7.1

Bifurcation Theory

H. Kielhöfer, Bifurcation theory: an introduction with applications to pde, second edition, Springer 2011

Yu. A. Kuznetsov, Elements of applied bifurcation theory, third edition, Springer 2004

M. Golubitsky & D. G. Schaeffer, Singularities and groups in bifurcation theory, Vol. 1, Springer 1985

— & I. Stewart, Vol. 2, Springer, 1988

Part 1 A model problem

Part 2 General theory (Lyapunov-Schmidt reduction, transcritical bifurcation, saddle-node bifurcation); follows Kielhöfer

Part 3 Application of the theory to the model problem

Part 4 The principle of exchange of stability (again following Kielhöfer)

Part 1 A model problem

$$\begin{cases} \dfrac{\partial u}{\partial t} = d\,\dfrac{\partial^2 u}{\partial x^2} + r f(u), \quad 0 < x < L \\ u(t,0) = 0 = u(t,L) \\ u(0,x) = u_0(x) \end{cases}$$

f "smooth"

$f(0) = 0$

$Df(0) = 1$

7.2 By scaling of x and t we can achieve that

$$d = 1 \quad , \quad L = 1 \quad , \quad r_{new} = \frac{r_{old} L^2}{d}$$

Note that variation of the new r can reflect variation of the three parameters d, L, r in the original problem. Our aim is to study how the number of steady states varies with the parameter r_{new} (but from now on we omit the subscript "new"). So we want to study, as a representative bifurcation problem,

$$(7.1) \quad \begin{cases} 0 = \phi'' + r f(\phi) \; , \quad 0 < x < 1 \\ \phi(0) = 0 = \phi(1) \end{cases}$$

where $f \in C^3$, or even smoother, and $f(0) = 0$, $Df(0) = 1$. Note that for every value of r (7.1) has the trivial solution $\phi \equiv 0$. The aim of (local) bifurcation theory is to find parameter values such that (7.1) has nontrivial small solutions (where "small" means that in the Banach space of functions that we consider, they are at a small distance from the trivial solution).

If we work with only one space of functions, the derivative is an unbounded linear operator. There are at least two ways to amend this troublesome feature:

- introduce another function space such that the second derivative becomes a <u>bounded</u> operator from one space to the other

- use the solution operator for the <u>linear</u> problem that i) includes the second derivative
  ii) includes the boundary conditions
  iii) is indeed solvable

  to reformulate (7.1) as a nonlinear fixed point problem

Here we follow the second approach. In lectures:4.pdf we established that for given $h \in L_2$ we have

$$\left.\begin{array}{l}\text{weak formulation}\\ \text{of} \quad \phi'' + h = 0\\ \qquad \phi(0) = 0 = \phi(1)\end{array}\right\} \iff \phi = \mathcal{G} h \tag{7.2}$$

where $\mathcal{G} : L_2 \to L_2$ is a <u>compact</u> linear operator that is <u>positive</u> and <u>self-adjoint</u>. The compactness was a consequence of $\mathcal{R}(\mathcal{G}) \subset H_0^1$ (here $\mathcal{R}$ denotes the <u>range</u>). We also found
(σ denotes <u>spectrum</u>) $\sigma(\mathcal{G}) = \{\lambda_k\}_{k=1}^{\infty}$
with $\lambda_k = \frac{1}{(k\pi)^2}$ eigenvalues with corresponding eigenfunctions $\phi_k(x) = \sqrt{2}\sin(k\pi x)$. The $\{\phi_k\}$ form a complete orthonormal system in L_2.

7.4 Interlude

G has the explicit representation

$$(Gh)(x) = \int_0^1 g(x,\xi)\, h(\xi)\, d\xi \qquad (7.3)$$

where

$$g(x,\xi) = \begin{cases} x\,(1-\xi)\,, & x < \xi \\ \xi\,(1-x)\,, & x > \xi \end{cases} \qquad (7.4)$$

(so note that $g(x,\xi) = g(\xi,x)$; we call g the <u>Green function</u>)

The function g can be found as follows. Formally, g is characterized by

$$\begin{cases} \dfrac{\partial^2 g}{\partial x^2}(x,\xi) = \delta(x-\xi) = \delta_\xi(x) \\ g(0,\xi) = 0 = g(1,\xi) \end{cases}$$

The kernel of the second derivative consists of functions of the form $ax+b$. From the boundary conditions we conclude that we should take

$$g(x,\xi) = \begin{cases} a(\xi)\,x & x<\xi \\ \tilde{a}(\xi)(x-1) & x>\xi \end{cases}$$

To determine $a(\xi)$ and $\tilde{a}(\xi)$ we impose the following two conditions:

i) the function $x \mapsto g(x,\xi)$ is continuous in $x = \xi$

ii) the derivative of $x \mapsto g(x,\xi)$ has a jump of $+1$ in $x = \xi$

Condition ii) is motivated by the relationship

$$\delta(x-\xi) = \frac{d}{dx} H(x-\xi)$$

(which is made precise in the theory of distributions)

7.5 where H is the Heaviside function, i.e., $H(x)=0$ for $x<0$ and $H(x)=1$ for $x>0$.

So we want that

$$\text{i)}\quad a(\xi)\,\xi = \tilde{a}(\xi)(\xi-1)$$

$$\text{ii)}\quad a(\xi)+1 = \tilde{a}(\xi)$$

which readily leads to $a(\xi)=\xi-1$ and $\tilde{a}(\xi)=\xi$.

<u>Exercise</u> Check that $(\mathcal{G}h)(0)=0=(\mathcal{G}h)(1)$ and that $\left(\frac{d^2}{dx^2}\mathcal{G}h\right)(x)=h(x)$

<u>End of interlude</u>

Before we can apply $\mathcal{G}$ to reformulate (7.1), we need to pay attention to a subtle aspect. The function f is a smooth map from $\mathbb{R}$ to $\mathbb{R}$. It induces a map, called the <u>Nemitskii operator</u> and here denoted by $\hat{f}$, that maps functions of the variable x to functions of x:

$$(\hat{f}(\phi))(x) = f(\phi(x)) \qquad (7.5)$$

If, for instance, $f(u)=u^2$, then $\hat{f}$ does <u>not</u> map L_2 to L_2. Even if f is bounded (in the sense that $f(\mathbb{R})$ is a bounded subset of $\mathbb{R}$), we are going to use Taylor expansion to approximate $f(u)$ for small values of u by, say, a second order polynomial and thus run into the same problem. In other words, even if $\hat{f}$ is defined on L_2, we cannot take its second derivative. This makes L_2 ill suited for

7.6/ dealing with nonlinear problems! Now note that the range of $\mathcal{G}$ is contained in H_0^1 and that, for a one-dimensional space variable, elements of H_0^1 are represented by a continuous function. This observation motivates to study (7.1) with the Banach space

$$(7.6) \qquad X = \{ \psi \in C[0,1] : \psi(0) = 0 = \psi(1) \}$$

(provided with the supremum norm) as the underlying space of functions.

Clearly $\mathcal{G}$ defined by (7.3) maps X into X. Clearly the eigenfunctions $\phi_k(x) = \sin(k\pi x)$ are elements of X. In fact, due to the regularizing effect of solving an elliptic equation, the spectral properties of $\mathcal{G}$ defined on L_2 survive when we restrict $\mathcal{G}$ to X. So by restricting to X we do not lose much concerning $\mathcal{G}$ (even though X is not a Hilbert space; it is this fact that makes it attractive to analyse <u>first</u> $\mathcal{G}$ as defined on L_2 and only then restrict to X), but we gain a lot concerning $\hat{f}$:

<u>Exercise</u> Assume $f: \mathbb{R} \to \mathbb{R}$ is C^1. Show that $\hat{f}: X \to X$ defined by (7.5) is C^1 and compute its derivative. Do you see any problem with extending this result to higher order derivatives?

We are now ready to reformulate (7.1) as

~~7.7~~

$$(7.7) \quad H(\phi, r) = 0$$

where $H : X \times \mathbb{R} \to X$ is the C^3 map defined by

$$(7.8) \quad H(\phi, r) = \phi - r\mathcal{G}\hat{f}(\phi)$$

The linearization of H (with respect to the first variable) at the trivial solution $\phi = 0$ is $D_1 H(0, r)$ with (since $Df(0) = 1$)

$$(7.9) \quad D_1 H(0,r)\psi = \psi - r\mathcal{G}\psi$$

So for all values of r except $(k\pi)^2$, $k = 1, 2, 3, \ldots$, $D_1 H(0,r) : X \to X$ is bijective and has a bounded inverse. Indeed, we can solve the equation $D_1 H(0,r)\psi = h$, i.e.,

$$\psi - r\mathcal{G}\psi = h$$

by representing h as $h = \sum_{j=1}^{\infty} d_j \phi_j$ with $d_j = \langle \phi_j, h \rangle$ and next determining the coefficients in $\psi = \sum_{j=1}^{\infty} c_j \phi_j$ from

$$c_j - \frac{r}{(j\pi)^2} c_j = d_j$$

as

$$c_j = \frac{d_j}{1 - \frac{r}{(j\pi)^2}}$$

If $r = (k\pi)^2$ this only works if $d_k = 0$. And if indeed $d_k = 0$ we still get a <u>unique</u> solution if we require that

 $c_k = 0$ as well. So we have the direct sum decomposition

$$(7.10) \qquad X = \mathcal{N}_k \oplus \mathcal{R}_k$$

where $\mathcal{N}_k = \operatorname{span}\{\phi_k\}$ ~~and~~ is the nullspace of $D_1 H(0,(k\pi)^2)$ and $\mathcal{R}_k$ is the range of $D_1 H(0,(k\pi)^2)$ and the restriction

$$D_1 H(0,(k\pi)^2)\Big|_{\mathcal{R}_k} : \mathcal{R}_k \to \mathcal{R}_k$$

is a bijective map, so has a bounded inverse.

The projection onto $\mathcal{N}_k$ associated with (7.10) is

$$(7.11) \qquad P_k \psi = \langle \phi_k, \psi \rangle \phi_k$$

<u>Summary of Part 1</u>: we want to investigate the problem (7.7) with H defined by (7.8).
The linearized map $D_1 H(0,r)$ given in (7.9) is invertible unless $r = (k\pi)^2$ for some $k \in \mathbb{N}$.
In the latter case we can decompose X as in (7.10) and the restriction of the linearized map to $\mathcal{R}_k$ is invertible.

Part 2 is devoted to general considerations for a slightly more general problem that has essentially the same structure.

7.9 <u>Part 2</u> General theory
2.1 Lyapunov – Schmidt reduction
2.2 Transcritical bifurcation
2.3 Saddle-node bifurcation

2.1 Lyapunov-Schmidt reduction

Let X, Y and Z be Banach spaces and let

$$F : X \times Y \to Z$$

be at least C^1 (later on we shall require C^2 or even C^3)
Let $(x_0, y_0) \in X \times Y$ be such that

(7.12) $$F(x_0, y_0) = 0$$

We want to determine ~~to determine~~ the structure of the local set of solutions of the equation

(7.13) $$F(x, y) = 0$$

where "local" means "in a neighbourhood of (x_0, y_0)".

If $D_1F(x_0, y_0) : X \to Z$ is bijective, we can invoke the IFT (Implicit Function Theorem) and conclude that all solutions in a neighbourhood of (x_0, y_0) are of the form $(\hat{x}(y), y)$ where $y \mapsto \hat{x}(y)$ is a C^1 map from a neighbourhood of y_0 in Y to X and $\hat{x}(y_0) = x_0$. So that case is settled.

Next <u>assume</u> that $D_1F(x_0, y_0)$ has a nontrivial nullspace $\mathcal{N}$. <u>Assume</u>, in addition, that $\mathcal{N}$ has finite dimension and that the range $\mathcal{R}$

of $D_1F(x_0, y_0)$ is closed and has finite codimension. ($D_1F(x_0,y_0)$ is called a Fredholm operator if it has these properties)

So we can decompose

$$(7.14) \quad \begin{aligned} X &= \mathcal{N} \oplus X_0 \\ Z &= \mathcal{R} \oplus Z_0 \end{aligned}$$

where X_0 ~~is closed~~ and $\mathcal{R}$ are closed and both $\mathcal{N}$ and Z_0 have finite dimension. By the Closed Graph Theorem the associated projections

$$P : X \to \mathcal{N} \quad , \quad Q : Z \to Z_0$$

are continuous. The key point is that

$$(I - Q) D_1F(x_0, y_0)\Big|_{X_0} : X_0 \to \mathcal{R}$$

is bijective.

For notational convenience we use the symbol v to denote a generic element of $\mathcal{N}$ and the symbol w to denote a generic element of X_0. In particular we write

$$x = v + w \qquad \text{with } v = Px \ , \ w = (I-P)x$$

Next we write $F(x,y) = F(v+w, y) = 0$ as the couple of equations

$$(7.15) \quad \begin{aligned} QF(v+w, y) &= 0 \\ (I-Q)\, F(v+w, y) &= 0 \end{aligned}$$

Motivated by the second of these we define the C^1 map

7.11

$$\tilde{F} : \mathcal{W} \times X_0 \times Y \to \mathbb{R}$$

by

$$(7.16) \qquad \tilde{F}(v, w, y) = (I - Q) F(v + w, y)$$

We also introduce the notation

$$(7.17) \qquad v_0 = P x_0 \quad , \quad w_0 = (I - P) x_0$$

From (7.16) it directly follows that

$$D_2 \tilde{F}(v, w, y) = (I - Q) D_1 F(v + w, y) \Big|_{X_0}$$

and hence

$$D_2 \tilde{F}(v_0, w_0, y_0) = (I - Q) D_1 F(x_0, y_0) \Big|_{X_0}$$

is a bijective map. So we are again in a position where we can apply the IFT.

<u>Theorem</u> There exists a C^1 function

$$\hat{w} : \mathcal{W} \times Y \to X_0$$

defined on a neighbourhood of (v_0, y_0) such that

i) $\hat{w}(v_0, y_0) = w_0$

ii) $\tilde{F}(v, \hat{w}(v, y), y) = 0$

iii) in a neighbourhood of (v_0, w_0, y_0) the equation $\tilde{F}(v, w, y) = 0$ has no other solutions

iv) the element $D_1 \hat{w}(v_0, y_0)$ of $\mathcal{L}(\mathcal{N}, X_0)$ equals zero

7.12 <u>Proof</u> i) – iii) are a direct consequence of the IFT applied to the equation $\tilde{F} = 0$ with known solution (v_0, w_0, y_0). If we differentiate the identity

$$(I - Q) F(v + \hat{w}(v,y), y) = 0$$

with respect to v we obtain

$$(I - Q) D_1 F(v + \hat{w}(v,y), y)(I_W + D_1 \hat{w}(v,y)) = 0$$

where $I_W : W \to X$ is the identity map and where $D_1 \hat{w}(v,y)$ is a bounded linear map from W to X_0. At $(v,y) = (v_0, y_0)$ this reduces to

$$(I - Q) D_1 F(x_0, y_0) D_1 \hat{w}(v_0, y_0) = 0$$

since W is the nullspace of $D_1 F(x_0, y_0)$. Since $(I - Q) D_1 F(x_0, y_0)$ is, as a map from X_0 to R, bijective, this last identity can only hold if $D_1 \hat{w}(v_0, y_0) = 0$ $\square$

Thus we have solved the second of the two equations in (7.15) for w as a function of v and y, locally near the known solution (v_0, w_0, y_0). By substituting $w = \hat{w}(v,y)$ in the first equation of (7.15) we obtain a <u>finite dimensional</u> problem, since this is an identity in Z_0 and $\dim Z_0 < \infty$. If $\dim Y < \infty$ the unknown (v, y) is finite dimensional too, since $v \in W$

and $\dim W < \infty$. In order to formalize this description in words we define a C^1 map

$$\Phi : W \times Y \to Z_0$$

by

$$\Phi(v,y) = QF(v + \hat{w}(v,y), y) \tag{7.18}$$

and note that we have reduced the equation (7.13) to the so-called <u>bifurcation equation</u>

$$\Phi(v,y) = 0 \tag{7.19}$$

This procedure is called <u>Lyapunov–Schmidt reduction.</u>

For later use we close this subsection with the following observation.

<u>Lemma</u> $D_1 \Phi(v_0, y_0)$ is the zero element of $\mathcal{L}(W, Z_0)$

<u>Proof</u> $D_1 \Phi(v_0, y_0) = Q D_1 F(x_0, y_0)\left(I_W + D_1 \hat{w}(v_0, y_0)\right) = 0$ since W is the nullspace of $D_1 F(x_0, y_0)$ and, by point iv) of the theorem above, $D_1 \hat{w}(v_0, y_0) = 0$ $\square$

Comment on terminology: $\dim W - \dim Z_0$ is called the <u>index</u> of the <u>Fredholm operator</u> $D_1 F(x_0, y_0)$

From now on we are going to assume that the index is zero.

2.2 Transcritical bifurcation

We now make the following additional assumptions:

$$F \text{ is } C^2$$

$$\dim W = \dim Z_0 = 1$$

$$\dim Y = 1, \text{ in fact } Y = \mathbb{R}$$

$$F(0, y) = 0 \quad \forall y \in Y = \mathbb{R}$$

So (7.13) has the branch (i.e., one-parameter family) of solutions $(0, y)$, $y \in \mathbb{R}$ and, as noted before, if $D_1 F(0, y_0)$ is bijective there are no other solutions for y in a neighbourhood of y_0. But here we consider the situation that $D_1 F(0, y_0)$ has a one-dimensional nullspace W, while the range R of $D_1 F(0, y_0)$ is complemented by a one-dimensional subspace Z_0.

Let $\hat{v}_0 \in W$ satisfy $\| \hat{v}_0 \| = 1$. Note that

$$W = \operatorname{span}[\hat{v}_0]$$

<u>Remark</u> For a function $g: \mathbb{R} \to Z$ we interpret the derivative $Dg(y)$ in a point $y \in \mathbb{R}$ as an element of Z even though, strictly speaking, $Dg(y)$ is a linear map from $\mathbb{R}$ to Z and the element should be denoted by $Dg(y)1$. In the same vein we shall interpret $D^2_{12} F(x, y)$ ~~als~~ as an element of $\mathcal{L}(X, Z)$. Note that, since F is C^2, $D^2_{12} F(x, y) = D^2_{21} F(x, y)$

7.15

Crandall – Rabinowitz Theorem

Assume that

$$D^2_{12} F(0, y_0)\, \hat{v}_0 \notin \mathcal{R}$$

Then there exists ~~an~~ a nontrivial C^1 curve

$$\{(\hat{x}(s), \hat{y}(s)) : -\delta < s < \delta\ ,\ (\hat{x}(0), \hat{y}(0)) = (0, y_0)\}$$

such that

$$F(\hat{x}(s), \hat{y}(s)) = 0 \quad , \quad -\delta < s < \delta$$

and all solutions of (7.13) in a neighbourhood of $(0, y_0)$ are either of the form $(0, y)$ or of the form $(\hat{x}(s), \hat{y}(s))$

<u>Proof</u> By Lyapunov – Schmidt reduction we know that all solutions near $(0, y_0)$ can be found by solving the <u>one-dimensional</u> bifurcation equation (recall (7.19) and (7.18))

$$\Phi(v, y) = QF(v + \hat{w}(v, y), y) = 0$$

where both v and y are one-dimensional variables. Note carefully that now v_0 introduced in (7.17) is zero and do not confuse v_0 with the unit vector $\hat{v}_0$ that spans W. Since, for all $y \in \mathbb{R}$, $F(0, y) = 0$ clearly $(I - Q)F(0, y) = 0$ and consequently

(7.20) $$\hat{w}(0, y) = 0 \quad , \quad y \in \mathbb{R}$$

(7.21) $$\Phi(0, y) = 0 \quad , \quad y \in \mathbb{R}$$

7.16/ Since Φ is C^1 (in fact C^2) and (7.21) holds we can write

$$\Phi(v,y) = \int_0^1 \frac{d}{dt}\Phi(tv,y)\,dt = \int_0^1 D_1\Phi(tv,y)\,v\,dt$$

and, by putting $v = s\hat{v}_0$

$$\Phi(s\hat{v}_0, y) = s\int_0^1 D_1\Phi(ts\hat{v}_0, y)\hat{v}_0\,dt$$

It follows that the bifurcation equation has the trivial solution corresponding to $s=0$ while nontrivial solutions are found by solving

$$(7.22) \qquad 0 = \tilde{\Phi}(s,y) := \int_0^1 D_1\Phi(ts\hat{v}_0, y)\hat{v}_0\,dt$$

Since F is C^2, it is guaranteed that Φ is C^2 and hence $\tilde{\Phi}$ is C^1. From the lemma on page 7.13 we infer that

$$\tilde{\Phi}(0,y) = 0$$

We want to solve $\tilde{\Phi}(s,y)=0$ for y as a function of s by invoking the IFT, so we now set out to compute $D_2\tilde{\Phi}(0,y_0)$ and show that it isn't equal to zero. As a first step we note that

$$D_2(D_1\Phi(v,y)\hat{v}_0) = (1) + (2) + (3)$$

with

$$(1) = QD_{11}^2F(v+\hat{w}(v,y),y)[\hat{v}_0 + D_1\hat{w}(v,y)\hat{v}_0,\ D_2\hat{w}(v,y)]$$

$$(2) \quad = Q\, D_1 F(v + \hat{w}(v,y), y)\, D^2_{21}\, \hat{w}(v,y)\, \hat{v}_0$$

$$(3) \quad = Q\, D^2_{21} F(v + \hat{w}(v,y), y) \left(\hat{v}_0 + D_1 \hat{w}(v,y)\, \hat{v}_0 \right)$$

(from the definition (7.18) of Φ we deduce

$$D_1 \Phi(v,y)\, \hat{v}_0 = Q\, D_1 F(v + \hat{w}(v,y), y) (\hat{v}_0 + D_1 \hat{w}(v,y) \hat{v}_0)$$

and next we apply D_2 and use the chain rule again to find these three terms, corresponding to the three places where y appears at the right hand side)

Now put $(v,y) = (0, y_0)$. Then $(1) = 0$ because (7.20) implies that $D_2 \hat{w}(0, y_0) = 0$. And $(2) = 0$ because Q annihilates the range $\mathcal{R}$ of $D_1 F(0, y_0)$. Point iv) of the theorem on page 7.11 shows that (3) reduces to

$$Q\, D^2_{21} F(0, y_0)\, \hat{v}_0$$

Recalling the assumption of the theorem that $D^2_{21} F(0, y_0) \hat{v}_0$ does not belong to $\mathcal{R}$, we conclude that $Q\, D^2_{21} F(0, y_0)\, \hat{v}_0 \neq 0$. Hence

$$D_2 \tilde{\Phi}(s,y) \Big|_{(0,y_0)} = \int_0^1 D_2 \left(D_1 \Phi(ts\hat{v}_0, y)\, \hat{v}_0 \right) \Big|_{(0,y_0)} dt \neq 0$$

So we can indeed solve $\tilde{\Phi}(s,y) = 0$ for y as a function of s. We call the solution $\hat{y}(s)$, note that $\hat{y}(0) = y_0$ and put

$$x(s) = s\hat{v}_0 + \hat{w}(s\hat{v}_0, \hat{y}(s)) \qquad (7.23)$$

□

So the nontrivial solutions of (7.13) are on a curve in $X \times \mathbb{R}$ parameterized by s, where s is the coordinate of the W component of x. As we show next, the X-component of the curve is, in $(0, y_0)$, tangent to W:

$$\frac{d}{ds} \hat{x}(s) \Big|_{s=0} = \hat{v}_0$$

Indeed, $\frac{d}{ds} \hat{x}(s) = \hat{v}_0 + D_1 \hat{w}(s\hat{v}_0, \hat{y}(s))\, \hat{v}_0$

$$+ D_2 \hat{w}(s v_0, \hat{y}(s)) \frac{d\hat{y}}{ds}(s)$$

and if we put $s=0$ the second term vanishes because of point iv) of the theorem on page 7.11 while the third term vanishes because (7.20) implies that $D_2 w(0, y_0) = 0$. In Part 4 we are going to derive formulas for $\frac{d\hat{y}}{ds}(0)$ and $\frac{d^2\hat{y}}{ds^2}(0)$ since these can provide useful information concerning the stability of the nontrivial solutions.

2.3 <u>Saddle-Node bifurcation</u> (also known as Fold bifurcation or Turning point bifurcation)

<u>Theorem</u> Let $Y = \mathbb{R}$, $\dim W = 1 = \dim Z_0$, let F be C^1 and assume that $D_2 F(x_0, y_0) \notin \mathcal{R}$. Then there exists a C^1 curve $\{(\hat{x}(s), \hat{y}(s)) : -\delta < s < \delta, (\hat{x}(0), \hat{y}(0)) = (x_0, y_0)\}$

such that $F(\hat{x}(s), \hat{y}(s)) = 0$ for $-\delta < s < \delta$ and all solutions of (7.13) in a neighbourhood of (x_0, y_0) belong to this curve.

<u>Proof</u> By Lyapunov-Schmidt reduction we reduce (7.13) to

$$\Phi(v,y) = QF(v + \hat{w}(v,y), y) = 0$$

which is one equation (since Z_0 has dimension 1) in two unknowns (since $Y = \mathbb{R}$ and $\dim \mathcal{W} = 1$). We are going to solve this equation for y as a function of v by way of the IFT, so we need to check that $D_2 \Phi(v_0, y_0) \neq 0$. Now

$$D_2 \Phi(v_0, y_0) = Q\left[D_1 F(x_0, y_0) D_2 \hat{w}(v_0, y_0) + D_2 F(x_0, y_0)\right]$$

Since Q maps the range $\mathcal{R}$ of $D_1 F(x_0, y_0)$ to zero, the first term vanishes. The assumption of the theorem guarantees that the second term does <u>not</u> vanish. So $D_2 \Phi(v_0, y_0) \neq 0$ and we can indeed solve for y as a function of v. Let $\hat{v}_0$ be a unit vector that spans $\mathcal{W}$ and introduce the parameter s via $v = v_0 + s\hat{v}_0$. □

Note that once again the tangent vector to the X-component of the curve in (x_0, y_0) is $\hat{v}_0$, so belongs to $\mathcal{W}$. But, as we now show, necessarily

$$\frac{d}{ds} \hat{y}(s)\Big|_{s=0} = 0$$

so if the second derivative does not vanish, the solutions either exist for $y > y_0$ or for $y < y_0$ (we return to this topic in Part 4).

If we denote the solution y as a function of v by $\tilde{y}(v)$ then $\hat{y}(s) = \tilde{y}(v_0 + s\hat{v}_0)$ and, by definition, $\Phi(v, \tilde{y}(v)) = 0$. By differentiation of this identity with respect to v we find that, after taking $v = v_0$,

$$D_1 \Phi(v_0, y_0) + D_2 \Phi(v_0, y_0) D\tilde{y}(v_0) = 0$$

The lemma on page 7.13 tells us that the first term vanishes. The proof of the theorem above was based on the observation that $D_2 \Phi(v_0, y_0) \neq 0$. So we must have that $D\tilde{y}(v_0) = 0$. And consequently

$$\frac{d}{ds} \hat{y}(s)\Big|_{s=0} = D\tilde{y}(v_0)\hat{v}_0 = 0$$

<u>Part 3</u> is simply an

<u>Exercise</u> Elaborate in detail how the Crandall–Rabinowitz Theorem of page 7.15 is applicable to the problem (7.7)–(7.8). In particular, specify how to choose the following ingredients of the general theory in order to deal with (7.7)–(7.8):
a) X, Y and Z b) the point(s) (x_0, y_0) c) the subspaces $\mathcal{N}$ and $\mathcal{R}$ d) the complementary spaces X_0 and Z_0 e) the projections P and Q f) the condition $D^2_{12} F(0, y_0)\hat{v}_0 \notin \mathcal{R}$

7.21 <u>Exercise</u> By way of the preceding exercise you now know that for every integer k the boundary value problem (7.1) has a nontrivial solution branch passing through $(0, (k\pi)^2)$.

- show that for $k \geq 2$ and small values of the parameter s, the solutions ϕ take both positive and negative values in $(0,1)$
- if $k=1$, what can you say about the sign of the solutions as a function of the spatial variable, for small values of the parameter s?

Part 4 The direction of bifurcation and the Principle of Exchange of Stability

Consider the setting of Section 2.2, i.e., transcritical bifurcation. Let

$$Z_0 = \text{span}\,[z_0]$$

(7.24)

$$Qz = \langle z_0^*, z \rangle z_0$$

with i) $\langle z_0^*, z_0 \rangle = 1$ ii) $\langle z_0^*, z \rangle = 0 \quad \forall z \in \mathcal{R}$

<u>Exercise</u> Show that

$$D\hat{y}(0) = -\frac{1}{2}\,\frac{\langle z_0^*, D_{11}^2 F(0, y_0)[\hat{v}_0, \hat{v}_0] \rangle}{\langle z_0^*, D_{21}^2 F(0, y_0)\, \hat{v}_0 \rangle} \tag{7.25}$$

Hint: Differentiate $\hat{\Phi}(s, \hat{y}(s)) = 0$ with respect to s.

7.22 **Exercise** Apply the result of the preceding exercise to the problem (7.7) – (7.8) and show that

$$D\hat{y}(0) = -\sqrt{2}\,(k\pi)^2 f''(0) \int_0^1 (\sin k\pi x)^3\,dx$$

Even though the following considerations only require that X is continuously embedded in Z, we assume from here on that

$$X = Z \qquad Z_0 = \mathcal{N} \qquad X_0 = \mathcal{R} \qquad P = Q$$

So in (7.24) we can take $z_0 = \hat{v}_0$ and write

(7.26) $$Qx = \langle z_0^*, x \rangle\, \hat{v}_0$$

with i) $\langle z_0^*, \hat{v}_0 \rangle = 1$ and ii) $\langle z_0^*, x \rangle = 0 \quad \forall x \in \mathcal{R}$

The identity

$$D_1 F(0, y_0)\,\hat{v}_0 = 0$$

can be reformulated as: $D_1 F(0, y_0)$ has eigenvalue zero. The assumption

(7.27) $$D_{12}^2 F(0, y_0)\,\hat{v}_0 \notin \mathcal{R} \qquad \left(\Leftrightarrow \langle z_0^*, D_{12}^2 F(0, y_0)\,\hat{v}_0 \rangle \neq 0\right)$$

of the Crandall – Rabinowitz Theorem can be reformulated as: zero is a simple eigenvalue of $D_1 F(0, y_0)$. If, apart from this simple zero eigenvalue, the spectrum of $D_1 F(0, y_0)$ has strictly negative real part, we "feel" that the stability (as an equilibrium of the abstract ODE $\frac{dx}{dt} = F(x,y)$) of the solutions of $F(x,y)=0$ near $(0,y_0)$ is governed by the sign of the real eigenvalue that reduces to zero for $(x,y)=(0,y_0)$. In the present exposition we do neither make precise what we mean by this abstract ODE and by stability nor do we substantiate this feeling. All we shall do is compute how the signs of the real eigenvalue along the trivial branch $(0,y)$ and along the nontrivial branch $(\hat{x}(s), \hat{y}(s))$, as provided by the Crandall-Rabinowitz Theorem, are related to one another via the direction of bifurcation (7.25).

As a first step we establish that it is exactly the simplicity of the eigenvalue zero that allows us to "continue" it when (x,y) varies in a neighbourhood of $(0,y_0)$.

Define $F_e : X \times Y \times \mathcal{R} \times \mathbb{R} \to X$ by

$$(7.28) \qquad F_e(x,y,w,\lambda) = D_1F(x,y)(\hat{v}_0 + w) - \lambda(\hat{v}_0 + w)$$

Here the index "e" indicates "eigenvalue problem", since when

$$(7.29) \qquad F_e(x,y,w,\lambda) = 0$$

we do know that $D_1F(x,y)$ has eigenvalue λ with corresponding eigenvector $\hat{v}_0 + w$, $w \in \mathcal{R}$ and, conversely, if $D_1F(x,y)$ has eigenvalue λ then the corresponding eigenvector has a component in $\mathcal{N}$ that, when nonzero, we normalize to be $\hat{v}_0$. So note that we eliminated the non-uniqueness of the eigenvector by normalizing its W component to be exactly $\hat{v}_0$ (and not just a multiple of $\hat{v}_0$). By continuity the eigenvector cannot have a vanishing W component for (x,y) close to $(0,y_0)$ and λ close to zero.

We want to solve equation (7.29) for (w,λ) as a function of (x,y) in a neighbourhood of $(0,y_0)$. Now note that

$$D_3 F_e(0,y_0,0,0) = D_1 F(0,y_0)\big|_{\mathcal{R}}$$

$$D_4 F_e(0,y_0,0,0) = -\hat{v}_0$$

Combining the three facts

i) $D_1 F(0,y_0)\big|_{\mathcal{R}} : \mathcal{R} \to \mathcal{R}$ is bijective

ii) $W = \operatorname{span}[\hat{v}_0]$

iii) $X = W \oplus \mathcal{R}$

we see that the linear map $\mathcal{R} \times \mathbb{R} \to X$

$$(w,\lambda) \longmapsto D_1 F(0,y_0) w - \lambda \hat{v}_0$$

is bijective. So once again we can apply the IFT.

<u>Theorem</u> There exist C^1 functions

$$\tilde{w} : X \times \mathbb{R} \to \mathcal{R} \quad , \quad \tilde{\lambda} : X \times \mathbb{R} \to \mathbb{R}$$

such that

i) $\tilde{w}(0,y_0) = 0$, $\tilde{\lambda}(0,y_0) = 0$

ii) $F_e(x,y,\tilde{w}(x,y),\tilde{\lambda}(x,y)) = 0$, i.e.,

7.25

$\hat{\lambda}(x,y)$ is an eigenvalue of $D_1 F(x,y)$
with corresponding eigenvector $\hat{v}_0 + \tilde{w}(x,y)$

iii) $\tilde{F}(x, y, w, \lambda) = 0$ has no other solutions near $(0, y_0, 0, 0)$

We call $\hat{\lambda}$ the <u>critical eigenvalue</u>. For ease of notation we define

$$\lambda_t(y) = \hat{\lambda}(0,y)$$

(7.30)

$$\lambda_n(s) = \hat{\lambda}(\hat{x}(s), \hat{y}(s))$$

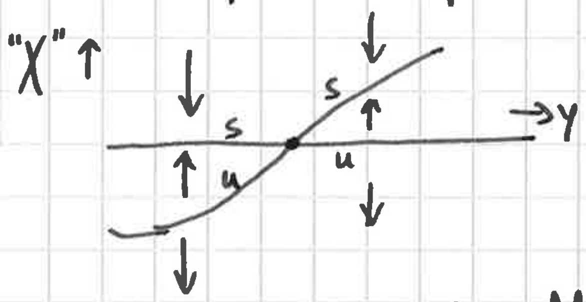
So $\lambda_t(y)$ is the critical eigenvalue along the <u>t</u>rivial (whence the subscript t) branch, while $\lambda_n(s)$ is the critical eigenvalue along the <u>n</u>ontrivial branch provided by the Crandall–Rabinowitz Theorem. We now compute

$$\dot{\lambda}_t(y_0) = \frac{d}{dy} \lambda_t(y) \Big|_{y=y_0}$$

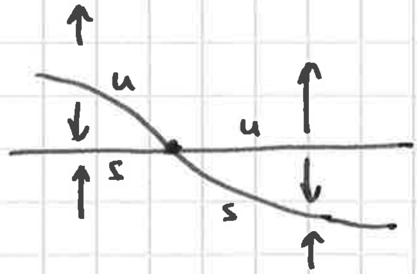
and

$$\dot{\lambda}_n(0) = \frac{d}{ds} \lambda_n(s) \Big|_{s=0}$$

and try to relate these two quantities to each other. The symbolic pictures



or



suggest heuristically that a relationship should exist.

7.26 Such pictures are described in words by

- supercritical branches are stable
- subcritical branches are unstable

and together such conclusions are designated as the

$$\boxed{\text{Principle of the Exchange of Stability}}$$

Exercise Show that

$$\dot{\lambda}_t(y_0) = \langle z_0^*, D_{21}^2 F(0,y_0)\,\hat{v}_0 \rangle$$

Hint: Differentiate $F_e(0, y, \tilde{w}(0,y), \tilde{\lambda}(0,y)) = 0$ with respect to y

Exercise Interpret (7.27) as: the real eigenvalue of $D_1 F(0,y)$ crosses the imaginary axis for $y = y_0$ with non-zero speed.

Exercise Show that

$$\dot{\lambda}_n(0) = \dot{\lambda}_t(y_0)\, D\hat{y}(0) + \langle z_0^*, D_{11}^2 F(0,y_0)[\hat{v}_0, \hat{v}_0] \rangle$$

Hint: Differentiate $F_e(\hat{x}(s), \hat{y}(s), \tilde{w}(\hat{x}(s),\hat{y}(s)), \tilde{\lambda}(\hat{x}(s),\hat{y}(s))) = 0$ with respect to s

Exercise Show that $\dot{\lambda}_n(0) = -\dot{\lambda}_t(y_0)\, D\hat{y}(0)$ (7.31)

Hint: use (7.25)

Exercise Relate (7.31) to the s and u labels in the pictures on the preceding page (NB the label s refers to "stability" and has nothing to do with the parameter/variable s that we use to describe the branch $(\hat{x}(s), \hat{y}(s))$)