(v_0\) s.t. \(J'(p(v_0)) \neq 0\), then there is a PPG \(\Psi(p(v_0))\) of \(J\) at \(p(v_0)\) w.r.t. \(\theta\) s.t.
\end{lemma}
(a) $\Psi(p(v_0)) \in L'$, $0 < \|\Psi(p(v_0))\| \leq M$ where $M \geq 1$ is the bound of the linear projection $\mathcal{P}$ from $B$ to $L'$;
(b) $\langle J'(p(v_0)), \Psi(p(v_0)) \rangle \geq \theta \|J'(p(v_0))\|$

It is interesting to note that such a PPG can be constructed only at a point $u = p(v)$. When the gradient of $J$ is replaced by PPG of $J$ in LMM, results on stepsize rule, LMMP, uniform stepsize rule and convergence parallel to the previous ones can be established [94]. Thus LMM can be modified to numerically find multiple saddles in Banach spaces.

### 4.2 Constrained nonlinear eigen-function problem

Let $F, G : B \to \mathbb{R}$ be $C^1$ functions. For given $\alpha \in \mathbb{R}$, consider finding eigen-pairs $(\lambda, u) \in \mathbb{R} \times (B \setminus \{0\})$ s.t.

$$F'(u) = \lambda G'(u), \quad \text{s.t. } G(u) = \alpha.$$ 

The problem is homogeneous if $F'(tu) = t^m F'(u), G'(tu) = t^\ell G'(u), t > 0$ for some $m, \ell > 0$.

**Case 1.** Iso-homogeneous cases $m = \ell$ [92]. Define the Raleigh quotient $\mathcal{R}(u) = \frac{F(u)}{G(u)}$.

Since $\mathcal{R}(tu) = \mathcal{R}(u)$ for any $t \neq 0$, the problem is degenerate everywhere. We have

$$\mathcal{R}'(u) = 0 \iff F'(u) = \lambda G'(u), \quad \left(\lambda = \mathcal{R}(u) = \frac{F(u)}{G(u)}\right).$$

Thus it becomes a problem to find multiple saddles of $\mathcal{R}$ in an order. LMM can be applied to do so. The order of eigen-functions coincides with the order of $\dim(L)$ [92].

**Case 2.** Non-iso-homogeneous, non-homogeneous cases [93]. Define its Lagrange functional

$$\mathcal{L}(\lambda, u) = F(u) - \lambda (G(u) - \alpha),$$

which may fail to have a mountain-pass structure. So we introduce a new *active Lagrange* functional $\mathcal{L}(\lambda, u) = F(u) - \ell(\lambda)(G(u) - \alpha)$ where $\ell(\lambda)$ is an *active Lagrange multiplier* whose selection is to let $\mathcal{L}(\lambda, u)$ have a mountain-pass structure, e.g., $\ell(\lambda) = |\lambda|^k, k = 2$ and to make the algorithm converge, see [93].

**Example 4.1.** (Eigen-pairs of $\Delta_p$) Consider finding eigen-pairs $(\lambda, u) \in \mathbb{R} \times (B \setminus \{0\})$ s.t.

$$-\Delta_p u(x) = \lambda |u(x)|^{q-2} u(x), u \in W_0^{1,p}(\Omega), \quad \Omega = (0, 2)^2.$$ 

**Case (a).** $p = q = 1.75 < 2$, we obtain $\lambda_1 = 4.2458, \lambda_2 = 9.3173, \lambda_3 = 9.4078$,

$\lambda_4 = 14.2805, \lambda_5 = 16.837, \lambda_6 = 17.2546, \lambda_7 = 23.3660$.

**Case (b).** $p = q = 2.5 > 2$, we obtain $\lambda_1 = 6.3547, \lambda_2 = 20.2896, \lambda_3 = 20.79854$,

$\lambda_4 = 35.9448, \lambda_5 = 48.2598, \lambda_6 = 49.6794, \lambda_7 = 51.1048$.

Observe the pattern order of eigen-solutions in Figures 10-11. We have done many computations, the pattern order remains the same as Figure 10 for all $1 < p < 2$ and as Figure 11 for
all \( p > 2 \). Thus when the \( p \)-value crosses 2, some pattern order switches, e.g., the second and the third eigen-solutions in Figures 10-11, etc. It indicates that the second eigen-value of
\[
-\Delta = -\Delta_p(p = 2)
\]
is a double eigen-value, see Figure 12. To separate the double eigen-value, one has to change the physical nature from Newtonian to non-Newtonian.

Figure 10: The first 7 eigenfunctions of \( \Delta_p \) in Case (a) \( 1 < p < 2 \).

Figure 11: The first 7 eigenfunctions of \( \Delta_p \) in Case (b) \( p > 2 \).

4.3 Multiple Solutions to Systems

When a nonlinear process involves multi-bodies (particles, molecules, species, etc.), it leads to a system. Comparing to single equations, nonlinear systems are much richer in varieties and complexities, and can be classified in many different ways.

We focus on developing computational theory and methods for finding multiple solutions
to a class of semilinear elliptic systems of the form [28,39,38,48]

\[
\begin{align*}
-\Delta u(x) &= F'_1(u(x), v(x), x), \quad x \in \Omega, \\
-\Delta v(x) &= G'_2(u(x), v(x), x), \quad x \in \Omega
\end{align*}
\]

(4.2)

with zero Dirichlet or Neumann B.C., where \( F'_1, G'_2 \) are their partial Frechet derivatives w.r.t the first and the second variables. In general such a nonlinear system is not variational. It has two energy functionals \( f, g \)

\[
\begin{align*}
f(u, v) &= \int_{\Omega} \left[ \frac{1}{2} |\nabla u(x)|^2 - F(u(x), v(x), x) \right] dx, \\
g(u, v) &= \int_{\Omega} \left[ \frac{1}{2} |\nabla v(x)|^2 - G(u(x), v(x), x) \right] dx,
\end{align*}
\]

(4.3)

where \( F, G \in C^1(\mathbb{R}^2 \times \Omega, \mathbb{R}) \) usually contain higher order terms satisfying some growth conditions and \( u \) and \( v \) interact each other in many different ways on energy profiles \( f, g \), e.g., strongly coupled, weakly coupled, cooperative, non-cooperative, etc. Here we focus on two variational classes, namely, cooperative systems and non-cooperative systems.

A Cooperative system and its energy functional

\[
-\Delta u(x) - \lambda u(x) = F'_u(u, v), \quad -\Delta v(x) - \mu v(x) = F'_v(u, v).
\]

\[
J(u, v) = \int_{\Omega} \left[ \frac{1}{2} (|\nabla u(x)|^2 - \lambda u^2(x) + |\nabla v(x)|^2 - \mu v^2(x)) - F(u(x), v(x)) \right] dx.
\]

(4.4)

A Non-cooperative system and its energy functional (strongly indefinite)

\[
-\Delta u(x) - \lambda u(x) = F'_u(u, v), \quad -\Delta v(x) - \mu v(x) = -F'_v(u, v).
\]

\[
J(u, v) = \int_{\Omega} \left[ \frac{1}{2} (|\nabla u(x)|^2 - \lambda u^2(x) - |\nabla v(x)|^2 + \mu v^2(x)) - F(u(x), v(x)) \right] dx.
\]

(4.5)

We are interested in finding co-existing solutions \((u, v)\) with \( u \neq 0, v \neq 0 \).
In [20,22,23] a game approach is developed to study the system (4.2) where the two components \( u \) and \( v \) are viewed as two players in a game with their objective functionals \( f \) and \( g \) in (4.3). Thus \( (u^*, v^*) \) is a solution to system (4.2) if \( (u^*, v^*) \) solves the system

\[
\begin{align*}
{f_u}'(u,v) &= 0 \quad \text{and} \quad {g_v}'(u,v) = 0,
\end{align*}
\]

which shows certain partial variational structure. We see that a solution \( (u^*, v^*) \) to

\[
\begin{align*}
\left\{ \begin{array}{l}
u^* = \arg \min_{u \in \mathcal{N}(u^*)} f(u, v^*), \\
v^* = \arg \min_{v \in \mathcal{N}(v^*)} g(u^*, v).
\end{array} \right.
\]

satisfies (4.6). Of course, a min in (4.7) can also be replaced by a max. For the cooperative system, \( f, g \) in (4.3) and \( J \) in (4.4) have the relation

\[
f(u, v) = J(u, v) - \alpha(v), \quad g(u, v) = J(u, v) - \beta(u),
\]

where

\[
\alpha(w) = \beta(w) = \int_{\Omega} \frac{1}{2} |\nabla w(x)|^2 dx.
\]

Since for Player \( u \), \( \alpha(v) \) is a constant and for Player \( v \), \( \beta(u) \) is a constant, the main objectives of Players \( u \) and \( v \) are the same, to minimize \( J(u, v) \). So the nature of a cooperative system is indeed a cooperative game.

For the non-cooperative game, \( f, g \) in (4.3) and \( J \) in (4.5) have the relation

\[
f(u, v) = J(u, v) - \alpha(v), \quad g(u, v) = -J(u, v) + \beta(u),
\]

where \( \alpha(\cdot), \beta(\cdot) \) are given in (4.9). So the main objectives of Player \( u \) and \( v \) are in opposite direction, i.e., Player \( u \) wants to minimize \( J(u, v) \) but Player \( v \) wants to maximize \( J(u, v) \), or a solution \( (u^*, v^*) \) to (4.7) coincides with a local saddle point of the form

\[
J(u^*, v) \leq J(u^*, v^*) \leq J(u, v^*), \quad \forall u \in \mathcal{N}(u^*), v \in \mathcal{N}(v^*).
\]

So the non-cooperative system is actually a local zero-sum game (reversed-cooperative).

According to game theory, a non-cooperative game is an exclusive notion to cooperative games. The reveal of this reversed-cooperative nature enables us to develop a new method in [23] to solve this type of problems for multiple solutions in an order.

Let \( H_1 \) and \( H_2 \) be Hilbert spaces, \( H = H_1 \times H_2 \), \( J \in C^1(H, \mathbb{R}) \), \( L_1 \subset H_1 \) and \( L_2 \subset H_2 \) be closed subspaces. Denote \( L = L_1 \times L_2 \).

**Definition 4.2.** A set-valued mapping \( P: S_{L^\perp} \to 2^H \) is called an \( L^\perp \) mapping of \( J \) if

\[
P(\bar{w}) = \left\{ (u,v) \in [L_1, \bar{u}] \times [L_2, \bar{v}] : J_u'(u, v) \perp [L_1, \bar{u}], J_v'(u, v) \perp [L_2, \bar{v}] \right\}, \forall \bar{w} = (\bar{u}, \bar{v}) \in S_{L^\perp}.
\]

A mapping \( p: S_{L^\perp} \to H \) is an \( L^\perp \) selection of \( J \) if \( p(w) \in P(w), \forall w \in S_{L^\perp} \). For \( w \in S_{L^\perp}, \) if \( p \) is locally defined near \( w \), then \( p \) is called a local \( L^\perp \) selection of \( J \) at \( w \).
Remark 4.1. (a) Definition 4.2 is bi-$\perp$ and stronger than the original one, since

$$J'_1 \perp [L_1, u], \ J'_2 \perp [L_2, v] \Rightarrow J' = (J'_1, J'_2) \perp [L, (u, v)], \ \forall (u, v).$$

It enables us to identify and capture the co-existing states, also gives us more flexibility in solving other nonlinear systems, e.g., multi-component systems.

(b) The bi-$\perp$ condition $(u^*, v^*) = p(\bar{u}, \bar{v})$ can be solved simultaneously from

$$u^* = \arg \max_{u \in [L_1, \bar{u}]} J(u, v^*), \quad v^* = \arg \min_{v \in [L_2, \bar{v}]} J(u^*, v).$$

Let $E_1^+, E_1^- \subset H_1, E_2^+, E_2^- \subset H_2$ where $E_1^+ \times E_2^+, E_1^- \times E_2^-$ are resp. m.p.d. and m.n.d. subspaces of $J''(0, 0)$ where $E_1^-$ is finite dimensional.

Cooperative Case: \text{dim}(E_2^-) < \infty. Let $L_1 \subset H_1$ contain $E_1^-$ and $L_2 \subset H_2$ contain $E_2^-$ be finite dimensional subspaces. We have a cooperative game characterization:

$$u^* = \arg \min_{u \in S_{L_1^+}} J(p(\frac{(u, v^*)}{\| (u, v^*) \|})) \text{ and } v^* = \arg \min_{v \in S_{L_2^+}} J(p(\frac{(u^*, v)}{\| (u^*, v) \|})), \quad (4.12)$$

It follows

$$J'_1(p(\frac{(u^*, v^*)}{\| (u^*, v^*) \|})) \perp L_1^+ \quad \text{and} \quad J'_2(p(\frac{(u^*, v^*)}{\| (u^*, v^*) \|})) \perp L_2^+ \Rightarrow J'(p(\frac{(u^*, v^*)}{\| (u^*, v^*) \|})) = 0.$$

Numerically we use negative gradients $-J'_1(\cdot), -J'_2(\cdot)$ to approximate the mins in $(4.12)$ with stepsize rules given as below: for each $(u, v) \in S_{L^+}$, denote the updates

$$u(s) = u - sJ'_1(p(u, v)) \quad \text{and} \quad v(t) = v - tJ'_2(p(u, v))$$

where $(s, t)$ is determined by the stepsize rules

$$J(p(\frac{(u(s), v)}{\| (u(s), v) \|})) - J(p(u, v)) \leq -\frac{1}{2}s_s\| J'_1(p(u, v)) \|^2,$$

$$J(p(\frac{(u, v(t))}{\| (u, v(t)) \|})) - J(p(u, v)) \leq -\frac{1}{2}t_v\| J'_2(p(u, v)) \|^2.$$

Modify LMM with the new stepsize rules, one can compute multiple solutions to the cooperative system. Also convergence of LMM can be similarly established and the equality $\text{MI}(u^*) + \text{dim}(H^0 \cap \bigcap_{i=1}^2 [L_i, u_i]) = \text{dim}(L) + 2$ is proved for instability analysis of a saddle $u^* = (u_1, u_2)$ computed by LMM to such a cooperative system [22].

Noncooperative Case: \text{dim}(E_2^+) < \infty. Let $L_1 \subset H_1$ contain $E_1^+, L_2 \subset H_2$ contain $E_2^+$ be finite dimensional subspaces. We have a zero-sum-like game characterization for noncooperative systems:

$$u^* = \arg \min_{u \in S_{L_1^+}} J(p(\frac{(u, v^*)}{\| (u, v^*) \|})) \text{ and } v^* = \arg \max_{v \in S_{L_2^+}} J(p(\frac{(u^*, v)}{\| (u^*, v) \|})), \quad (4.13)$$
\[ J(p(\frac{(u^*, v^*)}{\|u^*, v^*\|})) \leq J(p(\frac{(u^*, v^*)}{\|u^*, v^*\|})) \leq J(p(\frac{(u^*, v^*)}{\|u^*, v^*\|})) \forall (u, v) \in \mathcal{N}(u^*, v^*) \cap S_{L^1}. \quad (4.14) \]

It follows

\[ J'_1(p(\frac{(u^*, v^*)}{\|u^*, v^*\|})) \perp L^1_1 \quad \text{and} \quad J'_2(p(\frac{(u^*, v^*)}{\|u^*, v^*\|})) \perp L^1_2 \Rightarrow J'(p(\frac{(u^*, v^*)}{\|u^*, v^*\|})) = 0. \]

Numerically we use gradients \(-J'_1(\cdot)\) and \(J'_2(\cdot)\) to approximate the min and max in (4.13) with steps size rules given as below: for each \((u, v) \in S_{L^1}\), denote the updates

\[ u(s) = u - sJ'_1(p(u, v)) \quad \text{and} \quad v(t) = v + tJ'_2(p(u, v)) \]

where \((s, t)\) is determined by the steps size rules

\[
\begin{align*}
J(p(\frac{(u(s), v)}{\|(u, v)\|})) - J(p(u, v)) & \leq -\frac{1}{2}s_u \|J'_1(p(u, v))\|^2, \\
J(p(\frac{(u, v(t))}{\|(u, v(t))\|})) - J(p(u, v)) & \geq \frac{1}{2}t_v \|J'_2(p(u, v))\|^2.
\end{align*}
\]

Modify LMM with the new steps size rules, one can compute multiple solutions to the non-cooperative system. But its instability and convergence analysis are much more difficult.

**Example 4.2.** A cooperative system in nonlinear optics with multiple vector solitons [20].

\[
\begin{align*}
-\Delta u(x, y) &= -u(x, y) + \frac{u^2(x, y) + v^2(x, y)}{1 + \mu(u^2(x, y) + v^2(x, y))} u(x, y), \\
-\Delta v(x, y) &= -\gamma v(x, y) + \frac{u^2(x, y) + v^2(x, y)}{1 + \mu(u^2(x, y) + v^2(x, y))} v(x, y),
\end{align*}
\]

with zero Dirichlet B.C. where \(\Omega = (-10, 10)^2\), \(\gamma = 0.65\), \(\mu = 0.5\). Dipole- and multipole-mode vector solitons to (4.15) are computed and shown in Figure 13.

**Example 4.3.** A noncooperative system [22]:

\[
\begin{align*}
-\Delta u(x, y) &= \lambda u(x, y) - \delta v(x, y) + |u(x, y)|^{p-1}u(x, y), \\
-\Delta v(x, y) &= \delta u(x, y) + \gamma v(x, y) - |v(x, y)|^{q-1}v(x, y),
\end{align*}
\]

with zero Dirichlet B.C. where \(\Omega = (-1, 1)^2\). We choose \(p = q = 3, \lambda = \gamma = -0.5, \delta = 5\). 12 co-existing solutions are computed and shown in Figure 14 (a)-(l) where (a) \(J=1.1226\) (b) \(J=2.7412\) (c) \(J=2.7694\) (d) \(J=3.0446\), (e) \(J=3.1017\) (f) \(J=3.9376\) (g) \(J=4.1982\) (h) \(J=5.2226\), (i) \(J=5.3454\) (j) \(J=7.9394\) (k) \(J=8.4786\) (l) \(J=9.1233\).
Figure 13: Vector solitons to (4.15). Dash lines indicate the nodal lines.
Figure 14: 12 co-existing solution pairs (a)-l to Example 4.3.
5 Modified Newton Method for Multiple Solutions

The following facts on a Newton method are known:

(1) It does not assume a variational structure. Thus it can solve more general problems, but is blind to a saddle order and cannot tell the instability of saddles;

(2) It is locally fast convergent but strongly depends on an initial guess, which severely reduces its effectiveness to find multiple solutions. It is a consensus that for multiple solution problems, the complexity of obtaining an initial guess sufficiently close to an unknown solution is almost the same as that of finding the solution. Thus a Newton method alone is not suitable for solving such multiple solution problems, unless strategies to select initial guesses are designed and justified;

(3) Some Newton homotopy continuation methods (NHCM) are proposed to relieve its dependence on an initial guess [1]. However it is proved in [21] that a NHCM will help but cannot resolve the difficulty, i.e., an initial guess still has to be selected in the same basin of an unknown targeting solution;

(4) It is invariant to symmetry and such an invariance is insensitive to numerical errors [21,82], i.e., it traps a symmetry possessed by an initial guess. So it is double edged. If the symmetry of a solution is known, a Newton method can always find the solution by selecting an initial guess with the same symmetry. Otherwise it fails if the symmetry of an initial guess does not match that of a desired solution;

(5) Authors in [18,19,63] proposed to use eigen-functions of its linear part in the problem as an initial guess for a Newton method to find multiple solutions. But it works only for autonomous PDE and single eigen-function. Example 2.4 and Figures 6 and 9 showed so. When a linear operator has an eigen-value of multiplicity, any linear combination of its eigen-functions is still an eigen-function with a different symmetry. But in general the superposition principle does not hold for a nonlinear operator, it may have no solution with such a different symmetry. Once an eigen-function with a different symmetry is used as an initial guess, the Newton method fails;

(6) Authors in [88] proposed to use a linear approximation and bifurcation method to identify the symmetries of its eigen-functions and then use them as initial guesses in a Newton method. It can treat non autonomous PDEs. However other difficulties in (4) still remain, e.g., how to handle the case where two saddles have exactly the same symmetry?

Motivated by the local min-$\perp$ method [98] to use the information provided by previously found solutions, augmented singular transforms are introduced in [86,54] to change the local barrier structure of the original problem for finding new solutions.

Let $U$ be a Banach space, $V = U^*$, $\langle \cdot, \cdot \rangle$ be their duality relation and $\perp$ be defined by $\langle \cdot, \cdot \rangle = 0$. For $F : U \to V$, consider solving $F(u) = 0$ for multiple solutions. Assume $F(0) = 0$ and $S$ be the subspace spanned by previously found linearly independent solutions.
\( u_1, \ldots, u_k \in U \cap V. \) Denote \( S^\perp = \{ v \in U : \langle v, u \rangle = 0, \forall u \in S \} \) and \( S^+_1 = \{ u \in S^\perp : \| u \| = 1 \}. \)

For each \( u \in S^+_1 \) and \( t \neq 0 \), we define an augmented singular transform (AST)

\[
G(u, t, t_1, \ldots, t_k) = \frac{1}{t} F(t(u + \sum_{i=1}^k t_i u_i)). \tag{5.1}
\]

Note that \( G \) has a singularity at \( t = 0 \). It forms a barrier to separate \( U \setminus S \) from 0, since any \( u^* \notin S \) can be expressed as \( u^* = t(u + t_1 u_1 + \cdots + t_k u_k) \) for some \( u \in S^+_1 \) and constants \( t, t_1, \ldots, t_k \), and we have \( u^* \in (U \setminus S) \) if \( t \neq 0 \) and \( u^* = 0 \) if \( t = 0 \). Instead of solving \( F(u) = 0 \) where we need to find an initial guess sufficiently close to an unknown solution which is too hard to do so, we solve \( G(u, t, t_1, \ldots, t_k) = 0 \) where we only need to find an initial guess not in \( S \). It is easy to do so since \( S^\perp \) is known. Consequently a solution found is always NEW.

**Theorem 5.1.** [54] Let \( u^* = t^*(\bar{u} + \sum_{i=1}^k t_i^* u_i) \) be a solution to \( F(u) = 0 \) approximated by solving \( G(u, t, t_1, \ldots, t_k) = 0 \) where \( \bar{u} \in S^+_1 \). If \( \bar{u} + \sum_{i=1}^k t_i^* u_i \notin \ker(F'(0)) = K_0 \). Then \( t^* \neq 0 \) and \( u^* \notin S \).

**Remark 5.1.** (1) Since \( u \in S^+_1 \) and \( \sum_{i=1}^k t_i^* u_i \in S \), the condition \( \bar{u} + \sum_{i=1}^k t_i^* u_i \notin \ker(F'(0)) = K_0 \) is trivially satisfied if \( K_0 \subset S \) or simply \( K_0 = \{0\} \), i.e., \( u = 0 \) is nondegenerate; (2) AST formulation is independent of any numerical method applied. Also the condition in Theorem 5.2 is irrelevant to any solution to be found and can be simply checked.

**Theorem 5.2.** [54] Let \( F(u)(x) = L(u)(x) + N(x, u(x)) \) where \( L \) is linear and \( N(x, y) \) is nonlinear in \( y \) s.t. \( N(x, 0) = 0, N_{y}(x, 0) = 0 \). Thus \( F(0) = 0 \) and \( \ker(F'(0)) = \ker(L) \). If \( F(u) = 0 \) and \( N(x, u(x)) \neq 0 \), we have \( u \notin \ker(F'(0)) \) and Theorem 5.1 is applicable.

Denote the set-valued \( S \)-\( \perp \) mapping \( P : S^+_1 \to 2^U \) by

\[
P(u) = \{ w = t(u + t_1 u_1 + \cdots + t_k u_k) : F(w) \perp u, u_1, \ldots, u_k \}, \quad \forall u \in S^+_1,
\]

an \( S \)-\( \perp \) selection \( p : S^+_1 \to U \) by \( p(u) \in P(u) \) \( \forall u \in S^+_1 \). Let \( \mathcal{M}_G = \{ P(u) : u \in S^+_1 \} \) be the solution set. Then the problem is converted to find \( u \in S^+_1 \) s.t. \( F(p(u)) = 0 \). For this composite function, due to the constraint on \( u \), it is too difficult to apply a Newton method directly. Instead, we design a partial Newton-correction method (PNCM), which can be described in the following two steps. Choose an initial guess \( u^1 \in S^+_1 \) and set \( n = 1 \); Step 1: (Correction) find \( w^n = (u^n, t^n_1, \ldots, t^n_k) \) s.t. \( \bar{w}^n = t^n(u^n + \sum_{i=1}^k t^n_i u_i) \in \mathcal{M}_G \). Thus \( G(w^n) \perp [u^n, S] \); Step 2: (Partial Newton Iteration) update \( u^n \) by \( u^{n+1} = \frac{u^n + \nu^n}{\|u^n + \nu^n\|} \in S^+_1 \) where the Newton direction \( \nu^n \in S^\perp \) is solved from a linear system involving a partial derivative

\[
G'_{u}(w^n)\nu^n = -G(w^n). \tag{5.2}
\]
Since the new point $t^n(u^{n+1} + t^n_u u_1 + \cdots + t^n_k u_k)$ will be off the set $\mathcal{M}_G$, a correction is required. So we set $n = n + 1$ and return to Step 1. The algorithm iteration continues.

**Remark 5.2.** (1) For (5.2) to be solvable, we need the condition that $\bar{w}^n$ is near a nondegenerate solution. When the condition does not hold, the Newton direction $ν \in S^⊥$ can still be solved from (5.2) by the pseudoinverse [61]. Thus this part will not cause any difficulty. (2) Due to the affect of $S$, PNCM is no longer symmetry invariant.

**Example 5.1.** [54] Consider solving the following nonvariational equation:

$$F(u)(x) = -\Delta u(x) + x \cdot \nabla u(x) - |x|^r u(x)^3 = 0 \text{ in } Ω = (-1,1)^2,$$

(5.3)

where $U = H^1_o(Ω)$ and $V = W^{-1,2}(Ω)$. 6 solutions $u_1, \cdots, u_6$ are found by PNCM. Their profiles and the peak locations $(\bar{x}_1, \bar{x}_2)$ of an initial guess $u^0$ are shown in Figures 15-16 where $u_1, u_4$ and $u_5$ have the same symmetry as $u^0$ w.r.t. the line $x_1 = x_2$; $u_2$ in Figure 15 and its $u^0$ have different symmetry, but $u_3$ in Figure 15 has a very nice symmetry and its $u^0$ is asymmetric, $u_6$ is asymmetric.

![Figure 15: Example 5.1. Solutions $u_1$ and $u_2$ with $(\bar{x}_1, \bar{x}_2)$ at $(-0.5,-0.5)$, $u_3$ with $(\bar{x}_1, \bar{x}_2)$ at $(0,-0.5), (0,0.5), (-0.5,0)$ and $(0.5,0)$](image1.png)

![Figure 16: Example 5.1. Solutions $u_4$ with $(\bar{x}_1, \bar{x}_2)$ at $(-0.5,-0.5)$, $u_5$ and $u_6$ with $(\bar{x}_1, \bar{x}_2)$ at $(0,0.5), (0,-0.5)$ and $(-0.5,0)$](image2.png)
A Final Remark: It is understood that multiple solution problems have very rich varieties, thus there are still many types of problems, in particular, systems to be explored; new numerical methods and results appear constantly, some of them are well justified and matured, and others are not quite yet and still need more times. It is impossible to give a complete survey in this field. A long list of references is cited in this paper, we hope it will provide readers with more sufficient information and better systematic understanding on application background, mathematical results on the existence, computational methods and theory on multiple solution problems. Despite of this, it is still by no means exclusive. Due to the limit on the author’s knowledge and others, we sincerely apology to those whose results are not included herewith.

References


