Notes on abstract bifurcation theory

Junping Shi

The partial differential equation

$$\begin{cases} \Delta u + \lambda f(u) = 0, & \text{in} \quad \Omega, \\ u = 0, & \text{on} \quad \partial\Omega, \end{cases} \tag{1}$$

can be formulated to a functional equation:

$$F(\lambda, u) = 0, \tag{2}$$

where $F : \mathbf{R} \times X \to Y$, X and Y are Banach spaces, and λ is a real value parameter. When F is sufficiently smooth, we can often use the bifurcation theory based on differential calculus in Banach spaces. In this notes, we prove several local bifurcation theorems based on implicit function theorem in Banach spaces. In particular we obtain infinite dimension version of basic bifurcation phenomena like saddle-node, transcritical and pitchfork described before . In this notes, we always assume X and Y are Banach spaces.

1 Banach spaces and implicit function theorem

A *metric space* is a pair (M, d) where M is a set and $d : M \times M \to \mathbf{R}$ is a metric which satisfies for any $x, y, z \in M$: (i) $d(x, y) \geq 0$; (ii) $d(x, y) = 0$ if and only if $x = y$; (iii) $d(x, y) = d(y, x)$; and (iv) (triangle inequality) $d(x, z) \leq d(x, y) + d(y, z)$. A metric space (M, d) is said to be *complete* if any Cauchy sequence $\{x_n\} \subset M$ has a limit in M.

A *normed vector space* (over real numbers) is a pair $(V, ||\cdot||)$ where V is a linear vector space over real numbers and the norm $||\cdot|| : V \to \mathbf{R}$ is a function which satisfies for any $a \in \mathbf{R}$ and $x, y \in V$: (i) $||ax|| = |a| \cdot ||x||$; (ii) $||x|| \geq 0$, and $||x|| = 0$ if and only if $x = 0$ (the zero vector); and (iii) (triangle inequality) $||x + y|| \leq ||x|| + ||y||$. A normed vector space is a metric space with the metric $d(x, y) = ||x - y||$. A complete normed vector space is called a *Banach space* named after Stefan Banach (1892–1945).

An important tool of nonlinear analysis is the contraction mapping principle (or Banach fixed point theorem):

Theorem 1.1. (Contraction mapping principle) *Let (M, d) be a non-empty completed metric space. Assume that $T : M \to M$ is a contraction mapping, that is, there exists $k \in (0, 1)$ such that for any $x, y \in M$,*

$$d(Tx, Ty) \leq k \cdot d(x, y). \tag{3}$$

Then the mapping T has a unique fixed point x_* in M such that $Tx_* = x_*$.

Proof. Choose any $x_0 \in M$, and define $x_n = Tx_{n-1}$ for $n \geq 1$. From (3) and the principle of mathematical induction, one obtain that $d(x_{n+1}, x_n) \leq k^n d(x_1, x_0)$. This in turn implies that $\{x_n\}$ is a Cauchy sequence in M since $k < 1$, and thus $\{x_n\}$ has a limit $x_* \in M$ from the completeness of (M, d). Note that $0 \le d(x_{n+1}, Tx_*) = d(Tx_n, Tx_*) \le kd(x_n, x_*)$. Then $d(x_{n+1}, Tx_*) \to 0$ as $n \to \infty$ since $d(x_n, x_*) \to 0$ as $n \to \infty$. Since the limit of $\{x_n\}$ is unique, hence $Tx_* = x_*$ and x_* is a fixed point of T in M. If there is another fixed point $y_* \in M$ of T, then $0 \le d(x_*, y_*) = d(Tx_*, Ty_*) \le k \cdot d(x_*, y_*)$. But $k < 1$ so $d(x_*, y_*) = 0$, which implies $x_* = y_*$ from the definition of metric. Therefore T has a unique fixed point x_*, which is the limit of any iterated sequence $\{x_n\}$ defined as $x_n = Tx_{n-1}$ and any initial point $x_0 \in M$. □

The foundation of analytical bifurcation theory is the following implicit function theorem, which is a consequence of contraction mapping principle.

Theorem 1.2. (Implicit function theorem) *Let X, Y and Z be Banach spaces, and let $U \subset X \times Y$ be a neighborhood of (λ_0, u_0). Let $F : U \to Z$ be a continuously differentiable mapping. Suppose that $F(\lambda_0, u_0) = 0$ and $F_u(\lambda_0, u_0)$ is an isomorphism, i.e. $F_u(\lambda_0, u_0)$ is one-to-one and onto, and $F_u^{-1}(\lambda_0, u_0) : Z \to Y$ is a linear bounded operator. Then there exists a neighborhood A of λ_0 in X, and a neighborhood B of u_0 in Y, such that for any $\lambda \in A$, there exists a unique $u(\lambda) \in B$ satisfying $F(\lambda, u(\lambda)) = 0$. Moreover $u(\cdot) : A \to B$ is continuously differentiable, and $u'(\lambda_0) : X \to Y$ is defined as $u'(\lambda_0)[\psi] = -[F_u(\lambda_0, u_0)]^{-1} \circ F_\lambda(\lambda_0, u_0)[\psi]$.*

Proof. To solve u in the equation $F(\lambda, u) = 0$, we look for the solution (μ, v) of $F(\lambda_0 + \mu, u_0 + v) = 0$. We notice that

$$F(\lambda_0 + \mu, u_0 + v) = F(\lambda_0, u_0) + F_u(\lambda_0, u_0)v + R(\mu, v),$$

where $R(\mu, v) = F(\lambda_0 + \mu, u_0 + v) - F_u(\lambda_0, u_0)v$ is the remainder term. Since $F_u(\lambda_0, u_0)$ is invertible, then $F(\lambda, u) = 0$ is equivalent to

$$v + [F_u(\lambda_0, u_0)]^{-1} R(\mu, v) = 0.$$

Define $G(\mu, v) = -[F_u(\lambda_0, u_0)]^{-1} R(\mu, v) = [F_u(\lambda_0, u_0)]^{-1}(F_u(\lambda_0, u_0)v - F(\lambda_0 + \mu, u_0 + v))$. Then solving $F(\lambda, u) = 0$ is equivalent to finding the fixed points of $G(\mu, v)$. We show that for μ close to 0, $G(\mu, \cdot)$ is a contraction mapping in a neighborhood of $v = 0$. In the following we denote $[F_u(\lambda_0, u_0)]^{-1}$ by H. Notice that $G(\mu, v) = v - HF(\lambda_0 + \mu, u_0 + v)$, then

$$\begin{aligned} & ||G(\mu, v_1) - G(\mu, v_2)|| \\ = \ & ||v_1 - v_2 - H(F(\lambda_0 + \mu, u_0 + v_1) - F(\lambda_0 + \mu, u_0 + v_2))|| \\ = \ & \left|\left|(v_1 - v_2) - H \int_0^1 F_u(\lambda_0 + \mu, u_0 + tv_1 + (1-t)v_2) dt (v_1 - v_2)\right|\right|. \end{aligned} \tag{4}$$

Since $F_u(\lambda, u)$ is continuous near (λ_0, u_0), then there exists a ball $B_X \subset X$ centered at $\mu = 0$ and a ball $B_Y \subset Y$ centered at $v = 0$ such that when $(\mu, v) \in B_X \times B_Y$,

$$||I - HF_u(\lambda_0 + \mu, u_0 + v)|| \le \frac{1}{2}. \tag{5}$$

In (5), the norm $||\cdot||$ is the operator norm in the Banach space $L(Y,Y)$, which consists all the linear bounded operators from Y to Y. From (4) and (5), we find that $||G(\mu,v_1)-G(\mu,v_2)||\le ||v_1-v_2||/2$ for $\mu\in B_X$ and $v_1,v_2\in B_Y$.

Next we show that for μ in a neighborhood of 0 and $v\in B_Y$, then $G(\mu,v)\in B_Y$. In fact, for $\mu\in B_X$, and $v\in B_Y$ with B_X,B_Y defined above, we assume that $B_Y=\{y\in Y: ||y||\le\delta\}$, then

$$\begin{aligned} &||G(\mu,v)||\le ||G(\mu,0)||+||G(\mu,0)-G(\mu,v)|| \\ \le\ & ||HF(\lambda_0+\mu,u_0)||+\frac{1}{2}||v||. \end{aligned} \tag{6}$$

From the continuity of HF and that $HF(\lambda_0,u_0)=0$, we can find a ball $B'_X\subset B_X$ with $0\in B'_X$ so that when $\mu\in B'_X$, $||HF(\lambda_0+\mu,u_0)-HF(\lambda_0,u_0)||\le(1/4)\delta$. Hence $||G(\mu,v)||\le(3/4)\delta$ and $G(\mu,v)\in B_Y$. Therefore when $\mu\in B'_X$, $G(\mu,\cdot):B_Y\to B_Y$ is a contraction mapping, then from the contraction mapping principle (Theorem 1.1), there exists a unique $v(\mu)\in B_Y$ such that $G(\mu,v(\mu))=v(\mu)$. Hence the existence and uniqueness of the solution to $F(\lambda,u)=0$ for $\lambda\in A$ and $u\in B$ in the theorem follows by letting $A=\{\lambda_0+\mu:\mu\in B'_X\}$ and $B=\{u_0+v:v\in B_Y\}$.

To show that $u(\lambda)=u_0+v(\mu)$ is continuous, we see that for $\mu_1,\mu_2\in B'_X$,

$$\begin{aligned} &||v(\mu_1)-v(\mu_2)||=||G(\mu_1,v(\mu_1))-G(\mu_2,v(\mu_2))|| \\ \le\ & ||G(\mu_1,v(\mu_1))-G(\mu_1,v(\mu_2))||+||G(\mu_1,v(\mu_2))-G(\mu_2,v(\mu_2))|| \\ \le\ & \frac{1}{2}||v(\mu_1)-v(\mu_2)|| \\ &+||HF(\lambda_0+\mu_1,u_0+v(\mu_2))-HF(\lambda_0+\mu_2,u_0+v(\mu_2))||, \end{aligned}$$

hence

$$\begin{aligned} &||v(\mu_1)-v(\mu_2)|| \\ \le\ & 2\left|\left|H\int_0^1 F_\mu(\lambda_0+t\mu_1+(1-t)\mu_2,u_0+v(\mu_2))dt(\mu_1-\mu_2)\right|\right|, \end{aligned} \tag{7}$$

and the continuity of $u(\lambda)$ follows from the continuity of F_μ. To show the differentiability of $v(\mu)$, we notice that G is continuously differentiable near $(0,0)$ since F is assumed to be C^1, $G_\mu(\mu,v)=-H\circ F_\lambda(\lambda_0+\mu,u_0+v)$ and $G_v(\mu,v)=H\circ(F_u(\lambda_0,u_0)-F_u(\lambda_0+\mu,u_0+v))$. Hence for $\mu\in B'_X$ and $\psi\in X$ small, from the differentiability of G and continuity of F_u, we have

$$\begin{aligned} &\big|\big|v(\mu+\psi)-v(\mu)+[F_u(\lambda_0,u_0)]^{-1}\circ F_\lambda(\lambda_0,u_0)[\psi]\big|\big| \\ =\ & \big|\big|G(\mu+\psi,v(\mu+\psi))-G(\mu,v(\mu))+H\circ F_\lambda(\lambda_0,u_0)[\psi]\big|\big| \\ =\ & \Big|\Big|G_\mu(\mu,v(\mu))[\psi]+G_v(\mu,v(\mu))[v(\mu+\psi)-v(\mu)] \\ &+o(||\psi||)+o(||v(\mu+\psi)-v(\mu)||)+H\circ F_\lambda(\lambda_0+\mu,u_0+v(\mu))[\psi]\Big|\Big| \\ =\ & o(||\psi||)+o(||v(\mu+\psi)-v(\mu)||)=o(||\psi||). \end{aligned}$$

Therefore $v : B'_X \to B_Y$ is differentiable with the Fréchet derivative $v'(\mu) = -[F_u(\lambda_0, u_0)]^{-1} \circ F_\lambda(\lambda_0 + \mu, u_0 + v(\mu))$, which is continuous in μ from the continuity of F_λ and v. This proves that $u(\lambda)$ is C^1. □

Note that if we assume that F is continuously differentiable in u, and only continuous in λ, then the result still holds, but $u(\lambda)$ is only continuous. Similarly, if F is of class C^k, so is $u(\lambda)$; if F is analytic, so is $u(\lambda)$.

An important special case is when $X = \mathbf{R}$, then the implicit function theorem Theorem 1.2 implies that when the linearized operator $F_u(\lambda_0, u_0)$ is non-degenerate, then the set of solutions to $F(\lambda, u) = 0$ near (λ_0, u_0) is a C^1 curve $\{(\lambda, u(\lambda)) : \lambda \in (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)\}$. We apply it to (1):

Theorem 1.3. (Implicit function theorem) *Suppose that $f \in C^1(\mathbf{R})$, and $(\lambda_0, u_0) \in \mathbf{R} \times C_0^{2,\alpha}(\overline{\Omega})$ is a solution of (1), such that the equation*

$$\begin{cases} \Delta w + \lambda f'(u) w = 0, & \text{in} \quad \Omega, \\ w = 0, & \text{on} \quad \partial\Omega, \end{cases} \tag{8}$$

has only the trivial solution $w = 0$, then there exists $\varepsilon > 0$ such that for $\lambda \in (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)$, (1) has a unique solution $(\lambda, u(\lambda))$ near (λ_0, u_0), and $\{(\lambda, u(\lambda)) : \lambda \in (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)\}$ is a smooth curve.

Proof. Define $F(\lambda, u) = \Delta u + \lambda f(u)$ for $\lambda \in \mathbf{R}$ and $u \in X$. $F_u(\lambda_0, u_0)$ is a Fredholm operator of index zero. Since (8) has only the trivial solution, then $N(F_u(\lambda_0, u_0)) = \emptyset$ and $R(F_u(\lambda_0, u_0)) = Y$. From the open mapping theorem, $F_u(\lambda_0, u_0)$ is an isomorphism. Then the result follows from Theorem 1.2. □

2 Bifurcations on $\mathbf{R}^1$

The implicit function theorem provides a tool to describe the solution set of a nonlinear problem in an infinite dimensional space when the linearized operator is invertible. When the linearized operator is not invertible, but with only a kernel of finite dimension and a range space of finite codimension, the analytic bifurcation picture still resembles the ones in finite dimensional case described in Section 1.2. For that purpose, we first establish a result for finite dimensional bifurcation problem, and this result also gives a unified approach to the usual bifurcations of type saddle-node, transcritical and pitchfork, thus it is also of independent interest.

Theorem 2.1. *Suppose that $(\lambda_0, y_0) \in \mathbf{R}^2$ and U is a neighborhood of (λ_0, y_0). Assume that $f : U \to \mathbf{R}$ is a C^p function for $p \geq 1$, $f(\lambda_0, y_0) = 0$, and there is at most one critical point (λ_0, y_0) of f in U. Define S to be the connected component of $\{(\lambda, y) \in U : f(\lambda, y) = 0\}$ which contains (λ_0, y_0).*

1. *If $\nabla f(\lambda_0, y_0) \neq 0$, then S is a C^p curve passing through (λ_0, y_0).*

2. *If $\nabla f(\lambda_0, y_0) = 0$, we assume in addition that $p \geq 2$, and the Hessian $H = \nabla^2 f(\lambda_0, y_0)$ is non-degenerate with eigenvalues $\lambda_1, \lambda_2 \neq 0$, then*

   (a) *when $\lambda_1\lambda_2 > 0$, (λ_0, y_0) is the unique zero point of $f(x, y) = 0$ near (λ_0, y_0);*

   (b) *when $\lambda_1\lambda_2 < 0$, there exist two C^{p-1} curves $\{(\lambda_i(t), y_i(t)) : |t| \leq \delta\}$, $i = 1, 2$, such that S consists of exactly the two curves near (λ_0, y_0), $(\lambda_i(0), y_i(0)) = (\lambda_0, y_0)$. Moreover t can be rescaled so that $(\eta, \tau) = (\lambda_i'(0), y_i'(0))$, $i = 1, 2$, are the two linear independent solutions of*

$$f_{\lambda\lambda}(\lambda_0, y_0)\eta^2 + 2f_{\lambda y}(\lambda_0, y_0)\eta\tau + f_{yy}(\lambda_0, y_0)\tau^2 = 0. \tag{9}$$

Proof. Part (1) follows from the implicit function theorem (Theorem 1.2). Indeed, if $f_y(\lambda_0, y_0) \neq 0$, then S is in form $\{(\lambda, y(\lambda)) : |\lambda - \lambda_0| < \varepsilon\}$, and if $f_\lambda(\lambda_0, y_0) \neq 0$, then S is in form $\{(\lambda(y), y) : |y - y_0| < \varepsilon\}$. Part (2a) follows from standard multi-variable calculus since in this case, (λ_0, y_0) is a strict local maximum or minimum point of $f(x, y)$. So we only need to prove (2b).

Consider the system of differential equations:

$$\lambda' = \frac{\partial f(\lambda, y)}{\partial y}, \quad y' = -\frac{\partial f(\lambda, y)}{\partial \lambda}, \quad (\lambda(0), y(0)) \in U. \tag{10}$$

Then (10) is a Hamiltonian system with potential function $f(\lambda, y)$, and (λ_0, y_0) is the only equilibrium point of (10) in U. The Jacobian of (10) at (λ_0, y_0) is

$$J = \begin{pmatrix} f_{\lambda y}(\lambda_0, y_0) & f_{yy}(\lambda_0, y_0) \\ -f_{\lambda\lambda}(\lambda_0, y_0) & -f_{\lambda y}(\lambda_0, y_0) \end{pmatrix}. \tag{11}$$

Since $Trace(J) = 0$ and $Det(J) = Det(H) < 0$, then (λ_0, y_0) is a saddle type equilibrium of (10) and J has eigenvalues $\pm k$ for some $k > 0$. From the invariant manifold theory of differential equations, there exists a unique curve $\Gamma_s \subset U$ (the stable manifold) such that Γ_s is invariant for (10) and for $(\lambda(0), y(0)) \in \Gamma_s$, $(\lambda(t), y(t)) \to (\lambda_0, y_0)$ as $t \to \infty$; and similarly the unstable manifold is another invariant curve Γ_u for (10) and for $(\lambda(0), y(0)) \in \Gamma_u$, $(\lambda(t), y(t)) \to (\lambda_0, y_0)$ as $t \to -\infty$. Both Γ_s and Γ_u are C^{p-1} one-dimensional manifold by the stable and unstable manifold theorem ([Pe] page 107). $f(\lambda, y) = 0$ for $(\lambda, y) \in \Gamma_s \cup \Gamma_u$ since $f(\lambda, y)$ is the Hamiltonian function of the system and $\Gamma_s \cup \Gamma_u \cup \{(\lambda_0, y_0)\}$. On the other hand, for any $(\lambda, y) \notin \Gamma_s \cup \Gamma_u \cup \{(\lambda_0, y_0)\}$, $f(\lambda, y) \neq 0$ from the Morse lemma.

Finally we consider the tangential direction of Γ_s and Γ_u. We denote the two curves by $(\lambda_i(t), y_i(t))$, with $i = 1, 2$. Then

$$f(\lambda_i(t), y_i(t)) = 0. \tag{12}$$

Differentiating (12) in t twice, we obtain (we omit the subscript i for $\lambda_i(t)$ and $y_i(t)$ in the equation)

$$\begin{aligned} &f_{\lambda\lambda}(\lambda(t), y(t))(\lambda'(t))^2 + 2f_{\lambda y}(\lambda(t), y(t))\lambda'(t)y'(t) + f_{yy}(\lambda(t), y(t))(y'(t))^2 \\ &+f_\lambda(\lambda(t), y(t))\lambda''(t) + f_y(\lambda(t), y(t))y''(t) = 0. \end{aligned}$$

evaluating at $t = 0$ and $\nabla f(\lambda_0, y_0) = 0$, we obtain (9). □

Theorem 2.1 is proved in Liu, Wang and Shi [LSW], and a weaker result is proved in Nirenberg [Nir] Theorem 3.2.1, in which the crossing curves are shown to be C^{p-2}. Here we give an alternate proof using the invariant manifold theory.

In Theorem 2.1, if $f_y(\lambda_0, y_0) \neq 0$, then the C^p curve can be parameterized by λ; if $f_\lambda(\lambda_0, y_0) \neq 0$, then the C^p curve can be parameterized by y and indeed we have the saddle-node bifurcation; and if we assume that $f(\lambda, y_0) \equiv 0$, $f_y(\lambda_0, y_0) = 0$, and $f_{\lambda y}(\lambda_0, y_0) \neq 0$, then we obtain transcritical or pitchfork bifurcations. In this sense, Theorem 2.1 gives a unified unfolding of singularity in $\mathbf{R}^2$ with codimension 2. Using the implicit function theorem in Banach spaces, we will establish similar bifurcation results in Banach spaces in the following sections.

3 Saddle-node bifurcation

From the implicit function theorem (Theorem 1.2), a necessary condition for bifurcation is that

$$F_u(\lambda_0, u_0) \text{ is not invertible.} \tag{13}$$

When (13) is satisfied, we call (λ_0, u_0) a *degenerate solution* of $F(\lambda, u) = 0$. Here we discuss the case when the kernel of $F_u(\lambda_0, u_0)$ is nonempty, and in particular, we discuss the case that $\mu = 0$ is a simple eigenvalue of $F_u(\lambda_0, u_0)$, *i.e.*

(F1) $dim N(F_u(\lambda_0, u_0)) = codim R(F_u(\lambda_0, u_0)) = 1$, and $N(F_u(\lambda_0, u_0)) = span\{w_0\}$.

(F1) is equivalent to: $F_u(\lambda_0, u_0)$ has a one-dimension kernel, and it is a Fredholm operator with index zero. The range space $R(F_u(\lambda_0, u_0))$ is a subspace of Y of co-dimension one, then there exists $l \in Y^*$ (the space of linear functionals on Y), such that

$$u \in R(F_u(\lambda_0, u_0)) \Leftrightarrow \langle l, u \rangle = 0, \tag{14}$$

where $\langle l, u \rangle$ is the duality relation between Y^* and Y. In the following whenever (F1) is assumed, l is the associated linear functional in Y^*.

Theorem 3.1. (Saddle-node bifurcation theorem) *Let U be a neighborhood of (λ_0, u_0) in $\mathbf{R} \times X$, and let $F : U \to Y$ be a continuously differentiable mapping. Assume that $F(\lambda_0, u_0) = 0$, F satisfies (F1) at (λ_0, u_0) and*

(F2) $F_\lambda(\lambda_0, u_0) \notin R(F_u(\lambda_0, u_0))$.

1. *If Z is a complement of $span\{w_0\}$ in X, then the solutions of $F(\lambda, u) = 0$ near (λ_0, u_0) form a curve $\{(\lambda(s), u(s)) = (\lambda(s), u_0 + sw_0 + z(s)) : |s| < \delta\}$, where $s \mapsto (\lambda(s), z(s)) \in \mathbf{R} \times Z$ is a continuously differentiable function, $\lambda(0) = \lambda'(0) = 0$, and $z(0) = z'(0) = 0$.*
2. *If F is $k-$times continuously differentiable, so are $\lambda(s)$ and $z(s)$.*

3. If F is C^2 in u, then

$$\lambda''(0) = -\frac{\langle l, F_{uu}(\lambda_0, u_0)[w_0, w_0]\rangle}{\langle l, F_\lambda(\lambda_0, u_0)\rangle}, \tag{15}$$

where $l \in Y^*$ satisfying $N(l) = R(F_u(\lambda_0, u_0))$.

Proof. Define $G : U_1 \times (U_2 \times Z_1) \to Y$ by

$$G(s, \lambda, z) = F(\lambda, u_0 + sw_0 + z), \tag{16}$$

where U_1, U_2 are neighborhoods of 0 in $\mathbf{R}$ and Z_1 is neighborhood of 0 in Z so that the right hand side of (16) is well-defined ($(\lambda, u_0 + sw_0 + z) \in U$). Then G has the same smoothness as F and $G(0, \lambda_0, 0) = 0$. We claim that the partial derivative $G_{(\lambda,z)}(0, \lambda_0, 0) : \mathbf{R} \times Z \to Y$ is an isomorphism. We first show that $G_{(\lambda,z)}(0, \lambda_0, 0)$ is injective. Suppose that there exists $(\tau, \psi) \in \mathbf{R} \times Z$ such that $G_{(\lambda,z)}(0, \lambda_0, 0)[(\tau, \psi)] = 0$, then

$$\tau F_\lambda(\lambda_0, u_0) + F_u(\lambda_0, u_0)[\psi] = 0. \tag{17}$$

Applying l to (17), we obtain

$$\tau \langle l, F_\lambda(\lambda_0, u_0)\rangle = 0. \tag{18}$$

Since $F_\lambda(\lambda_0, u_0) \notin R(F_u(\lambda_0, u_0))$, then $\tau = 0$, and $\psi = 0$ since $\psi \in Z$ and $N(F_u(\lambda_0, u_0)) = span\{w_0\}$. Next we show that $G_{(\lambda,z)}(0, \lambda_0, 0)$ is surjective. Let $\theta \in Y$. Applying l to

$$\tau F_\lambda(\lambda_0, u_0) + F_u(\lambda_0, u_0)[\psi] = \theta, \tag{19}$$

we obtain

$$\tau = \frac{\langle l, \theta\rangle}{\langle l, F_\lambda(\lambda_0, u_0)\rangle}, \tag{20}$$

and $\psi = K[\theta - \tau F_\lambda(\lambda_0, u_0)]$, where K is the inverse of $F_u(\lambda_0, u_0)|_Z$. Thus (τ, ψ) is uniquely determined by θ. Since G is continuous, $G_{(\lambda,z)}(0, \lambda_0, 0)$ is a bijection, then $[G_{(\lambda,z)}(0, \lambda_0, 0)]^{-1}$ is also continuous by the open mapping theorem of Banach. Hence $G_{(\lambda,z)}(0, \lambda_0, 0)$ is an isomorphism. By Theorem 1.2, the first two statements of theorem are true. And $(\lambda'(0), z'(0))$ is determined by

$$G_{(\lambda,z)}(0, \lambda_0, 0)[(\lambda'(0), z'(0))] = -G_s(0, \lambda_0, 0) = -F_u(\lambda_0, u_0)[w_0] = 0, \tag{21}$$

so $\lambda'(0) = 0$ and $z'(0) = 0$ from the injectivity of $G_{(\lambda,z)}(0, \lambda_0, 0)$.

Differentiating $F(\lambda(s), u(s)) = 0$ with respect to s twice, we obtain

$$\begin{aligned} &\lambda''(s)F_\lambda + [\lambda'(s)]^2 F_{\lambda\lambda} + 2\lambda'(s)F_{\lambda u}[u'(s)] \\ &+F_{uu}[u'(s), u'(s)] + F_u[u''(s)] = 0. \end{aligned} \tag{22}$$

Let $s = 0$ in (22) and apply l to (22), then we obtain (15). □

Theorem 3.1 first appeared in Crandall and Rabinowitz [CR2]. When

(F4) $F_{uu}(\lambda_0, u_0)[w_0, w_0] \notin R(F_u(\lambda_0, u_0))$,

is satisfied, $\lambda''(0) \neq 0$, and the solution set $\{(\lambda(s), u(s)) : |s| < \delta\}$ is a parabola-like curve which reaches an extreme point at (λ_0, u_0). The degenerate solution (λ_0, u_0) in this case is called a *turning point* on the solution curve. When $\lambda''(0) > 0$, the bifurcation is *supercritical*; and when $\lambda''(0) < 0$, the bifurcation is *subcritical*. Similar to previous applications to (1), we have

Theorem 3.2. *Suppose that $f \in C^1(\mathbf{R}^+)$ and (λ_*, u_*) is a positive solution of* (1) *which satisfies*

$$\frac{\partial u_*}{\partial \nu}(x) < 0 \quad \text{for all } x \in \partial\Omega. \tag{23}$$

and suppose that the linearized equation (8) *has a unique (up to scale) nontrivial solution w, which satisfies*

$$\int_\Omega f(u_*) w dx \neq 0. \tag{24}$$

Then all the positive solutions of (1) *near (λ_*, u_*) have the form $(\lambda(s), u_* + sw + z(s))$ for $s \in (-\delta, \delta)$ and some $\delta > 0$, where $\lambda(0) = \lambda_*$, $\lambda'(0) = 0$, $z(0) = z'(0) = 0$. Moreover, if $f \in C^2(\mathbf{R}^+)$, then*

$$\lambda''(0) = -\frac{\lambda_* \int_\Omega f''(u_*) w^3(x) dx}{\int_\Omega f(u_*) w(x) dx}. \tag{25}$$

Proof. The setting is similar to that of Theorem 1.3. Recall from the proof of Theorem 1.3, $F_u(\lambda_*, u_*) = \Delta + \lambda_* f'(u_*)$, which is a Fredholm operator with index 0. (F1) is satisfied from the assumption that the solution space of (1) is one-dimensional. $R(F_u(\lambda_*, u_*)) = \{\phi \in C^\alpha(\overline{\Omega}) : \int_\Omega \phi(x) w(x) dx = 0\}$. Thus (24) is equivalent to (F2), and stated results follow from Theorem 3.1 except the positivity of $u(s) = u_* + sw + z(s)$, which follows from (23). $\square$

We also comment that at a bifurcation point described in Theorem 3.1, if $F \in C^3$ and $\lambda''(0) = 0$, then

$$\lambda'''(0) = -\frac{\langle l, F_{uuu}(\lambda_0, u_0)[w_0, w_0, w_0]\rangle + 3\langle l, F_{uu}(\lambda_0, u_0)[w_0, \theta]\rangle}{\langle l, F_\lambda(\lambda_0, u_0)\rangle}, \tag{26}$$

where $\theta \in Z$ is the solution of

$$F_{uu}(\lambda_0, u_0)[w_0, w_0] + F_u(\lambda_0, u_0)[\theta] = 0. \tag{27}$$

The solvability of (27) is equivalent to

(F4′) $F_{uu}(\lambda_0, u_0)[w_0, w_0] \in R(F_u(\lambda_0, u_0))$,

or $\lambda''(0) = 0$. In the case $\lambda'''(0) \neq 0$, a cusp type bifurcation occurs near the degenerate solution (λ_0, u_0). More discussion can be found in Shi [S].

4 Transcritical and pitchfork bifurcations

If there is a branch of trivial solutions $u = u_0$ for all λ, then nontrivial solutions can bifurcate from the trivial branch at a degenerate solution. Here is the theorem of *Bifurcation from a simple eigenvalue* by Crandall and Rabinowitz [CR1]:

Theorem 4.1. (Transcritical and pitchfork bifurcation theorem) *Let U be a neighborhood of (λ_0, u_0) in $\mathbf{R} \times X$, and let $F : U \to Y$ be a twice continuously differentiable mapping. Assume that $F(\lambda, u_0) = 0$ for $(\lambda, u_0) \in U$. At (λ_0, u_0), F satisfies (F1) and*

(F3) $F_{\lambda u}(\lambda_0, u_0)[w_0] \notin R(F_u(\lambda_0, u_0))$.

Let Z be any complement of span$\{w_0\}$ in X. Then the solution set of (2) near (λ_0, u_0) consists precisely of the curves $u = u_0$ and $\{(\lambda(s), u(s)) : s \in I = (-\epsilon, \epsilon)\}$, where $\lambda : I \to \mathbf{R}$, $z : I \to Z$ are C^1 functions such that $u(s) = u_0 + sw_0 + sz(s)$, $\lambda(0) = \lambda_0$, $z(0) = 0$, and

$$\lambda'(0) = -\frac{\langle l, F_{uu}(\lambda_0, u_0)[w_0, w_0]\rangle}{2\langle l, F_{\lambda u}(\lambda_0, u_0)[w_0]\rangle}, \tag{28}$$

where $l \in Y^*$ satisfying $N(l) = R(F_u(\lambda_0, u_0))$. If $\lambda'(0) = 0$, and in addition $F \in C^3$ near (λ_0, u_0), then

$$\lambda''(0) = -\frac{\langle l, F_{uuu}(\lambda_0, u_0)[w_0, w_0, w_0]\rangle + 3\langle l, F_{uu}(\lambda_0, u_0)[w_0, \theta]\rangle}{3\langle l, F_{\lambda u}(\lambda_0, u_0)[w_0]\rangle}, \tag{29}$$

where θ is the solution of (27).

When $\lambda'(0) \neq 0$ (thus (F4) is satisfied), then a *transcritical bifurcation* occurs; if instead (F4$'$) is satisfied, then $\lambda'(0) = 0$, and a *pitchfork bifurcation* occurs at (λ_0, u_0) if $\lambda'(0) = 0$ and $\lambda''(0) \neq 0$. We remark that in the original theorem of [CR1], under the weaker assumption that $F_{\lambda u}$ exists and continuous near (λ_0, u_0) instead of F being C^2, it was shown that same result holds but the curve of nontrivial solutions is only continuous not C^1.

Here we prove this important theorem as a consequence of a more general result based on Theorem 2.1. We assume F satisfies (F1) at (λ_0, u_0), then we have decompositions of X and Y: $X = N(F_u(\lambda_0, u_0)) \oplus Z$ and $Y = R(F_u(\lambda_0, u_0)) \oplus Y_1$, where Z is a complement of $N(F_u(\lambda_0, u_0))$ in X, and Y_1 is a complement of $R(F_u(\lambda_0, u_0))$. In particular, $F_u(\lambda_0, u_0)|_Z : Z \to R(F_u(\lambda_0, u_0))$ is an isomorphism. Since $R(F_u(\lambda_0, u_0))$ is codimension one, then there exists $l \in Y^*$ such that $R(F_u(\lambda_0, u_0)) = \{v \in Y : \langle l, v\rangle = 0\}$. Recall the condition (F2) in saddle-node bifurcation theorem, here we assume the opposite:

(F2$'$) $F_\lambda(\lambda_0, u_0) \in R(F_u(\lambda_0, u_0))$.

Then the equation

$$F_\lambda(\lambda_0, u_0) + F_u(\lambda_0, u_0)[v] = 0 \tag{30}$$

has a unique solution $v_1 \in Z$. The following "crossing curve bifurcation theorem" is proved in Liu, Shi and Wang [LSW]:

Theorem 4.2. *Let U be a neighborhood of (λ_0, u_0) in $\mathbf{R}\times X$, and let $F : U \to Y$ be a twice continuously differentiable mapping. Assume that $F(\lambda_0, u_0) = 0$, F satisfies (F1) and (F2′) at (λ_0, u_0). Let $X = N(F_u(\lambda_0, u_0)) \oplus Z$ be a fixed splitting of X, let $v_1 \in Z$ be the unique solution of* (30)*, and let $l \in Y^*$ such that $R(F_u(\lambda_0, u_0)) = \{v \in Y : \langle l, v\rangle = 0\}$. We assume that the matrix (all derivatives are evaluated at (λ_0, u_0))*

$$H_0 \equiv \begin{pmatrix} \langle l, F_{\lambda\lambda} + 2F_{\lambda u}[v_1] + F_{uu}[v_1, v_1]\rangle & \langle l, F_{\lambda u}[w_0] + F_{uu}[w_0, v_1]\rangle \\ \langle l, F_{\lambda u}[w_0] + F_{uu}[w_0, v_1]\rangle & \langle l, F_{uu}[w_0, w_0]\rangle \end{pmatrix} \tag{31}$$

is non-degenerate, i.e. $Det(H_0) \neq 0$.

1. *If H_0 is definite, i.e. $Det(H_0) > 0$, then the solution set of $F(\lambda, u) = 0$ near $(\lambda, u) = (\lambda_0, u_0)$ is the single point set $\{(\lambda_0, u_0)\}$.*
2. *If H_0 is indefinite, i.e. $Det(H_0) < 0$, then the solution set of $F(\lambda, u) = 0$ near $(\lambda, u) = (\lambda_0, u_0)$ is the union of two intersecting C^1 curves, and the two curves are in form of $(\lambda_i(s), u_i(s)) = (\lambda_0 + \mu_i s + s\theta_i(s), u_0 + \eta_i s w_0 + s v_i(s))$, $i = 1, 2$, where $s \in (-\delta, \delta)$ for some $\delta > 0$, $\theta_i(0) = 0$, $v_i(s) \in Z$, $v_i(0) = 0$ $(i = 1, 2)$, and (μ_i, η_i) $(i = 1, 2)$ are non-zero linear independent solutions of the equation*

$$\begin{aligned} &\langle l, F_{\lambda\lambda} + 2F_{\lambda u}[v_1] + F_{uu}[v_1, v_1]\rangle\mu^2 \\ &+ 2\langle l, F_{\lambda u}[w_0] + F_{uu}[w_0, v_1]\rangle\eta\mu + \langle l, F_{uu}[w_0, w_0]\rangle\eta^2 = 0. \end{aligned} \tag{32}$$

Proof. We start by reducing the equation in infinite-dimensional space to a finite dimensional one by a Lyapunov-Schmidt process. We denote the projection from Y into $R(F_u(\lambda_0, u_0))$ by Q. Then $F(\lambda, u) = 0$ is equivalent to

$$Q \circ F(\lambda, u) = 0, \quad \text{and } (I - Q) \circ F(\lambda, u) = 0. \tag{33}$$

We rewrite the first equation in form

$$G(\lambda, t, g) \equiv Q \circ F(\lambda, u_0 + tw_0 + g) = 0 \tag{34}$$

where $t \in \mathbf{R}$ and $g \in Z$. Calculation shows that $G_g(\lambda_0, 0, 0) = Q \circ F_u(\lambda_0, u_0)$ is an isomorphism from Z to $R(F_u(\lambda_0, u_0))$. Then $g = g(\lambda, t)$ in (34) is uniquely solvable from the implicit function theorem Theorem 1.2 for (λ, t) near $(\lambda_0, 0)$, and g is C^2. Hence $u = u_0 + tw_0 + g(\lambda, t)$ is a solution to $F(\lambda, u) = 0$ if and only if $(I - Q) \circ F(\lambda, u_0 + tw_0 + g(\lambda, t)) = 0$. Since $R(F_u(\lambda_0, u_0))$ is co-dimensional one, hence it becomes the scalar equation $\langle l, F(\lambda, u_0 + tw_0 + g(\lambda, t))\rangle = 0$.

From arguments above we have

$$f_1(\lambda, t) \equiv Q \circ F(\lambda, u_0 + tw_0 + g(\lambda, t)) = 0, \tag{35}$$

for (λ, t) near $(\lambda_0, 0)$. Differentiating f_1 and evaluating at $(\lambda, t) = (\lambda_0, 0)$, we obtain

$$0 = \nabla f_1 = (Q \circ (F_\lambda + F_u[g_\lambda]), Q \circ F_u[w_0 + g_t]). \tag{36}$$

Since $F_u[w_0] = 0$ and $g_t \in Z$, and $F_u(\lambda_0, u_0)|_Z$ is an isomorphism, then $g_t(\lambda_0, 0) = 0$. Similarly $g_\lambda \in Z$ and $F_\lambda \in R(F_u(\lambda_0, u_0))$ from (F2′), hence

$$F_\lambda(\lambda_0, u_0) + F_u(\lambda_0, u_0)[g_\lambda(\lambda_0, 0)] = 0. \tag{37}$$

Hence $g_\lambda(\lambda_0, 0) = v_1$ where v_1 is defined as in (30).

To prove the statement in Theorem 4.2, we apply Theorem 2.1 to

$$f(\lambda, t) = \langle l, F(\lambda, u_0 + tw_0 + g(\lambda, t))\rangle. \tag{38}$$

From the proofs above, $F(\lambda, u) = 0$ for (λ, u) near (λ_0, u_0) is equivalent to $f(\lambda, t) = 0$ for (λ, t) near $(\lambda_0, 0)$. To apply Theorem 2.1, we claim that

$$\nabla f(\lambda_0, 0) = (f_\lambda, f_t) = 0, \text{ and } Hess(f) \text{ is non-degenerate.} \tag{39}$$

It is easy to see that

$$\begin{aligned} &\nabla f(\lambda_0, 0) \\ =&(\langle l, F_\lambda(\lambda_0, u_0) + F_u(\lambda_0, u_0)[g_\lambda(\lambda_0, 0)]\rangle, \langle l, F_u(\lambda_0, u_0)[w_0 + g_t(\lambda_0, 0)]\rangle). \end{aligned} \tag{40}$$

Thus $\nabla f(\lambda_0, 0) = 0$ from (30) and $g_t(\lambda_0, 0) = 0$. For the Hessian matrix, we have

$$Hess(f) = \begin{pmatrix} f_{\lambda\lambda} & f_{\lambda t} \\ f_{t\lambda} & f_{tt} \end{pmatrix}. \tag{41}$$

Here

$$\begin{aligned} f_{\lambda t}(\lambda_0, 0) =&f_{t\lambda}(\lambda_0, 0) \\ =&\langle l, F_{\lambda u}[w_0 + g_t] + F_{uu}[w_0 + g_t, g_\lambda] + F_u[g_{\lambda t}]\rangle \\ =&\langle l, F_{\lambda u}[w_0] + F_{uu}[w_0, v_1]\rangle, \end{aligned} \tag{42}$$

since $g_t = 0$. Next we have

$$\begin{aligned} f_{\lambda\lambda}(\lambda_0, 0) =&\langle l, F_{\lambda\lambda} + 2F_{\lambda u}[g_\lambda] + F_{uu}[g_\lambda, g_\lambda] + F_u[g_{\lambda\lambda}]\rangle \\ =&\langle l, F_{\lambda\lambda} + 2F_{\lambda u}[v_1] + F_{uu}[v_1, v_1]\rangle. \end{aligned} \tag{43}$$

Finally,

$$f_{tt}(\lambda_0, 0) = \langle l, F_{uu}[w_0 + g_t, w_0 + g_t] + F_u[g_{tt}]\rangle = \langle l, F_{uu}[w_0, w_0]\rangle. \tag{44}$$

In summary, from our calculation,

$$Hess(f) = \begin{pmatrix} \langle l, F_{\lambda\lambda} + 2F_{\lambda u}[v_1] + F_{uu}[v_1, v_1]\rangle & \langle l, F_{\lambda u}[w_0] + F_{uu}[w_0, v_1]\rangle \\ \langle l, F_{\lambda u}[w_0] + F_{uu}[w_0, v_1]\rangle & \langle l, F_{uu}[w_0, w_0]\rangle \end{pmatrix}.$$

Therefore from Theorem 2.1, we conclude that the solution set of $F(\lambda, u) = 0$ near $(\lambda, u) = (\lambda_0, u_0)$ is a pair of intersecting curves if the matrix in (4) is indefinite, or is a single point if it is definite.

Now we consider only the former case of two curves. We denote the two curves by $(\lambda_i(s), u_i(s)) = (\lambda_i(s), u_0 + t_i(s)w_0 + g(\lambda_i(s), t_i(s)))$, with $i = 1, 2$. Then

$$F(\lambda_i(s), u_0 + t_i(s)w_0 + g(\lambda_i(s), t_i(s))) = 0. \tag{45}$$

From Theorem 2.1 the vectors $v_i = (\lambda_i'(0), t_i'(0))$ are the solutions of $v^T H_0 v = 0$, which are the solutions (μ, η) of (32). □

Now we show that Theorem 4.1 is a special case of Theorem 4.2. In fact, the assumption of $F(\lambda, u_0) \equiv 0$ implies that (F2′) is satisfied and $F_\lambda(\lambda_0, u_0) = F_{\lambda\lambda}(\lambda_0, u_0) = 0$, thus $v_1 = 0$ and $det(H_0) = -\langle l, F_{\lambda u}(\lambda_0, w_0)\rangle^2$. Hence the assumption (F3) implies that $det(H_0) \neq 0$ and H_0 is indefinite. The formula in (28) can be obtained from (32) as it becomes

$$2\langle l, F_{\lambda u}[w_0]\rangle \eta\mu + \langle l, F_{uu}[w_0, w_0]\rangle \eta^2 = 0. \tag{46}$$

We can choose one solution of (46) to be $(\mu, \eta) = (1, 0)$ which corresponds to the line of trivial solutions, and the other solution to be $(\mu, \eta) = (\lambda'(0), 1)$ with $\lambda'(0)$ being the expression in (28). Finally (29) can be obtained with further calculations. Indeed let $(\lambda(s), u(s))$ be the nontrivial solution curve. Then by differentiating $F(\lambda(s), u(s)) = 0$ three times, evaluating at $s = 0$ and applying l, one can obtain (29).

The implicit function theorem (transversal curve), saddle-node bifurcation (turning curve), and transcritical/pitchfork bifurcation (two crossing curves) illustrate the impact of different levels of degeneracy of the nonlinear mapping on the structure of local solution sets. In transcritical/pitchfork bifurcation, one solution curve is presumed. Indeed this is not necessary and a bifurcation structure with two crossing curves can be completely described via the partial derivatives of the nonlinear mapping. We remark that in the original result in [CR1], a slightly weaker smoothness condition is imposed on F: F is not necessarily C^2, but only the partial derivative $F_{\lambda u}$ exists. However to obtain (28) which indicates the direction of the bifurcation, more smoothness as ours is needed.

We illustrate the application of the transcritical and pitchfork bifurcation theorem by considering the diffusive logistic equation.

Example 4.3. *We consider the following diffusive logistic equation:*

$$\begin{cases} \Delta u + \lambda(u - u^p) = 0, & \text{in} \quad \Omega, \\ u = 0, & \text{on} \quad \partial\Omega, \end{cases} \tag{47}$$

where $p \geq 2$. For any $\lambda > 0$, $u = 0$ is always a solution to (47). We use Theorem 4.1 to analyze the bifurcation occurring at $\lambda = \lambda_1$. It is easy to verify that $F_u(\lambda, 0)w = \Delta w + \lambda w$, which is invertible if $\lambda \neq \lambda_i$. At $\lambda = \lambda_1$, $N(F_u(\lambda_1, 0)) = span\{\phi_1\}$, where $\phi_1 > 0$ is the principle eigenfunction. $R(F_u(\lambda_1, 0))$ is codimension one, and indeed $R(F_u(\lambda_1, 0)) = \{v \in Y : \int_\Omega \phi_1 v dx = 0\}$. Finally $F_{\lambda u}(\lambda_1, 0)[\phi_1] \notin R(F_u(\lambda_1, 0))$ since $\int_\Omega \phi_1 \cdot \phi_1 dx > 0$. Thus Theorem 4.1 can be applied to (47).

The calculation of $\lambda'(0)$ can be done by directly applying (28). But to illustrate the involved calculation, we calculate it directly. Differentiate (1) with respect to s once and twice, we obtain

$$\Delta u_s + \lambda f'(u)u_s + \lambda_s f(u) = 0, \tag{48}$$

$$\Delta u_{ss} + \lambda f'(u)u_{ss} + 2\lambda_s f'(u)u_s + \lambda f''(u)(u_s)^2 + \lambda_{ss} f(u) = 0. \tag{49}$$

Set $s = 0$ we obtain

$$\Delta u_s(0) + \lambda_1 f'(0)u_s(0) = 0, \tag{50}$$

$$(51) \qquad \Delta u_{ss}(0) + \lambda_1 f'(0) u_{ss}(0) + 2\lambda_s(0) f'(0) u_s(0) + \lambda_1 f''(0)[u_s(0)]^2 = 0.$$

By using integral by parts, and $u_s(0) = \phi_1$, we obtain

$$(52) \qquad 2\lambda_s(0) f'(0) \int_\Omega \phi_1^2(x)dx + \lambda_1 f''(0) \int_\Omega \phi_1^3(x)dx = 0,$$

and

$$(53) \qquad \lambda'(0) = -\frac{\lambda_1 f''(0) \int_\Omega \phi_1^3(x)dx}{2f'(0) \int_\Omega \phi_1^2(x)dx}.$$

In particular, for $f(u) = u - u^2$ when $p = 2$, $\lambda'(0) > 0$ and a transcritical bifurcation occurs. Thus for $\lambda \in (\lambda_1, \lambda_1 + \varepsilon)$ (which corresponds to $s \in (0, \varepsilon_1)$ since $\lambda'(0) > 0$), (47) has a positive solution with form $s\phi_1 + o(|s|)$.

For $p > 2$, (53) shows that $\lambda'(0) = 0$, and the smoothness of $\lambda(s)$ depends on p. When $p \geq 3$, $f \in C^3$ near $u = 0$, thus we have

$$(54) \qquad \lambda''(0) = -\frac{\lambda_1 f'''(0) \int_\Omega \phi_1^4(x)dx}{3f'(0) \int_\Omega \phi_1^2(x)dx},$$

by applying (29), where $\theta = 0$ since $F_{uu}(\lambda_1, 0)[\phi_1, \phi_1] = \lambda_1 f''(0)\phi_1^2 = 0$. Hence when $p = 3$, a pitchfork bifurcation occurs and $\lambda''(0) > 0$. Indeed one can show that for any $p \geq 2$, (47) has no positive solutions when $\lambda < \lambda_1$, and all positive solutions near $(\lambda, u) = (\lambda_1, 0)$ are on the right hand side of $\lambda = \lambda_1$, hence the bifurcation of positive solutions is always supercritical in that sense.

In general, we have the following result regarding (1):

Theorem 4.4. *Let $f \in C^2(\mathbf{R}^+)$, $f(0) = 0$ and $f'(0) > 0$. Then $\lambda_0 = \lambda_1/f'(0)$ is a bifurcation point. All positive solutions of (1) near $(\lambda_0, 0)$ have the form $(\lambda(s), s\phi_1 + sz(s))$ with $z(s)$ being a C^1 function satisfying $z(0) = 0$ for $s \in (0, \delta)$ and some $\delta > 0$, $\lambda(0) = \lambda_0$ and*

$$(55) \qquad \lambda'(0) = -\frac{\lambda_0 f''(0) \int_\Omega \phi_1^3(x)dx}{2f'(0) \int_\Omega \phi_1^2(x)dx}.$$

Example 4.5. *An alternate of the logistic growth is the Allee effect (see Section 1.4). For instance, consider the diffusive population model with weak Allee effect (see [SS]):*

$$(56) \qquad \begin{cases} \Delta u + \lambda u(1-u)(u+a) = 0, & \text{in} \quad \Omega, \\ u = 0, & \text{on} \quad \partial\Omega, \end{cases}$$

where $a > 0$. Again for any $\lambda > 0$, $u = 0$ is always a solution to (56). For $f(u) = u(1-u)(u+a)$, $f'(0) = a > 0$, hence $\lambda_0 = \lambda_1/a$ is a bifurcation point from Theorem 4.4. From (55), we obtain that

$$(57) \qquad \lambda'(0) = -\frac{\lambda_0 (1-a) \int_\Omega \phi_1^3(x)dx}{a \int_\Omega \phi_1^2(x)dx}.$$

Thus $\lambda'(0) > 0$ and the bifurcation of positive solutions is supercritical if $a > 1$, and $\lambda'(0) < 0$ and the bifurcation of positive solutions is subcritical (backward) if $0 < a < 1$. Later we shall show that the backward bifurcation indicates the nonuniqueness of positive solutions and it implies the bistability in the corresponding reaction-diffusion dynamics.

5 Global bifurcation

Bifurcation theorems in the last two sections are of local nature, as they only describe the structure of the solution set near the bifurcation point. Global bifurcation theorem gives information of the connected components of the solution set in the function spaces, and they are usually proved via topological tools such as Leray-Schauder degree theory. As preparation, we first review the concept of Leray-Schauder degree and a key topological lemma. No proof is given here, and the readers can consult standard references such as [Ch2, De].

Let X be a Banach space, and let U be an open bounded subset of X. Denote by $K(\overline{U})$ the set of compact operators from $\overline{U}$ to X, and define

$$M = \{(I - G, U, y) : U \subset X \text{ open bounded },G \in K(\overline{U}), \\ \text{and } y \notin (I - G)(\partial U)\}. \tag{58}$$

Then the Leray-Schauder degree $d : M \to \mathbf{Z}$ is a well-defined function, which satisfies the following properties:

1. $d(I, U, y) = 1$ if $y \in U$, and $d(I, U, y) = 0$ if $y \notin \overline{U}$;
2. (Additivity) $d(I - G, U, y) = d(I - G, U_1, y) + d(I - G, U_2, y)$ if U_1 and U_2 are disjoint open subsets of U so that $y \notin (I - G)(\overline{U} \backslash (U_1 \bigcup U_2))$;
3. (Homotopy invariance) Suppose that $h : [0, 1] \times \overline{U} \to X$ is compact and $y : [0, 1] \to X$ is continuous, and $y(t) \notin (I - G)(\partial U)$, then $D(t) = d(I - h(t, \cdot), U, y(t))$ is a constant independent of $t \in [0, 1]$.
4. (Existence) If $d(I - G, U, y) \neq 0$, then there exists $u \in U$ such that $u - G(u) = y$;
5. If for $G_1, G_2 \in K(\overline{U})$, $G_1(u) = G_2(u)$ for any $u \in \partial U$, then $d(I - G_1, U, y) = d(I - G_2, U, y)$.

Let L be a linear compact operator on X. From Riesz-Schauder theory, the set of eigenvalues of L is at most countably many, and the only possible limit point is $\lambda = 0$. For any eigenvalue λ of L, the subspace

$$X_\lambda = \bigcup_{n=1}^{\infty} \{u \in X : (L - \lambda I)^n u = 0\} \tag{59}$$

is finite dimensional, and $dim(X_\lambda)$ is the *algebraic multiplicity* of the eigenvalue λ. The *geometric multiplicity* of λ is defined as $dim\{u \in X : (L - \lambda I)u = 0\}$. The Leray-Schauder degree of a nonlinear mapping can be calculated from the following facts:

1. If L is a linear compact operator on X, then $d(I - \lambda L, B_R(0), 0) = (-1)^\beta$, where $B_R(v)$ is a ball centered at v with radius R, and β is the sum of algebraic multiplicity of eigenvalues μ of L satisfying $\lambda\mu > 1$.

2. Suppose that $G \in K(\overline{U})$, $u_0 \in U$ and $R > 0$ such that u_0 is the unique solution satisfies $u - G(u) = 0$ in $B_R(u_0)$, then the derivative $G'(u_0) : X \to X$ is a linear compact operator; if $\lambda = 1$ is not an eigenvalue of $G'(u_0)$, then $d(I - G, B_R(u_0), 0) = d(I - G'(u_0), B_R(0), 0)$ for some sufficiently small $R > 0$ (this number is also called *fixed point index* of u_0 with respect to G).

We also recall the following topological lemma (proof can be found in [Ch2, De]):

Lemma 5.1. *Let (M, d) be a compact metric space, and let A and B be close subsets of M such that $A \bigcap B = \emptyset$. Then there exist compact subsets M_A and M_B of M such that $M_A \bigcup M_B = M$, $M_A \bigcap M_B = \emptyset$, $M_A \supset A$, and $M_B \supset B$.*

Consider

$$F(\lambda, u) = u - \lambda L u - H(\lambda, u), \tag{60}$$

where $L : X \to X$ is a linear compact operator, and $H(\lambda, u)$ is compact on $U \subset \mathbf{R} \times X$ such that $||H(\lambda, u)|| = o(||u||)$ near $u = 0$ uniformly on bounded λ intervals. Note the conditions imply that $F_u(\lambda, 0) = I - \lambda L$, and if 0 is an eigenvalue of $F_u(\lambda_0, 0)$, then λ_0^{-1} must be an eigenvalue of the linear operator L. We will say that λ is a *characteristic value* of L, if λ^{-1} is an eigenvalue of L. Define $S = \{(\lambda, u) \in U : F(\lambda, u) = 0, u \neq 0\}$. We say $(\lambda_0, 0)$ is a *bifurcation point* for the equation (60) if $(\lambda_0, 0) \in \overline{S}$ ($\overline{S}$ is the closure of S).

Theorem 5.2. (Krasnoselski-Rabinowitz Global Bifurcation Theorem) *Let X be a Banach space, and let U be an open subset of $\mathbf{R} \times X$ containing $(\lambda_0, 0)$. Suppose that L is a linear compact operator on X, and $H(\lambda, u) : \overline{U} \to X$ is a compact operator such that $||H(\lambda, u)|| = o(||u||)$ as $u \to 0$ uniformly for λ in any bounded interval. If $1/\lambda_0$ is an eigenvalue of L with odd algebraic multiplicity, then $(\lambda_0, 0)$ is a bifurcation point. Moreover if C is the connected component of $\overline{S}$ which contains $(\lambda_0, 0)$, then one of the following holds:*

(i) C is unbounded in U;

(ii) $C \bigcap \partial U \neq \emptyset$; or

(iii) C contains $(\lambda_i, 0) \neq (\lambda_0, 0)$, such that λ_i^{-1} is also an eigenvalue of L.

Proof. First we prove that $(\lambda_0, 0)$ is a bifurcation point. Suppose not, then there exists a $R > 0$ such that in the region $O = \{(\lambda, u) : |\lambda - \lambda_0| \leq R, |u| \leq R\}$, the only solutions of $F(\lambda, u) = 0$ are $\{(\lambda, 0) : |\lambda - \lambda_0| \leq R\}$. We choose λ_-, λ_+ so that $\lambda_0 - R < \lambda_- < \lambda_0 < \lambda_+ < \lambda_0 + R$. From the homotopy invariance of the Leray-Schauder degree,

$$d(F(\lambda_-, \cdot), B_\rho(0), 0) = d(F(\lambda_+, \cdot), B_\rho(0), 0),$$

for any $\rho \in (0, R)$. For ρ small enough, $d(F(\lambda_\pm, \cdot), B_\rho(0), 0) = d(I - \lambda_\pm L, B_\rho(0), 0)$. But on the other hand,

$$|d(I - \lambda_+ L, B_\rho(0), 0) - d(I - \lambda_- L, B_\rho(0), 0)| = 1, \tag{61}$$

since λ_0^{-1} is the only eigenvalue of L in between λ_-^{-1} and λ_+^{-1}, and the algebraic multiplicity of λ_0^{-1} is odd. That is a contradiction. Thus $(\lambda_0, 0)$ is a bifurcation point.

Next we assume the stated alternatives do not hold, then C is bounded in U, $C \bigcap \partial U = \emptyset$, and $C \bigcap \{(\lambda, 0) \in U\} = \{(\lambda_0, 0)\}$. From the compactness of L and H, C is compact since it is bounded. Let $C_\varepsilon = \{(\lambda, u) \in U : dist((\lambda, u), C) < \varepsilon\}$. Let $A = C$ and $B = S \bigcap \partial C_\varepsilon$. From Lemma 5.1, there exists compact M_A and M_B such that $M_A \bigcap M_B = \emptyset$, $M_A \bigcup M_B = S \bigcap \overline{C_\varepsilon}$, $M_A \supset C$ and $M_B \supset S \bigcap \partial C_\varepsilon$. Hence there exists an open bounded $U_0 = M_A$ such that

$$C \subset U_0 \subset \overline{U_0} \subset U, \quad \text{and } \overline{S} \bigcap \partial U_0 = \emptyset. \tag{62}$$

Define $U_0(\lambda) = \{u \in X : (\lambda, u) \in U_0\}$ for $\lambda \in I$ where $I = \{\lambda \in \mathbf{R} : (\{\lambda\} \times X) \bigcap U_0 \neq \emptyset\}$. Then $D(\lambda) = d(F(\lambda, \cdot), U_0(\lambda), 0)$ is constant for $\lambda \in I$ since $\overline{S} \bigcap \partial U_0 = \emptyset$ and the homotopy invariance of $d(F, \Omega, 0)$, where $d(F(\lambda, \cdot), \Omega, 0)$ is the Leray-Schauder degree.

Since $(\lambda_0, 0)$ is the only intersection of C with the line $\{(\lambda, 0)\}$, U_0 can be chosen so that $U_0 \bigcap \{(\lambda, 0) \in U\} = [\lambda_0 - \delta, \lambda_0 + \delta] \times \{0\}$, and no any point λ in $[\lambda_0 - 2\delta, \lambda_0 + 2\delta]$ satisfies that λ^{-1} is an eigenvalue of L. We choose $\lambda_\pm$ which satisfy $\lambda_0 - \delta < \lambda_- < \lambda_0 < \lambda_+ < \lambda_0 + \delta$. We choose $\rho > 0$ small enough so that $F(\lambda, u) \neq 0$ for $\lambda \in [\lambda_+, \lambda_0 + 2\delta]$ and $u \in B_\rho(0) \backslash \{0\}$, and we also choose $\lambda^* > \lambda_0 + 2\delta$ such that $U_0(\lambda^*) = \emptyset$. From the homotopy invariance of the Leray-Schauder degree on $U_0 \backslash ([\lambda_+, \lambda^*] \times \overline{B}_\rho(0))$, we have

$$d(F(\lambda_+, \cdot), U_0(\lambda_+) \backslash \overline{B}_\rho(0), 0) = d(F(\lambda^*, \cdot), U_0(\lambda^*), 0) = 0. \tag{63}$$

For the same argument,

$$d(F(\lambda_-, \cdot), U_0(\lambda_-) \backslash \overline{B}_\rho(0), 0) = 0. \tag{64}$$

On the other hand, from the additivity of the Leray-Schauder degree,

$$D(\lambda_\pm) = d(F(\lambda_\pm, \cdot), U_0(\lambda_\pm) \backslash \overline{B}_\rho(0), 0) + d(F(\lambda_\pm, \cdot), B_\rho(0), 0). \tag{65}$$

Hence we obtain

$$d(F(\lambda_+, \cdot), B_\rho(0), 0) = d(F(\lambda_-, \cdot), B_\rho(0), 0). \tag{66}$$

For $\rho > 0$ small enough,

$$d(F(\lambda_\pm, \cdot), B_\rho(0), 0) = d(I - \lambda_\pm L, B_\rho(0), 0). \tag{67}$$

From the formula of Leray-Schauder degree of $I - \lambda L$, we have (61) again. But (61) is a contradiction with (67). Hence the alternatives in the theorem hold. $\square$

We apply Theorem 5.2 to (1). Recall that $K = (-\Delta)^{-1} : C^\alpha(\overline{\Omega}) \to C_0^{2,\alpha}(\overline{\Omega})$ is well defined as $K(f) = u$ that $u \in C_0^{2,\alpha}(\overline{\Omega})$ such that $-\Delta u = f$ for any $f \in C^\alpha(\overline{\Omega})$. We apply K to equation (1), and it becomes

$$G(\lambda, u) \equiv u - \lambda K f(u) = 0. \tag{68}$$

We set the domain of $G(\lambda, u)$ to be $\mathbf{R} \times C^\alpha(\overline{\Omega})$. Apparently if $u \in E \equiv C^\alpha(\overline{\Omega})$ satisfies (68), then u is a classical solution of (1). Combining with the maximum principle, we prove the global bifurcation theorem for the positive solutions of (1).

Theorem 5.3. *Let $f \in C^1(\mathbf{R}^+)$, $f(0) = 0$ and $f'(0) > 0$. Then $\lambda_0 = \lambda_1/f'(0)$ is a bifurcation point. Let $E = C^\alpha(\overline{\Omega})$, and let $S = \{(\lambda, u) \in \mathbf{R}^+ \times E : G(\lambda, u) = 0, u \neq 0\}$, where G is defined in (68). Then there exists a a connected component $\mathcal{C}$ of $\overline{S}$ such that $(\lambda_0, 0) \in \mathcal{C}$. Moreover, let $E^+ = \{u \in E : u(x) \geq 0$ in $\Omega\}$. Then $\mathcal{C}_+ = \mathcal{C} \bigcap (\mathbf{R}^+ \times E^+)$ is unbounded.*

Proof. First we extend f to $\mathbf{R}$ by an odd extension $f(u) = -f(-u)$ for $u < 0$. Then $f \in C^1(\mathbf{R})$, and the operator $G : \mathbf{R}^+ \times E \to E$ can be written as

$$G(\lambda, u) = u - \lambda f'(0) K u - \lambda K(f(u) - f'(0)u). \tag{69}$$

Then $H(\lambda, u) = \lambda K(f(u) - f'(0)u)$ is a nonlinear compact operator from the compactness of K, and apparently for λ in a bounded interval, $||H(\lambda, u)|| \to 0$ as $||u|| \to 0$ uniformly since $f \in C^1(\mathbf{R})$. When $\lambda = \lambda_0 \equiv \lambda_1/f'(0)$, $N(I - \lambda f'(0)K) \neq \emptyset$ where $\lambda_1 = \lambda_1(0)$ is the principal eigenvalue. Since K is symmetric, then the algebraic multiplicity is same as the geometric multiplicity, and it is $dim N(I - \lambda f'(0)K)$. It is known that the multiplicity of the principal eigenvalue is 1. Hence all conditions in Theorem 5.2 are satisfied, and there is a a connected component $\mathcal{C}$ of $\overline{S}$ such that $(\lambda_0, 0) \in \mathcal{C}$.

From the alternatives in Theorem 5.2, $\mathcal{C}$ is unbounded, or $\mathcal{C} \bigcap \partial(\mathbf{R}^+ \times E) \neq \emptyset$, or there is another λ_* such that $(\lambda_*, 0) \in \mathcal{C}$ and $\lambda_* f'(0)$ is another eigenvalue of $-\Delta$. If $(\lambda_a, u_a) \in \mathcal{C} \bigcap \partial(\mathbf{R}^+ \times E)$, then $\lambda_a = 0$, hence $u_a = 0$ from the uniqueness of Laplace equation. But near $(\lambda, u) = (0, 0)$, the only solutions of (1) are $(\lambda, 0)$ for $\lambda > 0$ from the implicit function theorem (Theorem 1.3), while $\overline{S}$ only contains $(\lambda_*, 0)$ such that $\lambda_* f'(0)$ is an eigenvalue of $-\Delta$. That is a contradiction since $-\Delta$ has no eigenvalues approaching 0. For the remaining cases, we note that from Hölder estimates, $\overline{S} \subset \mathbf{R} \times C_0^{2,\alpha}(\overline{\Omega})$. Hence $\mathcal{C} \subset \mathbf{R} \times C_0^{2,\alpha}(\overline{\Omega})$.

Define $E_2^+ = \{u \in C_0^{2,\alpha}(\overline{\Omega}) : u(x) \geq 0 \text{ in } \Omega\}$. Then the interior $int(E_2^+)$ of E_2^+ is non empty, and indeed $int(E_2^+) = \{u \in E_2^+ : u(x) > 0 \text{ in } \Omega, \partial u(x)/\partial\nu < 0 \text{ on } \partial\Omega\}$. Let $\mathcal{C}_+ = \mathcal{C} \bigcap (\mathbf{R}^+ \times E_2^+)$ and $\mathcal{C}_- = \mathcal{C} \bigcap (\mathbf{R}^+ \times (-E_2^+))$. From Theorem 4.4, $\mathcal{C}_+ \neq \emptyset$ and $\mathcal{C}_- \neq \emptyset$. We claim $\mathcal{C}_+ \bigcap (\mathbf{R}^+ \times \partial E_2^+) = \{(\lambda_0, 0)\}$. In fact, if $(\lambda, u) \in \mathcal{C}_+$, then $\lambda > 0$ from the argument in last paragraph, and u is a non-negative classical solution. From the maximum principle, either $u > 0$ or $u \equiv 0$. But if $u > 0$, (λ, u) is in the interior of $\mathbf{R}^+ \times \partial E^+$, that is contradiction. Hence $u \equiv 0$, and $\lambda f'(0)$ is an eigenvalue of $-\Delta$. Near the bifurcation point, all solutions have the form $(\lambda, s\phi_k)$ where ϕ_k is the corresponding eigenfunction. But ϕ_1 is the only eigenfunction of one sign, thus $\lambda f'(0) = \lambda_1$ and $\lambda = \lambda_0$. Similarly $\mathcal{C}_- \bigcap (\mathbf{R}^+ \times \partial(-E^+)) = \{(\lambda_0, 0)\}$.

Define $\mathcal{C}_1 = \mathcal{C}_+ \bigcup \{(\lambda_0, 0)\} \bigcup \mathcal{C}_-$. Then $\mathcal{C}_1 \subset O$, where $O = (\mathbf{R}^+ \times E_2^+) \bigcup (\mathbf{R}^+ \times (-E_2^+)) \bigcup O_1$ and $O_1 = \{(\lambda, u) : |\lambda - \lambda_0| + ||u|| \le \varepsilon\}$ for some small $\varepsilon > 0$. Moreover, $\mathcal{C}_1 \bigcap \partial O = \emptyset$ where ∂O is the boundary of O. Since $\mathcal{C}$ is the connected component of $\overline{S}$, then $\mathcal{C} = \mathcal{C}_1$ from the definition of connected component since $\mathcal{C}_1 \subset \mathcal{C}$ and $\mathcal{C}_1 \subset int(O)$. As a consequence, $\mathcal{C}$ cannot contain another $(\lambda_*, 0)$ so that $\lambda_* f'(0)$ is another eigenvalue of $-\Delta$. Hence $\mathcal{C}$ is unbounded. From our definition, $f(u)$ is an odd function, then $\mathcal{C}_-$ and $\mathcal{C}_+$ are symmetric in the sense that if $(\lambda, u) \in \mathcal{C}_+$ then $(\lambda, -u) \in \mathcal{C}_-$. Therefore $\mathcal{C}_+$ is unbounded. $\square$

It is useful to remark that Theorem 5.3 implies that the global continuum $\mathcal{C}_+$ is also unbounded in $\mathbf{R}^+ \times C^0(\overline{\Omega})$. In fact, if $\mathcal{C}_+$ is bounded in $\mathbf{R}^+ \times C^0(\overline{\Omega})$, $\mathcal{C}_+$ have uniform L^p norm for any $p > 1$, f is C^1 and f can be assumed as bounded, then $\mathcal{C}_+$ also satisfy uniform L^p estimates thus uniformly bounded in $W^{2,p}$ norm. From Sobolev embedding theorem, $\mathcal{C}_+$ is also bounded in C^α norm where $\alpha < 2 - (n/p)$, that contradicts Theorem 5.3. On the other hand, if one can establish uniform *a priori* estimates for the positive solutions of (1), *i.e.* given $\Lambda > 0$, for $\lambda \in [0, \Lambda]$, $||u||_\infty \le K_\Lambda$ for any positive solution (λ, u), then $\mathcal{C}_+$ can be extended to $\lambda = \infty$. That is, let p_+ be the projection of $\mathcal{C}_+$ onto the λ-axis, then $p_+ \supset (\lambda_0, \infty)$. In that case, the *a priori* estimates and global bifurcation theorem give the existence for not only large λ but a continuum of solutions.

Example 5.4. *First we continue the discussion started in Example 4.3. Consider*

$$\begin{cases} \Delta u + \lambda(u - u^p) = 0, & \text{in} \quad \Omega, \\ u = 0, & \text{on} \quad \partial\Omega, \end{cases} \tag{70}$$

where $p \ge 2$. *We have shown in Example 4.3 that* $\lambda = \lambda_1$ *is a bifurcation point, and the solution set near* $(\lambda_1, 0)$ *is a curve. Now applying Theorem 5.3, then the branch bifurcating from* $(\lambda_1, 0)$ *is indeed unbounded. Moreover the maximum principle implies that* $\max_{x \in \overline{\Omega}} u(x) < 1$ *for any positive solution* (λ, u). *From remark above, this implies the existence of a positive solution* (λ, u) *of* (70) *for any* $\lambda \in (\lambda_1, \infty)$, *and* $(\lambda, u) \in \mathcal{C}_+$, *the continuum emanating from* $(\lambda_1, 0)$. *In fact* $p_+ = (\lambda_0, \infty)$ *in this case since* (70) *has no positive solution when* $\lambda \le \lambda_1$.

References

[Ch2] Chang, Kung-Ching, Methods in nonlinear analysis. Springer Monographs in Mathematics. Springer-Verlag, Berlin, (2005).

[CR1] Crandall, Michael G.; Rabinowitz, Paul H., Bifurcation from simple eigenvalues. *J. Functional Analysis*, **8** (1971), 321–340.

[CR2] Crandall, Michael G.; Rabinowitz, Paul H., Bifurcation, perturbation of simple eigenvalues and linearized stability. *Arch. Rational Mech. Anal.* **52** (1973), 161–180.

[De] Deimling, Klaus, Nonlinear functional analysis. Springer-Verlag, Berlin-New York, (1985).

[LSW] Liu, Ping; Shi, Junping; Wang, Yuwen, Imperfect transcritical and pitchfork bifurcations. *Jour. Func. Anal.* **251** (2007), no. 2, 573–600.

[Nir] Nirenberg, Louis, Topics in nonlinear functional analysis. *Courant Lecture Notes in Mathematics*, **6**. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, (2001).

[Pe] Perko, Lawrence, Differential equations and dynamical systems. First edition. *Texts in Applied Mathematics*, **7**. Springer-Verlag, New York, 1991.

[R] Rabinowitz, Paul H., Some global results for nonlinear eigenvalue problems. *J. Func. Anal.* **7** (1971), 487–513.

[S] Shi, Junping, Persistence and bifurcation of degenerate solutions, *Jour. Func. Anal.* **169**, (1999), No. 2, 494–531.

[SS] Shi, Junping; Shivaji, Ratnasingham, Persistence in reaction diffusion models with weak Allee effect. *Jour. Math. Biol.* **52** (2006), no. 6, 807–829.