Year Long Project

Bifurcation Theory

Matthew Holmes
supervised by
Miles Wheeler

A project submitted in partial fulfillment
of the requirements for the degree of
Master of Mathematics

at the
University of Bath
Year 2022

Contents

Chapter 1

Introduction

A large proportion of mathematical education is devoted to the solving of equations. From a young age we are introduced to expressions such as $2x-4$, and encouraged to find values we can substitute in the place of x such that the expression evaluates to zero, another number, or perhaps even other expressions. Further time is spent solving equations with multiple variables, using more sophisticated methods. The solution sets of linear equations that we can encode as $A\mathbf{x} = \mathbf{b}$ are shown to have underlying geometric structure using the tools of linear algebra.

This project concerns the case of equations of many variables, potentially uncountably many, for finite equations we can naturally encode the problem as that of finding the solution set of a function:

$$f : \mathbb{R}^n \to \mathbb{R}^m$$

Solving for vector values $\mathbf{x}$ such that $f(\mathbf{x}) = \mathbf{0}$. However we now no longer insist upon linearity, the solutions of such functions are not easy to find, often having no closed form or process for solving aside from numerical methods. Thankfully the solution sets generally exhibit a certain level of well-behavedness: the vast majority of the solutions will lie on a manifold of fixed dimension.

For example the solution set of the following single equation in three variables:

$$f(x, y, z) = x\cos(z) + y\sin(z)$$

Takes the form of 2-dimensional manifold with a corkscrew shape (turning around the black line where $x = y = 0$). We may view this equation as a function $f : \mathbb{R}^3 \to \mathbb{R}^1$, then we call the solution set the "zero set" of the function.

Unfortunately there may exist areas of the solution set in which this homogeneity is compromised, centred around "bifurcation points". Here the solution set need not be a manifold, but have a more exotic topology. Such points pose an opportunity to better understand the qualitative behaviour of the system that led to the equations in the first place.

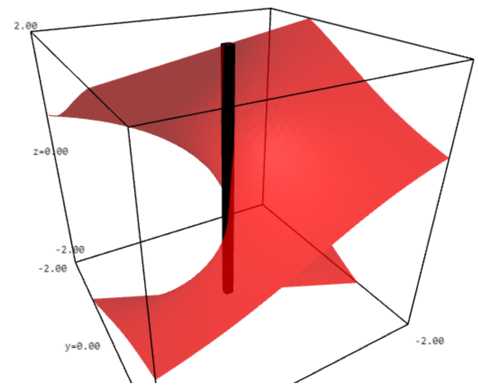


Figure 1.1: A “corkscrew” solution manifold

We call such points “bifurcations” since they usually arise at the intersection of two components of the solution set. Moving away from these bifurcations the solution set splits apart - or bifurcates, and we move back into portions of the solution set that look like manifolds.

If we consider a solution set that is generally a one-dimensional manifold, for example the red curves in 1.2, Then bifurcation points are easily spotted as any points of self intersection.

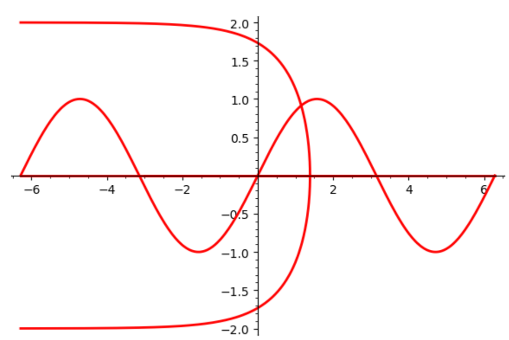


Figure 1.2: The self intersections of the red zero set are considered bifurcations (7 in total)

While finding such intersections may prove trivial for simple examples such as this, without a closed form for the solution set it proves much more challenging. By exploiting ideas from calculus and linear algebra we can develop sufficient conditions to identify a point as a bifurcation, without necessarily fully understanding the solution space as a whole.

In this project we explore methods for rigorously identifying bifurcations and extending them to the cases when the solutions lie in spaces of functions. We also implement the methods computationally, and explore how symmetry breaking can affect such bifurcations, leading to so called “imperfect bifurcations”

Chapter 2

Preliminaries

2.1 Derivatives and Taylor's Theorem

2.1.1 Frechet Derivatives

Let us consider functions of the form:

$$f : \mathbb{R}^n \to \mathbb{R}^m$$

Previous study of the calculus of such functions revolved around the concept of the partial derivative. First we treat f as not one, but m functions $f_i : \mathbb{R}^n \to \mathbb{R},\ \ i = 1, ..., m$ for each of which we may compute n partial derivatives - one for each input dimension.

$$\partial_{x_1} f_i, \dots, \partial_{x_n} f_i : \mathbb{R}^n \to \mathbb{R}$$

In this project we will treat these partial derivatives not as independent expressions but as one combined object, we consider the Frechet derivative of f at a:

$$\begin{aligned} Df_a &\in \mathcal{L}(R^n, \mathbb{R}^m) \\ Df_a &: R^n \to \mathbb{R}^m \end{aligned}$$

Such that

$$\lim_{\|h\| \to 0} \frac{\|f(a+h) - f(a) - Df_a(h)\|}{\|h\|} = 0 \tag{2.1}$$

Such Df_a is unique and when written as a matrix with respect to the standard basis, takes the form:

$$Df_a = \begin{pmatrix} \partial_{x_1} f_1 & \dots & \partial_{x_n} f_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m & \dots & \partial_{x_n} f_m \end{pmatrix} \tag{2.2}$$

This formulation allows for more concise arguments and opens up the possibility of applying results from linear algebra - many results in bifurcation theory depend on properties such as the kernel and image of Df_a

Notation

Such Frechet derivatives depend on the point of evaluation a, throughout this project we indicate this point as a subscript after the function in question.

Higher Order Derivatives

Higher order derivatives take the form of multilinear maps:

$$D^k f_a \in \mathcal{M}^k(\mathbb{R}^n, \mathbb{R}^m) \tag{2.3}$$

Sometimes is may be more helpful to consider them as nested linear maps instead:

$$D^k f_a \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \ldots, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \ldots)) \tag{2.4}$$

Both formulations can be shown to be equivalent.

2.1.2 Taylor's Theorem

Given a k-times differentiable function

$$f : \mathbb{R}^n \to \mathbb{R}^m$$

and point $a \in \mathbb{R}^n$, then we have that

$$f(x) = R_k(x, a) + \sum_{n=0}^{k} \frac{1}{n!} D^n f_a(x - a, ..., x - a) \tag{2.5}$$

Where $R_k(x, a)$ is an error term that we can bound above provided we assume that f is also $k + 1$ times differentiable. In which case:

$$\|R_k(x, a)\| \leq \frac{\|x - a\|\|^{n+1}}{(n + 1)!} \sup_{0 \leq t \leq 1} \|D^{n+1} f_{(1-t)a+tx}\| \tag{2.6}$$

Both 2.5 and 2.6 are proved in Analytic Theory of Global Bifurcation, Buffoni and Toland 2003, pages 31-32

We can interpret 2.6 to mean that as we increase k our Taylor approximations can arbitrarily well approximate f for values close to a. We can show that:

$$R_k(x, a) = o(\|x - a\|^k) \tag{2.7}$$

While not immediately a polynomial, we can easily extract the coefficients for a multi variable polynomial from a Taylor expansion of the form in 2.5. This will prove useful in later computational applications. However many theoretical results in bifurcation theory can be stated in terms of properties of the multilinear map notion of the derivative.

2.2 The Implicit Function Theorem

Given a k-times differentiable function

$$f : \mathbb{R}^{n+m} \to \mathbb{R}^m$$

For which (without loss of generality) $f(0) = 0$ Then viewing $\mathbb{R}^n$ as the product of two spaces:

$$\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$$

For which we denote $\mathbb{R}^n = X$ and $\mathbb{R}^m = Y$ so that:

$$f : X \times Y \to \mathbb{R}^m$$

Then the Implicit Function Theorem states that if the derivative with respect to Y is an isomorphism, then there exists open subsets of X and Y and a k-times differentiable function:

$$h : U_X \to U_Y \tag{2.8}$$

$$h(0) = 0 \tag{2.9}$$

That parametrises the zero set of f close to zero:

$$\{(x, y) \in U_X \times U_Y \mid \ f(x, y) = 0\} = \{(x, h(x)) \ \mid \ x \in U_X\} \tag{2.10}$$

In many ways the Implicit Function theorem is a natural extension to the Rank-Nullity theorem. Which relates the dimensionality of a linear operator's zero set, or kernel, to the dimension of the spaces it takes as domain and image, except now we permit non-linear functions.

2.2.1 Derivatives With Respect to Subspaces

We require $D_Y f_0$ is an isomorphism, which as a linear map means it must be invertible. Formally we can extract the derivative with respect to Y by considering the functions:

$$f(x_0, \cdot) : Y \to \mathbb{R}^m$$

From which $D_Y f_{(x_0,y_0)}$ is computed by fixing x_0 to get a function of only y and then computing the derivative at $y = y_0$.

If we choose the standard bases of linear maps $\mathbb{R}^{n+m} = \mathbb{R}^{m+n} \to \mathbb{R}^m$ then we can write the derivative of $f : \mathbb{R}^{m+n} \to \mathbb{R}^m$ as a $m \times (n + m)$ rectangular matrix:

$$Df_0 = \begin{pmatrix} \partial_{x_1} f_1 & \dots & \partial_{x_m} f_1 & \dots & \partial_{x_{n+m}} f_1 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \partial_{x_1} f_m & \dots & \partial_{x_m} f_m & \dots & \partial_{x_{n+m}} f_m \end{pmatrix} \tag{2.11}$$

From which we can focus just on the derivative in the $Y = \mathbb{R}^m$ part of the domain. Which corresponds to the left hand side $m \times m$ square part:

$$D_y f_0 = \begin{pmatrix} \partial_{x_1} f_1 & \dots & \partial_{x_m} f_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m & \dots & \partial_{x_m} f_m \end{pmatrix} \tag{2.12}$$

Should this matrix have full rank, or equivalently trivial kernel, then it will be invertible, and the Implicit Function Theorem holds. A matrix populated with random entries is very unlikely to be non-invertible, and so in general this requirement of invertibility is not very stringent.

2.2.2 Partial derivative notation

Many different conventions exist when writing partial derivatives. In this project we explicitly state the space in which the derivative is being computed; many books would write $D_x f$ or $\partial_x f$. We opt for

$$D_X f$$

Writing the space as a subscript after D, this yields a map:

$$D_X f : X \to \mathcal{L}(X, \mathbb{R}^m)$$

As before if we wish to specify the point at which the derivative is evaluated we use subscripts, for example:

$$D_X f_{x_0} \in \mathcal{L}(X, \mathbb{R}^m)$$

Which as a map we can evaluate to yield a result in $\mathbb{R}^m$

$$D_X f_{x_0}(x) \in \mathbb{R}^m$$

2.2.3 Manifolds

The parametrisation given in 2.10 shows us why close to zero, the solution set of f is a manifold. Consider:

$$\begin{aligned} p : U_X \subset R^n &\to \mathbb{R}^{n+m} \\ p(x) &= (x, h(x)) \end{aligned} \tag{2.13}$$

This is clearly continuous by the requirement that h from 2.8 was k-times differentiable. And invertible via projection onto $\mathbb{R}^n$, which is also continuous. Thus p is a homeomorphism and so a coordinate chart for the n-dimensional manifold:

$$M = \{(x, y) \in U_X \times U_Y \mid f(x, y) = 0\} = \{(x, h(x)) \mid x \in U_X\} \tag{2.14}$$

Further building the intuition that the Implicit Function Theorem preserves the result of the Rank-Nullity Theorem up to homeomorphism.

Historical Note

In fact the notion of a manifold as the zero set of a continuously differentiable function was how they were first introduced in the 1895 paper "Analysis Situs" of Poincare 1895, where a precursor to the general manifold was introduced as the level set of such functions.

Chapter 3

Bifurcations

In the previous chapter we saw how the implicit function theorem characterised the solution set of a function $f : \mathbb{R}^{n+m} \to \mathbb{R}^m$ locally as an n-dimensional manifold. Subject to the condition on the derivative. We now consider the failure case of the implicit function theorem that happens when this condition fails. Such points may turn out to be considered bifurcations, a term we will formalise now.

3.1 The Simple Bifurcation

A general treatment of all the possible setups permitted by the implicit function theorem would prove challenging. One immediate simplification is to fix the size of the dimension n, if we insist that $n = 1$ then we need only consider functions of the form:

$$f : \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m \tag{3.1}$$

In fact for practical applications this case is generally all that is required, in later chapters we will replace $\mathbb{R}^m$ with a general Banach space, so the theory can be applied to functions of functions.

We distinguish between the codomain and domain version of $\mathbb{R}^m$ by writing X or Y, some literature refers to the single dimensions as $\mathbb{F}$, which we will also adopt, writing:

$$f : \mathbb{F} \times X \to Y \tag{3.2}$$

3.1.1 A Standard Setup

We generally only consider functions with a "line of trivial solutions", functions for which:

$$\mathbb{F} \times \{0\} \tag{3.3}$$

is an a priori set of solutions. This 1-manifold of trivial solutions provide the "trunk" from which we look for branches of solutions splitting away from. This restriction places further weight on the the $\mathbb{F}$ dimension, as this now becomes the only degree of freedom when looking for a bifurcation. Now we can formalise what it means for a point in $\mathbb{F} \times \{0\}$ to be a bifurcation.

3.1.2 Definition - Bifurcation

A point $(\lambda_0, 0) \in \mathbb{F} \times \{0\}$ is a simple bifurcation if there is a sequence of solutions $\{(\lambda_n, x_n)\} \subset \mathbb{F} \times (X \setminus \{0\})$ converging to $(\lambda_0, 0)$ in $\mathbb{F} \times X$. Sometimes we may refer to just λ_0 as the bifurcation point.

This definition accounts for the fact we only care about local behaviour, the new strand of non-trivial solutions need only be accounted for very close to the trivial solutions.

Unless stated otherwise we use the terms "bifurcation" and "simple bifurcation" interchangeable, since the theory developed later only pertains to bifurcations of the form defined in 3.1.2.

3.2 A necessary condition

A cursory attempt at applying the implicit function theorem leads to a quick check that allows us to rule out points as bifurcations. The implicit function theorem is valid if f is differentiable and has invertible $D_X f_{(\lambda_0,0)}$. We generally only consider differentiable f so the latter condition is all that remains to check.

If $D_X f_{(\lambda_0,0)}$ is invertible then the implicit function theorem asserts that locally the solutions lie in a line parametrised by the $\mathbb{F}$ direction. Given that we have already assumed that $\mathbb{F} \times \{0\}$ are solutions passing through $(\lambda_0, 0)$ then these must be the solutions in question. If a new branch of solutions also existed then at the point $(\lambda_0, 0)$ the solution set would not be a 1-manifold.

So we conclude that non-invertibility of $D_X f_{(\lambda_0,0)}$ is a necessary condition for $(\lambda_0, 0)$ to be a bifurcation point. While not sufficient it formalises the idea that bifurcation theory deals with failure cases of the implicit function theorem.

Chapter 4

The Lyapunov-Schmidt Reduction

4.1 Motivation

The necessary condition established in 3.2 points us to investigate points with a level of degeneracy in the first derivative. Dimension reduction techniques allow us to salvage what remains well behaved, applying the implicit function theorem to leave only on the parts of the problem that are degenerate to be considered.

The Lyapunov-Schmidt reduction encapsulates this process formally. As seen in 2.12 for a point $(\lambda_0, 0)$ we can express the derivative with respect to X as a matrix of partial derivatives:

$$D_X f_{(\lambda_0,0)} = \begin{pmatrix} \partial_{x_1} f_1 & \dots & \partial_{x_m} f_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m & \dots & \partial_{x_m} f_m \end{pmatrix} \tag{4.1}$$

If $D_X f_{(\lambda_0,0)}$ is degenerate then we have that

$$\ker D_X f_{(\lambda_0,0)} \neq \{0\}$$

And,

$$\operatorname{Im} D_X f_{(\lambda_0,0)} \neq Y$$

The core idea of the Lyapunov-Schmidt reduction is to split the spaces X and Y according to the kernel and image of $D_X f_{(\lambda_0,0)}$ so that after choosing the right basis we can write $D_X f_{(\lambda_0,0)}$ as the block matrix:

$$\left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & A \end{array}\right) \tag{4.2}$$

Where A is invertible, and thus fit for application of the implicit function theorem on the spaces it takes as domain and codomain. Lyapunov-Schmidt gives us a way to relate this back to the original problem of finding solutions of f.

4.2 Statement of Theorem

Let $f : \mathbb{F} \times X \to Y$ be k-times differentiable, with a solution (λ_0, x_0) so that:

$$f(\lambda_0, x_0) = 0$$

Let L be the derivative of f at (λ_0, x_0) with respect to X:

$$L = D_X f_{(\lambda_0, x_0)}$$
$$L : X \to Y$$

If $\ker L \neq \{0\}$, and $q = \dim Y - \dim \operatorname{Im} L$, then there exists open sets U, V

$$(\lambda_0, x_0) \in U \subset X \tag{4.3}$$

$$(\lambda_0, 0) \in V \subset \mathbb{F} \times \ker L \tag{4.4}$$

such that the question of finding solutions of f in U is reduced to finding solutions of a new k-times differentiable function:

$$h : V \to \mathbb{R}^q \tag{4.5}$$

Where points $(\lambda, \xi) \in V$ that solve h, are mapped to solutions of f via:

$$\psi : V \to X \tag{4.6}$$

Which is also k-times differentiable, so that:

A point $(\lambda, x) \in U$ has $f(\lambda, x) = 0$ if and only if $\psi(\lambda, \xi) = x$ for a point $(\lambda, \xi) \in V$ with $h(\lambda, \xi) = 0$

4.2.1 Remarks

The content of this theorem is based in the X space. The $\mathbb{F}$ direction is present in both the full space $\mathbb{F} \times X$ and the reduced "kernel space" $V = \mathbb{F} \times \ker X$. Whatever λ comprises part of a solution to h is left unchanged when we map to the solution in $\mathbb{F} \times X$.

The power of this theorem is we reduce a system of $m + 1$ equations to just $q + 1$. In the case of the application to Crandall-Rabinowitz this q is taken to be $q = 1$ and so vastly reduces the complexity of the question.

The proof illuminates how h and ψ are constructed.

4.3 Proof of theorem

By the linearity of $L = D_X f_{(\lambda_0, x_0)}$ there exist subspaces $W \subset X$ and $Z \subset Y$ such that:

$$X = \ker L \oplus W \quad \text{and} \quad Y = Z \oplus \operatorname{Im} L \tag{4.7}$$

See that $\dim Z = \dim Y - \dim \operatorname{Im} L = q$.

With respect to bases derived from these direct sums, the matrix of L would have form:

$$\left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & A \end{array}\right) \tag{4.8}$$

Where A maps from the W subspace to $\operatorname{Im} L$. Formally we introduce a projection map

$$P : Y \to Y \tag{4.9}$$

$$\ker P = \operatorname{Im} L \tag{4.10}$$

Using block matrix notation as in 9.3.1 we would have that:

$$P = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right) \qquad (I - P) = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & I \end{array}\right) \tag{4.11}$$

We see that $(I - P)Lx = Lx$ for all $x \in X$. So by viewing $(I - P)$ as a map into $\operatorname{Im} L$ we see that $(I - P)L$ is a bijective linear map from W into $\operatorname{Im} L$. Moreover these projection maps allow us to extract the degenerate directions when solving f. But first we must detach the spaces $\ker L$ and W when talking about points in X.

See that for any point $x \in X$ we may express it as:

$$x = x_0 + \xi + w$$

Where $\xi \in \ker L$ and $w \in W$. Since bases give rise to unique coordinates we can define:

$$G(\lambda, \xi, w) = (I - P)f(\lambda, x_0 + \xi + w)$$

Which both separates the degenerate from well behaved dimensions and only considers parts of the evaluation of f that lie in the image of it's derivative.

See that $G(\lambda_0, 0, 0) = (I - P)f(\lambda_0, x_0) = 0$. Now consider taking the derivative with respect to the non-kernel subspace W at $(\lambda_0, 0, 0)$, finding

$$D_W G_{(\lambda_0, 0, 0)} \in \mathcal{L}(W, Y)$$

$$D_W G_{(\lambda_0,0,0)} = (I-P) D_W f_{(\lambda_0,x_0)} \tag{4.12}$$

But then observe that since we can view any $w \in W$ as a point $w \in X$, and W contains all non-kernel directions of L, we have:

$$D_W f_{(\lambda_0,x_0)}(w) = D_X f_{(\lambda_0,x_0)}(w) = Lw \tag{4.13}$$

Since if we recall how to construct the matrices that representation derivatives with respect to subspaces as per 2.12 we see that

$$D_W f_{(\lambda_0,x_0)}(w) = \begin{pmatrix} 0 \\ A \end{pmatrix} w = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & A \end{array}\right) \begin{pmatrix} 0 \\ w \end{pmatrix} = Lw \tag{4.14}$$

So combining 4.12 and 4.13 we see that:

$$D_W G_{(\lambda_0,0,0)}(w) = ((I-P)L)(w) \tag{4.15}$$

And thus we see that $D_W G_{(\lambda_0,0,0)}$ extracts the A quadrant from 9.3.1, forming a bijection from W onto $\operatorname{Im} L$. Observing that $\operatorname{Im} L$ is also the codomain of G due to the premultiplication by $(I-P)$ we can apply the Implicit Function theorem to see that locally $\mathbb{F} \times \ker L$ parametrises the zero set of G, that is:

There exists open sets U, V,

$$(\lambda_0, x_0) \in U \subset X$$

$$(\lambda_0, 0) \in V \subset \mathbb{F} \times \ker L$$

And a k-times differentiable mapping:

$$\phi : V \to W$$

Such that $G(\lambda, \xi, \phi(\lambda, \xi)) = 0$ for all $(\lambda, \xi) \in V$, in fact the theorem states that the whole portion of the zero set in U is accounted for by ϕ, recalling the definition of G this means:

$$\begin{aligned} &\{(\lambda, x_0 + \xi + w) \in U \mid (I-P) f(\lambda, x_0 + \xi + w) = 0\} \\ &= \{(\lambda, x_0 + \xi + w) \mid (\lambda, \xi) \in V \text{ and } w = \phi(\lambda, \xi)\} \end{aligned}$$

We now can define the reduced dimensional problem h and its mapping to solutions of f, let:

$$\psi : V \to X$$

$$\psi(\lambda, \xi) = x_0 + \xi + \phi(\lambda, \xi)$$

Which "almost solves" f, with $(I-P) f(\lambda, x_0 + \xi + \phi(\lambda, \xi)) = 0$ for any choice of input, forcing the component we must pick in the well behaved derivative directions W. So all that remains is to clean up any non-zeros that remain in Z, with the only degree of freedom being the choices of $(\lambda, \xi) \in V$ we feed into ψ, accordingly set

$$
\begin{aligned}
h &: V \to Z \\
h(\lambda, \xi) &= Pf(\lambda, \psi(\lambda, \xi)) \in Z
\end{aligned} \tag{4.16}
$$

Choosing a point which generates a solution to $(I - P)f$ only if it that solution also solves Pf, and so combining we see that:

$$
\begin{aligned}
h(\lambda, \xi) = 0 &\text{ if and only if } Pf(\lambda, \psi(\lambda, \xi)) = 0 \\
&\text{ if and only if } Pf(\lambda, x_0 + \xi + \phi(\lambda, \xi)) = 0 \\
&\text{ if and only if } F(\lambda, x_0 + \xi + \phi(\lambda, \xi)) = 0
\end{aligned}
$$

Since $(\lambda, x_0 + \xi + \phi(\lambda, \xi))$ is also a solution of $(I - P)f$

Choosing a basis for the q-dimensional space Z which is the codomain of h we identify Z with $\mathbb{R}^q$ and are done

$\square$

4.3.1 Remarks on proof

The key idea in this proof is trying to reduce our freedom to choose solution points. Initially we may consider moving anywhere in $\mathbb{F} \times X$ on the hunt for solutions. But the application of the Implicit Function Theorem shows that we need not consider the well behaved W part. The choice of degenerate part of the solution pulls the strings on the part in W, via ϕ, since any solution to f must also solve $(I - P)f$, of which $V \subset \mathbb{F} \times \ker L$ parametrises the solutions.

Hence we only need consider searching V for "seeds" to solutions of Pf, this becomes our reduced problem.

This proof provides an algorithm to obtain both h and ψ, indicating a potential computational implementation.

Chapter 5

The Crandall-Rabinowitz Theorem

Now equipped with the Implicit Function Theorem and Lyapunov-Schmidt reduction we can state and prove a sufficient condition for a point to be a bifurcation. Like previous chapters we denote $X = Y = \mathbb{R}^m$, $\mathbb{F} = \mathbb{R}$

5.1 Statement of Theorem

If we have:

$$f : \mathbb{F} \times X \to Y$$

a k-times differentiable function with $k \geq 2$ with trivial solutions $f(\lambda, 0) = 0 \in Y$ for all $\lambda \in \mathbb{F}$, where $\dim X = \dim Y$.

Then at a point $(\lambda_0, 0)$ if we have that

$$L = D_X f_{(\lambda_0,0)}$$

Has one dimensional kernel

$$\ker L = \langle \xi_0 \rangle = \{\xi \in X \mid \xi = s\xi_0 \text{ for some } s \in \mathbb{F}\}$$

$$\xi_0 \in X \setminus \{0\}$$

And the following transversality condition holds:

$$D^2_{\mathbb{F},X} f_{(\lambda_0,0)}(1, \xi_0) \notin \operatorname{Im} L \tag{5.1}$$

Then this is sufficient for $(\lambda_0, 0)$ to be a bifurcation point.

Precisely, for an open neighbourhood U

$$(\lambda_0, 0) \in U \subset \mathbb{F} \times X$$

and an $\epsilon > 0$ we can parametrise the branch of non-trivial solutions

$$\{(\lambda, x) \in U \mid f(\lambda, x) = 0, x \neq 0\} = \{(\Lambda(s), \chi(s) \mid 0 < |s| < \epsilon\} \tag{5.2}$$

With Λ and χ both $(k-1)$-times differentiable on $(-\epsilon, \epsilon)$

5.1.1 Concerning transversality

The most novel requirement of this theorem is the so called transversality condition, pertaining to the bilinear map $D^2_{\mathbb{F},X} f_{(\lambda_0,0)}$.

Using symmetry (viewing $D^2_{\mathbb{F},X} f_{(\lambda_0,0)}$ as a tensor) we can we see that:

$$D^2_{\mathbb{F},X} f_{(\lambda_0,0)}(1,\xi_0) = D^2_{X,\mathbb{F}} f_{(\lambda_0,0)}(\xi_0,1) \tag{5.3}$$

and then considering the nested linear map interpretation of higher order derivatives, we see that

$$D^2_{X,\mathbb{F}} f_{(\lambda_0,0)}(\xi_0,1) = D_{\mathbb{F}}[D_X f_{(\lambda,0)}(\xi_0)]_{\lambda_0}(1) \tag{5.4}$$

Where the expression inside the square brackets is a function of $\lambda \in \mathbb{F}$ and thus differentiable with respect to $\mathbb{F}$

Since $\mathbb{F}$ is the one-dimensional $\mathbb{R}$ this means we can express the expression in 5.1 as a standard derivative, combining 5.3 and 5.4 we see that:

$$D^2_{\mathbb{F},X} f_{(\lambda_0,0)}(1,\xi_0) = \lim_{t\to 0} \frac{D_X f_{\lambda_0+t,0}(\xi_0) - D_X f_{(\lambda_0,0)}(\xi_0)}{t} \in Y \tag{5.5}$$

The transversality condition 5.1 asserts that this value cannot wholly lie in $\operatorname{Im} L$.

The transversality condition serves two purposes, the first is ruling out the possibility that the solution is a 2-dimensional manifold parametrised by an open subset of $\langle \xi_0 \rangle \times \mathbb{F}$, in which case we would have $D^2_{\mathbb{F},X} f_{(\lambda_0,0)}(1,\xi_0) = 0$

The second is in that it raises possibility of solution parametrised by movement in $\langle \xi_0 \rangle$, since requires that the degeneracy of derivatives f is limited to just the first order. In fact this is enough to implement the Implicit Function Theorem on an altered problem to yield a new branch of solutions. The presence of a component outside the image of L is essentially in setting this up.

5.2 Proof of Theorem

We begin the proof by applying the Lyapunov-Schmidt reduction to the equation $f(\lambda, x) = 0$ around the point $(\lambda_0, 0)$. Which locally reduces the problem to solving a new function h. We see that all solutions in U are determined by solutions in V of h where:

$$(\lambda_0, 0) \in U \subset X$$

$$(\lambda_0, 0) \in V \subset \mathbb{F} \times \ker L = \mathbb{F} \times \langle \xi_0 \rangle$$

$$h : V \to \mathbb{R}^q$$

$$\psi : V \to X$$

A solution $(\lambda, \xi) \in V$ of h maps to a solution $(\lambda, \psi(\lambda, \xi)) \in \mathbb{F} \times X$ of f, and this is an if and only if relationship.

We see that $q = 1$, since $q = \dim Y - \dim \operatorname{Im} L = \dim X - \dim \operatorname{Im} L = \dim \ker L = 1$ via assumption and rank nullity.

We will also require the function ϕ from the proof, that resulted from applying the Implicit Function Theorem to $(I-P)f$, and forces our choice of non-kernel component once we fix a point in $V \in \mathbb{F} \times \ker L = \mathbb{F} \times \langle \xi_0 \rangle$. Recall $(I - P)$ is projection onto $\operatorname{Im} L$. If we set W so that:

$$X = \langle \xi_0 \rangle \oplus W$$

Then

$$\phi : \mathbb{F} \times \langle \xi_0 \rangle \to W$$

ϕ is such that for all $\xi \in \ker L$ we have that $f(\lambda, \xi + \phi(\lambda, \xi))$ evaluates to zero in the parts of Y accessible by the Implicit Function Theorem, namely $\operatorname{Im} L$. And ϕ accounts for the whole portion of the zero set of $(I - P)f$ in U.

We now will show that this zero set is non-trivial and has a branch splitting away from $(\lambda_0, 0)$, constructed from a branching that occurs in $\mathbb{F} \times \langle \xi_0 \rangle$. First we establish some results:

We know that $f(\lambda, 0) = 0$ for all $\lambda \in \mathbb{F}$. By orthogonality the only way to achieve $\xi + w = 0$ for $\xi \in \ker L$ and $w \in W$ is if $\xi = w = 0$, and thus the $(\lambda, 0)$ solution written as $(\lambda, \xi + \phi(\lambda, \xi))$ enforces that both ξ and $\phi(\lambda, \xi)$ are zero. We conclude that $\phi(\lambda, 0) = 0$ for all $(\lambda, 0) \in U$.

Recall the definition of h from 8.1 ($x_0 = 0$)

$$h(\lambda, \xi) = Pf(\lambda, \xi + \phi(\lambda, \xi))$$

We see that:

$$h(\lambda, 0) = Pf(\lambda, 0 + 0 + \phi(\lambda, 0)) = Pf(\lambda, 0) = 0 \tag{5.6}$$

For all $(\lambda, 0) \in V$. Furthermore, it is shown in A.1 and A.3 that:

$$D_{\langle \xi_0 \rangle} h_{(\lambda_0, 0)} = 0 \tag{5.7}$$

and,

$$D^2_{\mathbb{F}, \langle \xi_0 \rangle} h_{\lambda_0, 0}(1, \xi_0) \neq 0 \tag{5.8}$$

This last condition depends on the Transversality condition holding. The goal now is concoct a new solution to h, parametrised by $\langle \xi_0 \rangle$, this will be our branch of non-trivial solutions. As usual we aim to use the Implicit function theorem, unfortunately 5.6 thwarts the simple application, since a line of solutions exists already. So we employ a trick, observing that A.1 and A.3 ensure we could get a local solution to:

$$\tilde{g} : \mathbb{F} \times \langle \xi_0 \rangle \to Y$$

$$\tilde{g}(\lambda, \xi) = D_{\langle\xi_0\rangle} h_{(\lambda,0)}(\xi)$$

Parametrising the $\mathbb{F}$ component by the $\langle\xi_0\rangle$ dimension. The problem is that solutions to $\tilde{g}$ need not solve h, but the definition of the derivative points us to a simpler function whose solutions will also solve h:

$$g(\lambda, s\xi_0) = \begin{cases} \frac{h(\lambda, s\xi_0)}{s} & \text{if } s \neq 0 \\ D_{\langle\xi_0\rangle} h_{(\lambda,0)}(\xi_0) & \text{if } s = 0 \end{cases}$$

That this functions satisfies the constraints on derivatives for the Implicit Function theorem remains to be shown, in fact we don't even know it is continuous yet. Luckily we see that it has a different formulation, arising from considering the fact (recalling $h(\lambda, 0) = 0$):

$$h(\lambda, \xi) = \int_0^1 \frac{d}{dt} h(\lambda, t\xi)\, dt$$

Since we can write all $\xi \in \langle\xi_0\rangle$ in the form $\xi = s\xi_0$ for some $s \in \mathbb{R}$, we see that (for $s \neq 0$):

$$\begin{aligned} \frac{d}{dt} h(\lambda, t\xi) &= \lim_{i\to 0} \frac{h(\lambda, (t+i)\xi) - h(\lambda, t\xi)}{i} \\ &= \lim_{i\to 0} \frac{h(\lambda, ts\xi_0 + is\xi_0) - h(\lambda, ts\xi_0)}{i} \\ &= \lim_{is\to 0} \frac{h(\lambda, (ts)\xi_0 + (is)\xi_0) - h(\lambda, ts\xi_0)}{iss^{-1}} \\ &= s\frac{d}{d\xi_0} h(\lambda, ts\xi_0) \\ &= s[D_{\langle\xi_0\rangle} h(\lambda, t\xi)](\xi_0) \end{aligned}$$

Thus we see that for $s \neq 0$:

$$g(\lambda, \xi) = \frac{h(\lambda, s\xi_0)}{s} = \frac{1}{s}\int_0^1 \frac{d}{dt} h(\lambda, t\xi)\, dt = \int_0^1 [D_{\langle\xi_0\rangle} h(\lambda, t\xi)](\xi_0)\, dt$$

Which also agrees with the definition at $s = 0$:

$$\int_0^1 [D_{\langle\xi_0\rangle} h(\lambda, t\cdot 0)](\xi_0)\, dt = [D_{\langle\xi_0\rangle} h(\lambda, 0)](\xi_0)$$

So we conclude:

$$g(\lambda, s\xi_0) = \int_0^1 [D_{\langle\xi_0\rangle} h(\lambda, t\xi)](\xi_0)\, dt = \begin{cases} \frac{h(\lambda, s\xi_0)}{s} & \text{if } s \neq 0 \\ D_{\langle\xi_0\rangle} h_{(\lambda,0)}(\xi_0) & \text{if } s = 0 \end{cases} \tag{5.9}$$

A $(k-1)$-times differentiable function since h is k-times differentiable.

With this in hand we now check that g satisfies the conditions of the implicit function theorem. First observing that by A.1

$$g(\lambda_0, 0) = D_{\langle\xi_0\rangle} h_{(\lambda_0,0)}(\xi_0) = 0$$

Thus $(\lambda_0, 0)$ is a solution of g. Now we check that the $\mathbb{F}$ derivative is invertible at $(\lambda_0, 0)$, which as a linear map with one dimensional domain is true if $D_{\mathbb{F}}g_{(\lambda_0,0)}(1) \neq 0$, evaluating:

$$D_{\mathbb{F}}g_{(\lambda_0,0)}(1) = D_{\mathbb{F}}[D_{\langle\xi_0\rangle}h_{(\lambda,0)}(\xi_0)]_{\lambda_0}(1) = D^2_{\mathbb{F},\langle\xi_0\rangle}h_{(\lambda_0,0)}(1, \xi_0) \neq 0$$

By A.3, as required.

Thus we may apply the Implicit Function Theorem to get a $(k-1)$ times differentiable mapping:

$$\tilde{\Lambda} : \{s\xi_0 \in \langle\xi_0\rangle : |s| < \epsilon\} \to \mathbb{F}$$

$$g\left(\tilde{\Lambda}(s\xi_0), s\xi_0\right) = 0 \text{ if } |s| < \epsilon$$

$$\tilde{\Lambda}(0) = \tilde{\lambda}_0$$

From 5.9 we see that for $s \neq 0$ any solution to g is also a solution to h, so combined with $h(\lambda_0, 0) = 0$ we have that:

$$h(\tilde{\Lambda}(s\xi_0), s\xi_0) = 0 \text{ if } |s| < \epsilon$$

Thus we have parametrised a new solution of h by $\langle\xi_0\rangle$. The Implicit Function Theorem ensures that these are the only solutions of g in the induced open set, and thus by its definition for $s \neq 0$ these are the only solutions of h when $s \neq 0$ in the induced open set.

We can use ψ from the Lyapunov-Schmidt decomposition to map these solutions to solutions of f. ψ depends on both the λ and ξ variables, so to highlight that ultimately it is only the single dimension $\langle\xi_0\rangle$ that parametrises the solution, we define:

$$\chi(s) = \psi(\tilde{\Lambda}(s\xi_0), s\xi_0)$$

Since ψ and $\tilde{\Lambda}$ are both at least $(k-1)$ times differentiable, so is χ by the chain rule.

Finally, setting

$$\Lambda(s) = \tilde{\Lambda}(s\xi_0)$$

We arrive at:

$$f(\Lambda(s), \chi(s)) = 0 \text{ for } 0 < |s| < \epsilon$$

As $\langle\xi_0\rangle$ locally parametrises all the non-trivial solutions, the conditions $x \neq 0$ and $s \neq 0$ are equivalent, thus:

$$\{(\lambda, x) \in U \mid f(\lambda, x) = 0, x \neq 0\} = \{(\Lambda(s), \chi(s) \mid 0 < |s| < \epsilon\}$$

As required

□

Chapter 6

Bifurcation Theory in Function Spaces

With some adjustments we can can extend Lyapunov-Schmidt decompositions and the Crandall-Rabinowitz theorem to equations with unknowns that lie in spaces of functions. In this chapter we will discuss how to extend the theory of bifurcations to spaces of functions, beginning with some definitions.

We now consider X and Y to be arbitrary Banach spaces, over the field $\mathbb{R}$.

Definition (linear operator)

a *linear operator* T between Banach spaces X, Y is a function $T : X \to Y$ such that:

$$T(\lambda x + \mu y) = \lambda Tx + \mu Ty \quad \text{for all} \lambda, \mu \in \mathbb{R}, x, y \in X \tag{6.1}$$

6.0.1 Lemma

Given a finite-dimensional subspace X_1 of a Banach space X there exists a subspace X_2 of X such that $X = X_1 \oplus X_2$

Definition (codimension)

The dimension of the space X_1 in lemma 6.0.1 is called the *codimension* of X_2

Definition (Fredholm operator)

An linear operator $T : X \to Y$ is called *Fredholm* if:
$\ker(T)$ has finite dimension n
$\mathrm{Im}(T)$ is closed and has finite codimension r
We then say T has *Fredholm index* $p = n - r$

6.1 Extending Theorems to Spaces of Functions

The three main theorems stated so far: The implicit function theorem, Lyapunov-Schmidt theorem and Crandall-Rabinowitz theorem can all be extended to the general case where X and Y are Banach spaces. For use in the this project they will assumed without proof. However the adjustments required are minimal, and will be outlined now.

6.1.1 Implicit Function Theorem - Outline

If we now let X, Y, Z be arbitrary Banach spaces and consider a k-times differentiable function

$$f : X \times Y \to Z \tag{6.2}$$

Such that (without loss of generality) $f(0,0) = 0$. Then if $D_X f_{(0,0)}$ is a homeomorphism there exists a k-times differentiable function between subsets of X and Y that parametrises the zero set close to zero.

Full statement and proof is found in Buffoni and Toland 2003, pages 37-38

We see that this theorem remains largely the same, the main difference being that $D_X f$ is a derivative with respect to a potentially infinite dimensional space. This will be investigated further next section.

6.1.2 Lyapunov-Schmidt Theorem - Outline

Let $f : \mathbb{F} \times X \to Y$ be a k-times differentiable function, and X, Y Banach spaces with a known solution $f(\lambda_0, x_x) = 0$.

If $L = D_X f_{(\lambda_0, x_x)} : X \to Y$ is a Fredholm operator, with non-trivial kernel: $\ker L \neq \{0\}$. Then we can reduce the question of finding solutions of f in an open set around (λ_0, x_0) to that of finding solutions of a new k-times differentiable function:

$$h : V \to \mathbb{R}^q$$

Where $V \subset \mathbb{F} \times \ker L$ is open too and q is the codimension of $\mathrm{Im}\, L$.

Full statement and proof is found in Buffoni and Toland 2003, pages 104-105.

See that where before we defined $q = \dim Y - \dim \operatorname{Im} L$, we now require the notion of codimension, since Y may be infinite dimensional. The key observation being that under this finite codimensionality assumption we can yield a new finite dimensional problem. The proof is very similar, using the general implicit function theorem. The notion of "infinite dimensional" block matrices is more abstract, but since the non-zero block is still finite dimensional this idea remains helpful.

6.1.3 Crandall-Rabinowitz Theorem - Outline

Given $f : \mathbb{F} \times X \to Y$ k-times differentiable, $K \geq 2$, with trivial solutions $f(\lambda, 0) = 0$.

Then if $L = D_X f_{(\lambda_0,0)}$ is a Fredholm operator of index zero, with one dimensional kernel:

$$\ker L = \langle \xi_0 \rangle = \{\xi \in X \mid \xi = s\xi_0 \text{ for some } s \in \mathbb{F}\}$$

And transversality holds:

$$D^2_{\mathbb{F},X} f_{(\lambda_0,0)}(1, \xi_0) \notin \operatorname{Im} L \tag{6.3}$$

Then $(\lambda_0, 0)$ is a bifurcation point.

Full statement and proof is found in Buffoni and Toland 2003 pages 106-107.

We see that the general statement requires the notion of Fredholm index, since $\dim X = \dim Y$ is not a well posed question when such dimensions are potentially infinite. The restriction that L has Fredholm index zero is a sufficient analogue.

The proof for the general case is much the same as before, we still can apply a Lyapunov-Schmidt reduction to begin with, with the constraint on L's Fredholm index ensuring that we can reduce the space we investigate to the two dimensional $V \subset \mathbb{F} \times \ker L$.

Chapter 7

Example - bending an elastic rod

7.1 The Problem

The bending of an elastic rod is a classic example in Bifurcation Theory, thought of a problem which has solutions in the space of continuous functions.

As per Buffoni and Toland 2003 pages 2-3 we consider an elastic rod of length $L > 0$, fixed at the left side and the other end free to move in only the x axis. We assume that the rod is incompressible, and thus has constant length, but may bend.

We then consider applying a lateral force from the right hand side, attempting to induce bending of the rod.

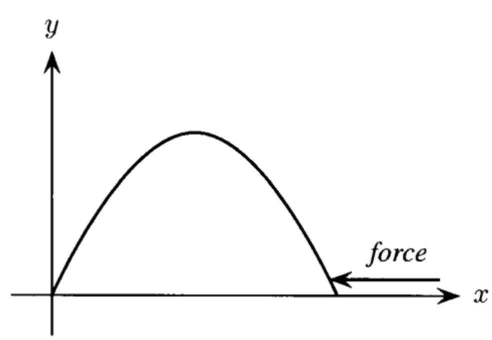


Figure 7.1: The rods bends under the action of the force. Buffoni and Toland 2003, page 2

We can describe the rods shape as being given by the angles $\phi(s)$ given by the tangent to the rod and the horizontal. Where $s \in [0, L]$ is a measure of distance along the rod.

We see that the corresponding $(x(s), y(s))$ Cartesian coordinates can be obtained via:

$$x(s) = \int_0^s \cos\phi(t)\, dt \;\text{ and }\; y(s) = \int_0^s \sin\phi(t)\, dt \tag{7.1}$$

Denoting the force applied by λ it can be show that:

$$\phi'(s) = \lambda y(s) \tag{7.2}$$

Here the right hands side is the moment of the applied force.

Thus by observing that $y'(s) = \sin\phi(s)$ and $y(0) = y(L) = 0$, we arrive at the following boundary-value problem:

$$\phi''(x) + \lambda \sin(\phi(x)) = 0 \;\text{ for }\; x \in [0, L] \tag{7.3}$$

$$\phi'(0) = \phi'(L) = 0 \tag{7.4}$$

7.4 are obtained since at the endpoints we see that $y(0) = y(L) = 0$, and $\phi'(s) = \lambda y(s)$. To translate into the language of the Crandall-Rabinowitz theorem we denote:

$$\mathbb{F} = \mathbb{R}$$

$$X = \{\phi \in C^2[0, L] \mid \phi'(0) = \phi'(L) = 0\}$$
$$Y = C[0, L]$$

Then we see that it becomes a problem of finding solutions of

$$F : \mathbb{R} \times X \to Y$$

Where $F(\lambda, \phi) = \phi'' + \lambda \sin(\phi)$

We see that the zero function is indeed a solution for any choice of $\lambda > 0$, and thus establish the "line" of trivial solutions.

7.2 Bifurcation Theory in $\mathbf{C}^k$

We will now apply the general case Crandall-Rabinowitz theorem as discussed in 6.1.3, first we confirm that

we need to first check the necessary condition:

$$D_X F_{(\lambda,0)} = 0 \tag{7.5}$$

But now we must consider how to compute derivatives with respect to the space:

$$X = \{\phi \in C^2[0, L] \mid \phi'(0) = \phi'(L) = 0\}$$

7.2.1 Linearising the problem

We seek a linear map of functions:

$$D_X F_{(\lambda,0)} \in \mathcal{L}(X,Y)$$

That is the best linear approximation of F. we see that the operator $\phi(x) \to \phi''(x)$ is already linear, thus all remains to linearise is the:

$$\phi(x) \to \sin(\phi(x))$$

part, since λ is treated as a fixed constant when taking the derivative with respect to X. This part has best linear approximation:

$$\phi(x) \to \phi(x)$$

which is verified by considering the Taylor expansion of $\sin x$ at $x = 0$, since we are computing the derivative at the zero function, so always computing the Taylor expansion at zero.

Thus we arrive at:

$$D_X F_{(\lambda,0)}(\phi) = \phi'' + \lambda\phi \ \in Y \tag{7.6}$$

7.2.2 The Fredholm index

A requirement of the general case of Crandall-Rabinowitz 6.1.3 is that $D_X F_{(\lambda,0)}$ has Fredholm index zero. This is a highly non-trivial assertion. Fortunately for this example 7.6 takes the form of an elliptic operator, the Fredholm indices of which are discussed at length in Kielhöfer 2012 section III.1 The Fredholm Property of Elliptic Operators. For the purpose of this example we will assume without proof that $D_X F_{(\lambda,0)}$ has Fredholm index zero.

7.2.3 The necessary condition

Considering 7.6 we see that degeneracy in the first derivative occurs only when the function $\phi(x) \in X$ solves the boundary-value problem:

$$\phi''(x) + \lambda\phi(x) = 0 \ \text{ for } \ x \in [0, L] \tag{7.7}$$

$$\phi'(0) = \phi'(L) = 0 \tag{7.8}$$

Where requirement 7.8 is necessary by virtue of $\phi \in X = \{\phi \in C^2[0, L] \mid \phi'(0) = \phi'(L) = 0\}$

It is a known result that this problem has non-trivial solutions (λ, ϕ) if and only if

$$\lambda \in \left\{ \lambda_k = \left(\frac{k\pi}{L}\right)^2 \mid k \in \mathbb{N} \right\}$$

In which case we have that the 1-dimensional subspace of X spanned by:

$$\phi_k(x) = \cos\left(\frac{k\pi x}{L}\right)$$

is the degenerate direction, i.e. such that:

$$L(c\phi_k) = D_X F_{(\lambda,0)}(c\phi_k) = 0$$

Thus $\ker L = \langle \phi_k \rangle$ and so the necessary condition is met.

7.2.4 The transversality condition

We now aim to verify the transversality condition:

$$D^2_{\mathbb{F},X} F_{(\lambda_k,0)}(1, \phi_k) \notin \mathrm{Im}(L)$$

First we evaluate $D_X F_{(\lambda,0)}$, where we need to take the derivative with respect to $\mathbb{F}$ of $D_X F_{(\lambda,0)} : \phi \to \phi'' + \lambda\phi$. Recall that:

$$D^2_{\mathbb{F},X} F_{(\lambda_k,0)}(1, \phi_k) = \lim_{t \to 0} \frac{\phi_k'' + (\lambda_k + t)\phi_k - \phi_k'' - (\lambda_k)\phi_k}{t} = \phi_k$$

Now we want to find $\mathrm{Im}\, L$, see that any $v \in \mathrm{Im}\, L \subset Y = C([0, L])$ is such that:

$$v = u'' + \lambda_k u \tag{7.9}$$

For some $u \in X$. Now consider:

$$\begin{aligned}
\int_0^L v(s)\phi_k(s)ds &= \int_0^L \left[u''(s) + \lambda_k u(s)\right] \phi_k(s)ds \\
&= \int_0^L u''(s)\phi_k(s)ds + \int_0^L \lambda_k u(s)\phi_k(s)ds \\
&= \left[u'\phi_k\right]_0^L - \int_0^L u'(s)\phi_k'(s)ds + \int_0^L \lambda_k u(s)\phi_k ds \\
&= 0 \cdot \phi_k(L) - 0 \cdot \phi_k(0) - \left[u\phi_k'\right]_0^L + \int_0^L u(s)\phi_k''(s)ds + \int_0^L \lambda_k u(s)\phi_k ds \\
&= -u(0) \cdot 0 + u(0) \cdot 0 + \int_0^L u(s)\phi_k''(s) + \lambda_k u(s)ds \\
&= \int_0^L \left(\phi_k''(s) + \lambda_k \phi_k(s)\right) u(s)ds \\
&= \int_0^L 0 \cdot u(s)ds \\
&= 0
\end{aligned}$$

Where we apply integration by parts twice, then apply the boundary property $\phi'(0) = \phi'(L) = 0$ and the first derivative of $\phi_k'(s) = -\frac{k\pi}{L}\sin\left(\frac{k\pi s}{L}\right)$ which also evaluates to 0 at $s = 0$ and $s = L$

Thus:

$$\int_0^L v(s)\phi_k(s)ds = 0 \tag{7.10}$$

is a necessary condition for any $v \in \operatorname{Im} L$. But for any λ_k potential bifurcation we have that $D^2_{\mathbb{F},X}F_{(\lambda_k,0)}(1,\phi_k) = \phi_k$, evaluating for condition 7.10:

$$\int_0^L v(s)\phi_k(s)ds = \int_0^L \phi_k^2(s) > 0 \tag{7.11}$$

Thus we have that $D^2_{\mathbb{F},X}F_{(\lambda_k,0)}(1,\phi_k) = \phi_k \notin \operatorname{Im} L$, and thus at all points λ_k we have a bifurcation of solutions. Each corresponding to a higher frequency bending mode.

Chapter 8

A Computational Implementation of the Implicit Function Theorem

The Implicit Function Theorem is an ever present component in the previous results. We now return to finite dimensional problems and implement an approximate Implicit Function Theorem in the computer algebra software SageMath (The Sage Developers 2021). This will afford the flexibility to tackle larger problems where "by hand" computations are not feasible. While approximations, all expressions will remain in closed form, without using floating point arithmetic.

8.1 The Problem

The Implicit Function Theorem states that, if we denote $\mathbb{R}^n = X$ and $\mathbb{R}^m = Y$, functions of the form

$$f : X \times Y \to \mathbb{R}^m$$

Have solution sets parametrised locally by the X dimensions, provided the derivative with respect to Y is an isomorphism.

We only consider applying the theorem at zero, without loss of generality (since we can shift the function), in which case the Implicit Function Theorem ensures the existence of open subsets of X and Y and a k-times differentiable function that parametrises the zero set of f close to zero:

$$h : U_X \to U_Y \tag{8.1}$$

$$h(0) = 0 \tag{8.2}$$

If we can find this h then we have found the solutions to f, locally at least. In practice, a closed form expression of this h is too challenging a problem, however we may approximate it. The problem we attempt to solve is to produce an arbitrarily accurate Taylor Polynomial approximation of h.

The goal is to produce a program that for a given order k, can produce a closed form expression for the k-th Taylor Polynomial at zero:

$$h_k = h(0) + Dh_0(x) + \frac{D^2 h_0(x^2)}{2!} + \cdots + \frac{D^k h_0(x^k)}{k!} = \sum_{i=0}^{k} \frac{h_0^{(i)}(x^i)}{i!}$$

8.2 Building h From f

This implementation depends on one key fact, which is easily seen when we consider that:

$$f(x, h(x)) = 0 \;\; \forall\, x \in U_X$$

Thus if we now view f as a function:

$$f(x, h(x)) : X \to \mathbb{R}^m$$

Locally we have that $f \equiv 0$, thus:

$$D_X^k f(x, h(x)) = 0 \;\; \forall k \in \mathbb{N},\; \forall x \in U_X$$

So explicitly computing these k-th order derivatives we can arrive at expressions such as:

$$0 = f(x, h(x)) \tag{8.3}$$
$$0 = f_X + f_Y h_X \tag{8.4}$$
$$0 = f_{XX} + 2 f_{XY} h_X + f_{YY} h_X^2 + f_Y h_{XX} \tag{8.5}$$

Since we know the function f, we may evaluate it's derivatives, furthermore we have that f_Y is invertible, thus 8.4 and 8.5 yield:

$$\begin{aligned} h_X &= -f_Y^{-1} f_X \\ h_{XX} &= -f_Y^{-1} \left(f_{XX} + 2 f_{XY} h_X + f_{YY} h_X^2 \right) \end{aligned}$$

Allowing us to generate terms of the Taylor expansion of h, as discussed in 8.1. We see that two tasks now arise: first we must be able to generate expressions for arbitrary derivatives:

$$D_X^k f(x, h(x))$$

And then also compute and manipulate arbitrary derivatives of the function f.

8.3 Generating Expressions for $D_X^k f(x, h(x)))$

First we will establish how to compute these derivatives, and automate this process since "by hand" computations quickly become laborious.

8.3.1 Iterating the derivative

An iterative approach is taken here, computing:

$$D_X^k f(x,h(x)) = D_X\left[D_X^{(k-1)} f(x,h(x))\right]$$

To aid this we make the following claim about the general form of the derivative:

Lemma

The expression for $D_X^k f(x,h(x))$ can be written as a sum of terms, where each term has general form:

$$f_{S_1 S_2 \cdots S_k}(x,h(x)) \cdot \left[P_1(X_{i_{1,1}},\cdots,X_{i_{1,l_1}}),\cdots,P_k(X_{i_{k,1}},\cdots,X_{i_{k,l_k}})\right] \tag{8.6}$$

Where $c \in \mathbb{R}$,

$$S_i \in \{X, Y\}$$

And the P_j's are either the identity map $id : X \to X$ (so $l_j = 1$), or a l_j-th derivative of h, i.e a l_j-linear map.

The P_i maps can be viewed as "premultiples", preparing the input from: $X \times \cdots \times X$ to be fed in the f partial. The $X_{i\ldots}$ indicates which version of X the input value for that slot is taken from.

We use a pair of of sub-sub-scripts for the X inputs so we can for example distinguish between $X_{i_{1,1}}$ and $X_{i_{2,1}}$, since P_1 and P_2 may take different inputs in their first slot. Since we will be generating lots of expressions in this form we drop the D^k from these terms for brevity.

Remarks

Previously little attention was payed to the ordering, but now it is imperative to keep track of how the derivatives develop.

Translating the expressions from 8.3, 8.4, 8.5 into this notation (omitting the $(x,h(x))$ parts for brevity), we have:

$$0 = f$$

$$0 = f_X \cdot [id(X_1)] + f_Y \cdot [(h_X(X_1)]$$

$$\begin{aligned} 0 = f_{XX} \cdot [id(X_1), id(X_2)] + f_{XY} \cdot [id(X_1), h_X(X_2)] \\ + f_{YX}\,[h_X(X_1), id(X_2)] + f_{YY}\,[h_X(X_1), h_X(X_2)] + f_Y \cdot [(h_{XX}(X_1, X_2)] \end{aligned} \tag{8.7}$$

We see that a k-th order derivative requires k inputs, as it is a k-linear map.

8.7 now has 5 terms, since we ignore the equivalence between f_{XY} and f_{YX}.

A note on notation

Since all partial derivatives are computed at 0 we drop this from the expressions for conciseness. Furthermore we now concatenate the subscripts instead of using commas.

Also here taking a further derivative is indicated by appending X or Y to the end, contrary to common use, the reason for this is that it makes the book-keeping a lot easier during the proof.

Proof of lemma

We first apply the product rule to see that:

$$D_X \left(f_{S_1 S_2 \cdots S_k} \cdot \left[P_1(X_{i_{1,1}}, \cdots, X_{i_{1,l_1}}), \cdots, P_k(X_{i_{k,1}}, \cdots, X_{i_{k,l_k}}) \right] \right) \tag{8.8}$$

$$= (D_X f_{S_1 S_2 \cdots S_k}) \cdot \left[P_1(X_{i_{1,1}}, \cdots, X_{i_{1,l_1}}), \cdots, P_k(X_{i_{k,1}}, \cdots, X_{i_{k,l_k}}) \right] \tag{8.9}$$

$$+ f_{S_1 S_2 \cdots S_k} \cdot \left(D_X \left[P_1(X_{i_{1,1}}, \cdots, X_{i_{1,l_1}}), \cdots, P_k(X_{i_{k,1}}, \cdots, X_{i_{k,l_k}}) \right] \right) \tag{8.10}$$

First consider 8.9, recalling that we are taking derivatives of $f(x, h(x))$, and thus must differentiate with respect to each slot, using the chain rule for the second slot:

$$(D_X f_{S_1 S_2 \cdots S_k}) \cdot \left[P_1(X_{i_{1,1}}, \cdots, X_{i_{1,l_1}}), \cdots, P_k(X_{i_{k,1}}, \cdots, X_{i_{k,l_k}}) \right] \tag{8.11}$$

$$= f_{S_1 S_2 \cdots S_k X} \cdot \left[P_1(X_{i_{1,1}}, \cdots, X_{i_{1,l_1}}), \cdots, P_k(X_{i_{k,1}}, \cdots, X_{i_{k,l_k}}), id(X_{k+1}) \right] \tag{8.12}$$

$$+ f_{S_1 S_2 \cdots S_k Y} \cdot \left[P_1(X_{i_{1,1}}, \cdots, X_{i_{1,l_1}}), \cdots, P_k(X_{i_{n,1}}, \cdots, X_{i_{k,l_k}}), h_X(X_{k+1}) \right] \tag{8.13}$$

Where both 8.12 and 8.13 are in the form of 8.6.

Now consider 8.10:

$$f_{S_1 S_2 \cdots S_k} \cdot \left(D_X \left[P_1(X_{i_{1,1}}, \cdots, X_{i_{1,l_1}}), \cdots, P_k(X_{i_{k,1}}, \cdots, X_{i_{k,l_k}}) \right] \right) \tag{8.14}$$

$$= f_{S_1 S_2 \cdots S_k} \sum_{j=1}^{k} \left[P_1(X_{i_{1,1}}, \cdots, X_{i_{1,l_1}}), \cdots, (D_X P_j)(X_{i_{j,1}}, \cdots, X_{i_{j,l_j}}, X_{k+1}), \cdots, P_k(X_{i_{k,1}}, \cdots, X_{i_{k,l_k}}) \right] \tag{8.15}$$

Observe that the X_{k+1} is now embedded inside the premultiples.

See that when P_j is the identity we have: $D_X P_j = D_X id = 0$, and so as expected the sum only picks up non-trivial premultiples. If P_j is a l_j-th derivative of h, it now becomes a $(l_j + 1)$-th, with $X_{i_{j,l_j+1}} = X_{k+1}$.

Thus by the the linearity of derivatives, we have that 8.15 is a sum of terms in the form of 8.6

Therefore we have shown that $D_X^k f(x,h(x))$ can be written as a sum of terms, where each term has general form:

$$f_{S_1 S_2 \cdots S_k}(x,h(x)) \cdot \Big[P_1(X_{i_{1,1}}, \cdots, X_{i_{1,l_1}}), \cdots, P_k(X_{i_{k,1}}, \cdots, X_{i_{k,l_k}})\Big] \tag{8.16}$$

As required.

□

While rather complicated, the fixed rules here indicate that we could iterated this process arbitrarily using a program, that is the next step.

8.3.2 Implementing in SageMath

The code for this program was written for use in SageMath a Python overhaul that implements computer algebra, (The Sage Developers 2021). However, for the initial representation of $D_X^k f(x,h(x))$ in the form outlined in the Lemma, mainly base Python was used. The classes used here can be found in the file implicit.py in the Bifurcations GitHub repository (Holmes 2022).

Generating expressions

A `Expression` class was implemented to generate arbitrarily large expressions for $D_X^k f(x,h(x))$. Calling the `diff()` method we can compute expressions for $D_X^k f(x,h(x))$, the first 3 are:

```
lprint(e.diff())
```

$$f_X \cdot [id(X_1)] + f_Y \cdot [h_X(X_1)]$$

```
lprint(e.diff().diff())
```

$$\begin{aligned} f_{XX} \cdot [id(X_1), id(X_2)] + f_{XY} \cdot [id(X_1), h_X(X_2)] \\ + f_{YX} \cdot [h_X(X_1), id(X_2)] + f_{YY} \cdot [h_X(X_1), h_X(X_2)] \\ + f_Y \cdot [h_{XX}(X_1, X_2)] \end{aligned}$$

```
lprint(e.diff().diff().diff())
```

$$
\begin{aligned}
f_{XXX} \cdot [id(X_1), id(X_2), id(X_3)] &+ f_{XXY} \cdot [id(X_1), id(X_2), h_X(X_3)] \\
+f_{XYX} \cdot [id(X_1), h_X(X_2), id(X_3)] &+ f_{XYY} \cdot [id(X_1), h_X(X_2), h_X(X_3)] \\
+f_{XY} \cdot [id(X_1), h_{XX}(X_2, X_3)] &+ f_{YXX} \cdot [h_X(X_1), id(X_2), id(X_3)] \\
+f_{YXY} \cdot [h_X(X_1), id(X_2), h_X(X_3)] &+ f_{YX} \cdot [h_{XX}(X_1, X_3), id(X_2)] \\
+f_{YYX} \cdot [h_X(X_1), h_X(X_2), id(X_3)] &+ f_{YYY} \cdot [h_X(X_1), h_X(X_2), h_X(X_3)] \\
+f_{YY} \cdot [h_{XX}(X_1, X_3), h_X(X_2)] &+ f_{YY} \cdot [h_X(X_1), h_{XX}(X_2, X_3)] \\
+f_{YX} \cdot [h_{XX}(X_1, X_2), id(X_3)] &+ f_{YY} \cdot [h_{XX}(X_1, X_2), h_X(X_3)] \\
&+ f_Y \cdot [h_{XXX}(X_1, X_2, X_3)]
\end{aligned}
$$

See that the last term in each expression contains a k-th derivative of h, while the f_Y term can always be divided out. Later expressions use derivatives of h, but always of order less than k, which we can compute inductively. Next we implement a means to perform symbolic computations using arbitrary derivatives of f and h

8.4 Symbolic Tensors

Just as we may combine a basis and a linear map to yield a matrix, multilinear maps can be expressed as arrangements of values in the underlying field. The objects that arise have the structure of multi-dimensional arrays, and we call them tensors. A symbolic tensor class was implemented to manipulate these objects, since most implementations only permit floating point arithmetic.

8.4.1 Tensors over a single space

For the purposes of this project we will refer to the representations of multilinear maps with respect to bases as tensors. Just as we distinguish matrices and linear maps.

First recall that a multilinear map over $\mathbb{R}^n$ is a map:

$$M : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}^m \tag{8.17}$$

That is linear in each slot. As a function mapping into $\mathbb{R}^m$ we can split M into m components, and so generally will only consider the case $m = 1$; the complexity arises from the structure of the domain.

The number of copies of $\mathbb{R}^n$, say k, is the order of the mapping, and we write

$$M \in \mathcal{M}^k(\mathbb{R}^n) \tag{8.18}$$

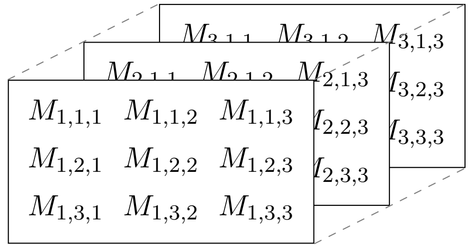


Figure 8.1: General form of 3rd order tensor over $\mathbb{R}^3$, we have each $M_{i_1,i_2,i_3} \in \mathbb{R}$

Hyper-array representation

Before representing the map, we must fix a basis, the canonical choice is $e_1, ..., e_n$. Then we may view M as a k dimensional hyper-array populated with n^k entries. Abusing notation we denote an element in the array:

$$M_{i_1,...,i_k} = M(e_{i_1}, ..., e_{i_k}) \tag{8.19}$$

Where $i_1, ..., i_k \in \{1, ..., n\}$, this array is the tensor corresponding to M.

8.4.2 Tensors over subspaces

For the applications in this project, we mainly consider derivatives of f with respect to two subspaces, X and Y. Therefore we need a means to represent high order partial derivatives with respect to these spaces. These take the form of "hyper-rectangular" tensors.

We generally consider:

$$\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m = X \times Y$$

Using the method of 8.19 we can represent any map $M \in \mathcal{M}^k(X \times Y)$ as a hyper-array with $(n+m)^k$ entries. We thus consider multilinear maps:

$$M : S_1 \times \cdots \times S_K \to \mathbb{R}$$

Where $S_i \in \{X, Y\}$.

We see that M can be embedded into a map:

$$\tilde{M} : (X \times Y) \times \cdots (X \times Y)$$

Such that if $S_i = X$ then:

$$\begin{aligned}\tilde{M}(\dots, x+y, \dots) &= \tilde{M}(\dots, x, \dots) + \tilde{M}(\dots, y, \dots) \\ &= \tilde{M}(\dots, x, \dots) + 0 \\ &= \tilde{M}(\dots, x, \dots)\end{aligned}$$

Where we have used linearity in the i-th slot, and set x and y to be the components of the vector in X and Y respectively. If $S_i = Y$, then the x component would evaluate to zero.

We can think of maps like M as occupying hyper-rectangular subsets of the larger hyper array corresponding to $\tilde{M}$, with the remainder padded with zeros. For the implementation these padding zeros will not be stored.

8.4.3 Implementation in SageMath

To compute expressions such as $f_{YYX} \cdot [h_X(X_1), h_X(X_2), id(X_3)]$, we must be able to store tensors such as f_{YYX}, and evaluate them at arbitrary values. The class `SymbolicXYTensor` was implemented to perform both these functions, and can be found in the file implicit.py in the Bifurcations GitHub repository (Holmes 2022).

Storing tensors

To store the hyper-array, the Pandas dataframe package was used. We see that that a multi-indexed dataframe has the structure of a hyper-array. Where each index level accounts for a different space. Since the symbolic tensors may store any python object, including standard integers, we may compare our implementation to the `numpy.array` object. In figure 8.2 we construct a 3 dimensional hyper-array using both `numpy.array` and `SymbolicXYTensor`. While the data frame implementation is less concise, it allows us to easily keep track of which of the X, Y spaces are used in each dimension, and makes tensor-vector multiplication simpler.

Tensor-vector multiplication

Given a generic k-th order tensor:

$$M : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$$

We can evaluate the tensor in one slot to yield a $(k-1)$-th dimensional tensor. We see this by observing that multilinear maps can also be viewed as nested linear maps, and thus denoting:

$$M(x_1, \ldots, x_{k-1}, x_k) = M^{x_k}(x_1, \ldots, x_{k-1})$$

For a fixed x_k. We wish to obtain the map M^{x_k}, since iterating this process will allow to evaluate the tensor at every x_i, eventually yielding a value $M(x_1, \ldots, x_k) \in \mathbb{R}$

```python
b = SymbolicXYTensor(x_dim = 3, y_dim = 2, xy_order = 'YYX')
b.data['data'] = [1, 2, 1, 0, 0, 1, -1, -1 , 2,1, 0 ,3 ]
b.data
```

1	2	3	data
y1	y1	x1	1
		x2	2
		x3	1
	y2	x1	0
		x2	0
		x3	1
y2	y1	x1	-1
		x2	-1
		x3	2
	y2	x1	1
		x2	0
		x3	3

(a) `SymbolicXYTensor` implementation

```python
x = np.array([[[1, 2, 1],
               [0, 0, 1]],
              [[-1, -1, 2],
               [1, 0, 3]]])
```

```python
print(x)
print(x.shape)
```

```
[[[ 1  2  1]
  [ 0  0  1]]

 [[-1 -1  2]
  [ 1  0  3]]]
(2, 2, 3)
```

(b) `numpy.array` hyperarray implementation

Figure 8.2: both implementations contain the same object, however the symbolic tensor also stores information about the subspaces

If we use the hyper array representation as per 8.19, then we have that:

$$M^{x_k}_{i_1,\ldots,i_{k-1}} = M(e_{i_1}, \ldots, e_{i_{k-1}}, x_k) \tag{8.20}$$

$$= \sum_{j=1}^{n} (x_k)_j M(e_{i_1}, \ldots, e_{i_{k-1}}, e_j) \tag{8.21}$$

$$= \sum_{j=1}^{n} M_{i_1,\ldots,i_{k-1},i_j} \cdot (x_k)_j \tag{8.22}$$

Where $(x_k)_j$ is the j-th component of x_k, so that $x_k = \sum_{j=1}^{n}(x_k)_j e_j$. Thus 8.21 follows from the linearity of M. This idea is easily extended to the case of a tensor over the subspaces X and Y.

The `vec_mult` method implements this approach, 8.22 is computed using a `groupby()` operation. In 8.3 we observe how the lowest multi-index level is removed; if the `vec_mult` method is called k times, the final call will output a scalar. We now have a method to evaluate arbitrary tensors.

Here we see the payoff from using the dataframe structure, since the collapsing multiplication can be phrased simply as a `groupby()` operation over a level of the multi-index.

```
b = SymbolicXYTensor(x_dim = 3, y_dim = 2, xy_order = 'YYX')
b.data['data'] = [1, 2, 1, 0, 0, 1, -1, -1 , 2,1, 0 ,3 ]
b.data
```

1	2	3	data
y1	y1	x1	1
		x2	2
		x3	1
	y2	x1	0
		x2	0
		x3	1
y2	y1	x1	-1
		x2	-1
		x3	2
	y2	x1	1
		x2	0
		x3	3

```
print(np.tensordot(x, np.array([1,1,1]), axes = 1))
```

```
[[4 1]
 [0 4]]
```

```
b.vec_mult([1,1,1]).data
```

1	2	data
y1	y1	4
	y2	1
y2	y1	0
	y2	4

Figure 8.3: The implemented `vec_mult` method produces the same output as the `numpy` implementation

The `SymbolicXYVectorTensor` class

So far we have built the machinery to work with multilinear maps of the form:

$$M : S_1 \times \cdots \times S_K \to \mathbb{R} \tag{8.23}$$

However for the problems we are interested in, we need to consider maps into higher dimensional spaces:

$$M : S_1 \times \cdots \times S_K \to \mathbb{R}^l \tag{8.24}$$

These can be thought of as having l components, each of which is a map with form as in 8.23. The `SymbolicXYVectorTensor` class does just this, keeping track of multiple `SymbolicXYTensor` objects, and performing operations like `vec_mult()` on each component tensor. The `evaluate` method was added to this class to evaluate the tensor in every slot, by performing multiple `vec_mult()` operations.

Now that we can interact with these objects, we will focus specifically on the tensors of interest for the implicit function theorem: derivatives of f. We consider the case where:

$$f : X \times Y \to \mathbb{R}$$

We then consider partial derivatives $M = f_{S_1,\ldots,S_k}$, seeing that:

$$M_{i_1,\ldots,i_k} = \partial_{s_{1,i_1}} \cdots \ \partial_{s_{k,i_k}} f$$

Where $\partial_{s_{j,i_j}} = \partial_{x,i_j}$ or $\partial_{s_{j,i_j}} = \partial_{y,i_j}$, we use s_j to first distinguish whether the derivative is with respect to the space X or Y, then the next subscript denotes with respect to which component we differentiate.

When we populate a tensor with values obtained this we, we obtain a realisation of the map $f_{S_1,\dots,S_k}$ with respect to the canonical bases.

The method `fill_from_functions` implements this, taking all derivatives at zero, using the SageMath inbuilt `diff` function to extract the correct evaluated partial derivative at each point. And thus allows us to store arbitrary partial derivatives of f in tensor form.

8.5 Obtaining Taylor Polynomials from Derivatives

Now we have the means to compute all but the last term in expressions such as:

$$\begin{aligned} f_{XX} \cdot [id(X_1), id(X_2)] + f_{XY} \cdot [id(X_1), h_X(X_2)] \\ + f_{YX} \cdot [h_X(X_1), id(X_2)] + f_{YY} \cdot [h_X(X_1), h_X(X_2)] \\ + f_Y \cdot [h_{XX}(X_1, X_2)] \end{aligned} \tag{8.25}$$

All that remains is to use this information to produce the coefficients of the Taylor polynomial approximation of h from this. For example here we aim to compute the tensor corresponding to h_{XX}. Which we can then use to find the second order terms in the Taylor polynomial of h. We see that 8.25 gives us a means to compute for example $D_X^2 h_0 = h_{XX}$, provided we know lower order derivatives of h.

To extract the coefficient of the polynomial that is obtained by evaluating $D_X^n h_0(x, \dots, x)$ we split x into the m variable components $x = x_1 + \cdots + x_m$ and separate out m^n terms using the linearity of $D_X^n h_0$, each will correspond to an element in the hyper-array representation.

For example if $n = 3$, then the x_1^3 coefficient is only obtained by evaluating $D_X^3 h_0(x_1, x_1, x_1) = ax_3^3$ (this corresponds to a "diagonal entry" in the 3D array. Meanwhile, the $x_1x_2x_3$ coefficient is obtained by summing over the 6 permutations that exist of ways to evaluate $D_X^n h_0(x_i, x_j, x_k) = bx_ix_jx_k$, where i, j, k are different.

In practise since these correspond to different orderings of computing $\partial_{x_1}\partial_{x_2}\partial_{x_3}$, which by properties of continuously differentiable functions they will be all the same, a property that will be exploited later when increasing the efficiency of the algorithm.

8.6 Implicit Function Theorem Implementation

Now we have all the requisite components to implement the approximate implicit function theorem, this is obtained by calling the `get_hkx_polynomial` function.

This is a recursive function on k, the order of the Taylor polynomial approximation of h requested. The function recursively calls itself for lower values of k, then appends on the order k terms. The base case $k = 0$ will always yield the zero polynomial since we know that the constant terms of h must be zero since the theorem asserts that $h(0) = 0$.

To find the order k terms, first the expressions for $D^k_X f(x, h(x))$ are computed as per 8.3, then the symbolic tensors for each term in the sum are computed separately. They are stored so can be viewed after the function call, and reused at different levels of the recursion.

Finally functions are called that evaluate tensors; collect polynomial terms; and manipulate polynomial terms. Yielding the output of a Taylor polynomial approximation of h, to order k.

8.6.1 Testing

The code was tested in the Jupyter notebook Implicit Testing.ipynb (Holmes 2022), where it is can be observed approximating the solutions to functions with increasing accuracy as the order of the approximation is increased. After specifying the variables corresponding to the X and Y spaces

8.6.2 Optimisations

While this implementation is by no means intended to be optimised in terms of computational cost, a few alterations were made to provide speedup.

Precomputing f as a Taylor polynomial

A large cost is associated with computing various partial derivatives of f, we can outsource this to the inbuilt Taylor approximation function in SageMath. From which the symbolic tensors entries can be filled in by reading off the entries.

Exploiting symmetry

As discussed in section 8.5 the tensors derived from derivatives of f have symmetries induced by the various orders in which we can take derivatives.

By using the `itertools.combinations_with_replacement` function we can loop over only distinct terms in the polynomial and then rescale by how many terms each is equivalent to, this number of terms is computed as:

```
Integer(factorial(order) / prod([factorial(combo.count(item))
    for item in set(combo)]))
```

Chapter 9

Computational Bifurcation Theory

Now equipped with a computational implementation of the implicit function theorem, we can begin to approach finite dimensional problems in bifurcation theory using the computer algebra package SageMath. We will aim to both identify bifurcation points and find approximations for the new curve of non-trivial solutions.

The first problem that arises is that our implicit function theorem places fairly strict requirements on the types of functions processed. The implementation requires that we can split the domain into two spaces $\mathbb{R}^n \times \mathbb{R}^m$, where the derivative with respect to $\mathbb{R}^m$ is an isomorphism.

In the setup of Crandall-Rabinowitz there are no such constraints, the kernel directions by no means neatly fit into only certain canonical coordinate directions. Furthermore Crandall-Rabinowitz dictates that the Jacobian with respect to X can not be subjective (by considering dimensions of X and Y and rank-nullity). Thus we will need to trim down the codomain to apply the implicit function theorem.

9.1 The Problem Setup

Crandall-Rabinowitz concerns functions of the form:

$$f : \mathbb{F} \times X \to Y$$

Since we are now working explicitly, we now will replace these with real spaces, so that:

$$f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

The $\mathbb{R}$ component of the domain will always be identified with the variable λ, and we assume that $f(\lambda, 0) = 0$ for all $\lambda \in \mathbb{R}$.

9.1.1 Example

To aid the description of the process an example will be worked through , the “by hand”” computations can be found in B.0.2, the computational approach used now be seen in the notebook Crandall Rabinowitz.ipynb (Holmes 2022).

$$f(\lambda, x, y, z) = \begin{bmatrix} f_1(\lambda, x, y, z) \\ f_2(\lambda, x, y, z) \\ f_3(\lambda, x, y, z) \end{bmatrix} \in \mathbb{R}^3$$

$$\begin{aligned} f_1(\lambda, x, y, z) &= -y\cos\left(\frac{1}{3}\sqrt{\pi^2 + 9\,x^2 + 9\,y^2}\right) + x\sin\left(\frac{1}{3}\sqrt{\pi^2 + 9\,x^2 + 9\,y^2}\right) \\ f_2(\lambda, x, y, z) &= x\cos(\lambda) + y\sin(\lambda) \\ f_3(\lambda, x, y, z) &= y^3 + x^2 + \lambda z + z^2 + 4\,z \end{aligned}$$

We observe that $f(\lambda, 0, 0, 0) = 0)$ for all $\lambda \in \mathbb{R}$

9.2 The Necessary Condition

As discussed in 3.2, a bifurcation at λ_0 is only possible if $D_X f_{(\lambda_0,0)}$ is non-invertible. Thus the first step we can take is to compute the Jacobian with respect to the $\mathbb{R}^n$ portion of the domain and evaluate at the zero vector, leaving the λ variable free. Recall that a matrix is invertible if and only if it has non-zero determinant. Thus considering the function:

$$J : \mathbb{R} \to \mathbb{R}^{n^2}$$

$$J(\lambda) = D_X f_{(\lambda,0)}$$

Where we identify $\mathbb{R}^{n^2}$ with the space of matrices and view $D_X f_{(\lambda,0)}$ as the Jacobian of f at $(\lambda, 0)$, with respect to the canonical bases.

Then λ_0 can possibly be a bifurcation only if

$$\det(J(\lambda_0)) = 0$$

Thus if we can find a closed form for the solutions of $\det(J(\lambda))$ then we can characterise all possible potential bifurcation points. However we recall this is by no means sufficient to confirm that such points are indeed bifurcations.

9.2.1 Example

Following from the example we now compute the Jacobian at $x = y = z = 0$

$$J(\lambda) = \begin{pmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 1 \\ \cos(\lambda) & \sin(\lambda) & 1 \\ 0 & 0 & \lambda + 4 \end{pmatrix}$$

Which has determinant:

$$\det(J(\lambda)) = \frac{1}{2}\sqrt{3}(\lambda + 4)\sin(\lambda) + \frac{1}{2}(\lambda + 4)\cos(\lambda)$$

We previously saw that there is a bifurcation point at $\lambda = -\frac{\pi}{6}$, we can verify that

$$\det\left(J\left(-\frac{\pi}{6}\right)\right) = 0$$

9.3 Regularising the Problem

Once a point λ_0 is located we wish determine if is indeed a bifurcation, and if so approximate the new curve of solutions. To do this we will regularise the functions; in the spirit of Lyapunov-Schmidt we wish to use bases derived from the subspace decompositions:

$$X = \mathbb{R}^n = \ker J(\lambda_0) \oplus W \;\text{ and }\; Y = \mathbb{R}^n = Z \oplus \operatorname{Im} J(\lambda_0) \tag{9.1}$$

Now we non longer refer to the derivative as L, instead choosing $J(\lambda_0)$ to emphasise that it is nothing more than a square matrix of order n, not an abstract linear map.

See that if we find invertible matrices A and B such that $B \circ J(\lambda_0) \circ A$ is a block matrix as in 9.3.1 then we can compute:

$$\tilde{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n \tag{9.2}$$

$$\tilde{f}(\lambda, x) = B \circ f(\lambda, \cdot) \circ A(x) \tag{9.3}$$

Which when viewed as mapping only into it's image, will have all the desired properties of the implicit function theorem.

Furthermore if we can characterise the zero set of $\tilde{f}$ then simply applying the map A to this will produce the zero set of f, as desired. Using that B invertible means that multiplying by B at the end has no impact on the zero set, since it has trivial kernel.

In fact establishing that $\tilde{f}$ has a bifurcation at λ_0 is sufficient to determine the f has a bifurcation at λ_0, since invertible matrices are homeomorphisms, preserving the topology of the zero set.

9.3.1 Example

Taking $\lambda_0 = -\frac{\pi}{6}$ we see that:

$$J(\lambda_0) = \begin{pmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 1 \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 1 \\ 0 & 0 & -\frac{1}{6}\pi + 4 \end{pmatrix} \tag{9.4}$$

Which has kernel:

$$\ker J(\lambda_0) = \langle \xi_0 \rangle = \langle \left(1, \sqrt{3}, 0\right)^T \rangle \tag{9.5}$$

Fixing the kernel

First we seek the matrix A that ensures the kernel of the Jacobian of $\tilde{f}$ is restricted to one canonical basis vector, we see that if we extend $\{\xi_0\}$ to an orthogonal basis, say $\{\xi_0, u_2, u_3\}$, then we can construct A as follows:

$$A = \begin{pmatrix} | & | & | \\ \xi_0 & u_2 & u_3 \\ | & | & | \end{pmatrix}$$

We observe that $Ae_1 = \xi_0$, and $Ae_i = u_i$ for $i \in \{2, 3\}$, where the u_i's lie in the subspace perpendicular to $\ker J(\lambda_0)$

To do this in SageMath we run (assuming that $J = J(\lambda_0)$):

```
K = J.right_kernel().basis()
Kperp = matrix(K).right_kernel().basis()
A = matrix(K+Kperp).transpose()
# transpose since SageMath use row vectors
```

If we do this we find:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ \sqrt{3} & -\frac{1}{3}\sqrt{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Fixing the image

Now we construct the matrix B, this will reorder the function $f(\lambda, \cdot) \circ A$ so as to restrict the image to 2 canonical basis vector directions. Given a basis of the image of the Jacobian of $f(\lambda, \cdot) \circ A$ at λ_0, say $\{v_1, v_2\}$, we can orthogonally extend to $\{v_1, v_2, z_3\}$, then we can construct B as:

$$B = \begin{pmatrix} | & | & | \\ v_1 & v_2 & z_3 \\ | & | & | \end{pmatrix}^{-1}$$

Will be such that $Bv_i = e_i$ for $i \in 1, 2$, so the image is restricted to two canonical basis directions.

Doing this in SageMath:

```
f1A(x,y,z,l) = f1(*list(A*vector([x,y,z])),l)
f2A(x,y,z,l) = f2(*list(A*vector([x,y,z])),l)
f3A(x,y,z,l) = f3(*list(A*vector([x,y,z])),l)
```

```
Ja = jacobian([f1A,f2A,f3A],(x,y,z))(x=0,y=0,z=0,l=-pi/6)

I = Ja.image().basis() # not image in our understanding
I = [Ja*vec for vec in I] # fixes that

Iperp = matrix(I).right_kernel().basis()
B = matrix(I + Iperp).transpose().inverse()
# transpose since uses row vectors
```

Which yields:

$$B = \begin{pmatrix} \frac{1}{4}\sqrt{3} & \frac{1}{4}\sqrt{3} & \frac{3\sqrt{3}}{\pi-24} \\ 0 & 0 & -\frac{6}{\pi-24} \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}$$

Result

Now we can construct $\tilde{f}$:

```
[ft1,ft2,ft3] = B*vector([f1A,f2A,f3A])
```

The resultant equations are somewhat unwieldy so are omitted, however we can compute the new Jacobian, $\tilde{J}(\lambda_0)$

$$\tilde{J}(\lambda_0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Which we see is of form:

$$\left(\begin{array}{c|c} 0 & I \\ \hline 0 & 0 \end{array}\right)$$

Which up to reordering of the subspaces specified, looks like matrix obtained in the first half of the proof of Lyapunov-Schmidt, as desired.

9.4 Applying the Implicit Function Theorem

We now have $X = \mathbb{R}^n = \ker \tilde{J}(\lambda_0) \oplus W$ where $\tilde{J}(\lambda_0)$ is the Jacobian of $\tilde{f}$ at λ_0. However by the normalised kernel, we can write:

$$\ker \tilde{J}(\lambda_0) \oplus W = \mathbb{R} \times \mathbb{R}^{n-1}$$

Similarly by the normalised the image:

$$Z \oplus \operatorname{Im} \tilde{J}(\lambda_0) = \mathbb{R} \times \mathbb{R}^{n-1}$$

If we now restrict the image of $\tilde{f}$ to $\operatorname{Im} \tilde{J}(\lambda_0)$ we obtain the function So we can view $\tilde{f}$ as a function:

$$\tilde{f}_{\text{rest}} : \mathbb{R}^2 \times \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$$

Where $\mathbb{R}^2$ contains the λ and ξ_0 kernel directions, and the codomain is $\operatorname{Im} \tilde{J}(\lambda_0)$. Recall we write $\ker \tilde{J}(\lambda_0) = \langle \xi_0 \rangle$. Now we are in the setup of the implicit function theorem. Thus we can apply the computational implementation of the implicit function theorem and obtain a polynomial approximation h

$$h : \mathbb{R}^2 \to \mathbb{R}^{n-1} \tag{9.6}$$

That parametrises the zero set of $\tilde{f}_{\text{rest}}$ close to zero.

9.4.1 Remark

While we have parametrised the zero set of $\tilde{f}_{\text{rest}}$, this is not the same as the zero set of $\tilde{f}$. An important distinction.

9.4.2 Example

In our example we have $n = 3$, we omit the last component of $\tilde{f}$, and treat λ, x as "X" variables, and y, z as "Y" variables, as defined in the statement of the implicit function theorem. Thus after running the computational implementation of the implicit function theorem, obtain a Taylor approximation of $h : \mathbb{R}^2 \to \mathbb{R}^2$. . As input variables to $\tilde{f}$, we cannot interpret them in the same way as when we defined f in terms of x, y, z, λ.

If we set the order of the approximation to be 2 then the output is:

$$h(\lambda, x) = \begin{bmatrix} -\frac{3\,(\pi-24)\lambda x}{2\left(\sqrt{3}\pi-24\,\sqrt{3}\right)} - \frac{9\,x^2}{\sqrt{3}\pi-24\,\sqrt{3}} \\ \frac{6\,x^2}{\pi-24} \end{bmatrix}$$

Note

Without loss of generality we have shifted the λ input of $\tilde{f}$ by $-\frac{\pi}{6}$ to ensure that all Taylor polynomial approximations are computed at zero. This is minor alteration that can easily be undone once the analysis is complete.

9.5 A Corollary

This h approximation will prove useful in approximating a new curve of solutions - provided that they exist. Crandall-Rabinowitz requires transversality, however we now show that we can substitute this for a constraint on a Taylor approximation generated from the h produced, to rigorously prove that a bifurcation occurs.

9.5.1 Notation

At this point we run into notational difficulties, since the implicit function theorem and Crandall-Rabinowitz theorem both refer to a function h, these are not the same function, and yet unfortunately both required in the following corollary.

From now on in this section we rename the h as produced in 9.6, to ϕ, so we have:

$$\phi : \mathbb{R}^2 \to \mathbb{R}^{n-1} \tag{9.7}$$

We we now exclusively refer to h as used in the proof of Crandall-Rabinowitz 5.2, which is defined in the proof of Lyapunov-Schmidt 4.3. We see that the newly christened ϕ is in fact the same ϕ that is used when defining this h.

9.5.2 Corollary

Given a function $f : \mathbb{F} \times X \to Y$ k-times differentiable, $k \geq 2$, with trivial solutions $f(\lambda, 0) = 0 \in Y$ for all $\lambda \in \mathbb{F}$, where $\dim X = \dim Y$

And a point $(\lambda_0, 0)$ such that $L = D_X f_{\lambda_0,0)}$ has one dimensional kernel $\ker L = \langle \xi_0 \rangle$.

We may then reduce this to solving a new function h, by applying a Lyapunov-Schmidt reduction to the equation at $(\lambda_0, 0)$:

$$h : \mathbb{F} \times \ker L \to \mathbb{R} \tag{9.8}$$

If the Taylor approximation of this h has a non-zero coefficient for the $\lambda\xi_0$ term, then λ_0 is a bifurcation point.

9.5.3 Remarks

This corollary mimics the Crandall-Rabinowitz theorem, except now we replace the transversality assumption for a restriction on the Taylor polynomial of h. We will go on to see that we can compute the Taylor approximation of this h, armed only with f and ϕ, and thus establish if a point is indeed a bifurcation in an entirely automated fashion.

Whilst computing the value of $D^2_{\mathbb{F},X} f_{(\lambda_0,0)}(1, \xi_0)$ and using the standard formulation may be more direct, we will see that access to ϕ and h allow us to approximate the new curve of solutions.

9.5.4 proof of corollary

This corollary follows almost immediately upon inspection of the proof of Crandall-Rabinowitz 5.2, since we only need to reprove arguments that depended upon the now revoked transversality assumption 5.1.

However the only time we use transversality is whilst proving A.3, which asserts that:

$$D^2_{\mathbb{F},\langle\xi_0\rangle} h_{\lambda_0,0}(1,\xi_0) \neq 0$$

However we see that this is precisely the assumption we have just made! Since after scaling, the value of $D^2_{\mathbb{F},\langle\xi_0\rangle} h_{\lambda_0,0}(1,\xi_0)$ will be the coefficient of the $\lambda\xi_0$ term in the Taylor polynomial of h.

And thus the conclusions of the Crandall-Rabinowitz theorem follow, as required. $\square$

9.6 Computing h From ϕ

We now require a way to compute $h : \mathbb{R}^2 \to \mathbb{R}$, given $\phi : \mathbb{R}^2 \to \mathbb{R}^{n-1}$. To do this we refer to the proof of the Lyapunov-Schmidt theorem 4.3. The first half of the proof is concerned with ensuring that the derivative has block matrix form, which is what we have achieved by producing $\tilde{f}$.

Then the function $\psi : \mathbb{R}^2 \to X(= \mathbb{R}^n)$ is defined, which we will see appends the information about λ and ξ into ϕ, and maps into the full space - not just Im $\tilde{J}(\lambda_0)$. We then finally define:

$$h(\lambda, \xi) = P\tilde{f}(\lambda, \psi(\lambda, \xi)) \in Z$$

Where P denotes projecting into the Z. Since we have that $Z = \mathbb{R}$, we see that:

$$h : \mathbb{R}^2 \to \mathbb{R}$$

9.6.1 Example

Recall that inputs have been shifted so that $\lambda = 0$ now corresponds to the potential bifurcation point.

In our example the projection P just corresponds to picking out the third component of

$\tilde{f}$, which has become:

$$\begin{aligned}
\tilde{f}_3(\lambda, x, y, z) = &-\frac{1}{2}(x+y)\cos\left(-\frac{1}{6}\pi+\lambda\right) \\
&-\frac{1}{6}\left(3\sqrt{3}x-\sqrt{3}y\right)\cos\left(\frac{1}{3}\sqrt{\pi^2+36\,x^2+12\,y^2}\right) \\
&-\frac{1}{6}\left(3\sqrt{3}x-\sqrt{3}y\right)\sin\left(-\frac{1}{6}\pi+\lambda\right) \\
&+\frac{1}{2}(x+y)\sin\left(\frac{1}{3}\sqrt{\pi^2+36\,x^2+12\,y^2}\right)-\frac{1}{2}z+\frac{1}{2}e^z-\frac{1}{2}
\end{aligned}$$

We only need to consider the Taylor expansion of h up to second order however, so it suffices to only consider the Taylor expansion of $\tilde{f}_3$ to second order, at $x = y = z = 0$, $\lambda = 0$, which can be computed to be:

$$\tilde{f}_3(\lambda, x, y, z) \approx -\lambda x + \frac{1}{4}z^2$$

ϕ gives us approximations for z and y in terms of λ and x, from the application of the implicit function theorem. We can then compute h by substituting the approximations for z, and (vacuously) y into $\tilde{f}_3$:

$$h \approx \tilde{f}_3(\lambda, x, \phi(x,\lambda)_1, \phi(x,\lambda)_2) \approx \left(\frac{9}{(\pi-24)^2}\right)x^4 - \lambda x$$

And thus by corollary 9.5.2, we conclude that $\lambda_0 = -\frac{\pi}{6}$ is indeed a bifurcation point (in the unshifted original functions)

Note that the x^4 term here is somewhat spurious, the process above can only ensure that terms of order 2 or below are exactly correct, luckily that is all we require to apply the corollary.

9.6.2 Approximating the new curve of solutions for $\tilde{f}$

We now compute approximate solutions, continuing from the example of 9.6.1.

Second order approximation

Concatenating h to only the second order approximation we are left with λx, which has solution set $\{x = 0\} \cup \{\lambda = 0\} \subset \mathbb{R}^2$, the non-trivial part being the $\{\lambda = 0\}$ component, we can use this to approximate the non-trivial solutions.

```
eps = 0.01

y_est = phi[0](x = eps, l = 0)
z_est = phi[1](x = eps, l = 0)

print(float(ft1(eps, y_est, z_est, 0)))
print(float(ft2(eps, y_est, z_est, 0)))
print(float(ft3(eps, y_est, z_est, 0)))
```

```
2.3832034103186824e-07
1.634707100460723e-06
1.9098957265272637e-06
```

Figure 9.1: We choose a point on the $\{\lambda = 0\}$ line of new solutions, `l` here denotes λ, observing that the approximation has low error for each component of $\tilde{f}$

Third order approximation

Now we attempt to approximate the new curve of solutions. We repeat the process from before, except now computing an order 3 Taylor polynomial approximation of ϕ, $\tilde{f}_3$ and thus h, which we find to be:

$$\frac{6}{\pi}x^3 - \lambda x$$

If we plot the zero set of this polynomial, we see the classic "pitchfork"" bifurcation diagram:

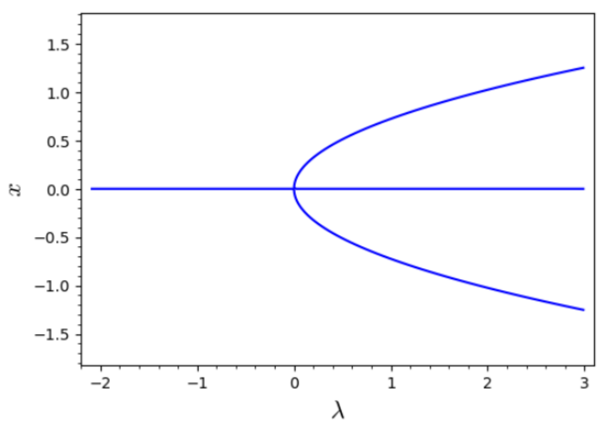


Figure 9.2: Zero set of $\frac{6}{\pi}x^3 - \lambda x$

And we see that:

$$\lambda = \frac{6}{\pi}x^2$$

parametrises the non-trivial component by x, and thus we can approximate a non-trivial solution as before, reducing the error.

```
lprint(better)
```

$$\left[\lambda = \frac{6\,x^2}{\pi}\right]$$

```
l_est = float(better[0].right()(x = eps))
y_est = float(phi[0](x = eps, l = l_est))
z_est = float(phi[1](x = eps, l = l_est))
```

```
print(float(ft1(eps, y_est, z_est, l_est)))
print(float(ft2(eps, y_est, z_est, l_est)))
print(float(ft3(eps, y_est, z_est, l_est)))
```

```
1.1792126274142682e-07
-1.3610204373291223e-07
6.221208791179489e-11
```

Figure 9.3: approximate solutions now using cubics

Converting to solutions of f

We now check that these solutions of $\tilde{f} = B^{-1} \circ f \circ A$ do indeed correspond to solutions of f, we saw that we can parametrise the new non-trivial solutions to $\tilde{f}$, by f, yielding

$$(x, y(x), z(x), \lambda(x))$$

To get solutions to f we apply A to the $(x, y(x), z(x))$ terms, viewing them as a vector in $\mathbb{R}^3$, recalling that invertibility of A and B ensures that the solution set structure is preserved. In 9.4 we see this was applied using the cubic approximations found in 9.6.2.

```
l_est = better[0].right()
xt_est = x
yt_est = out[0](x = xt_est, l = l_est)
zt_est = out[1](x = xt_est, l = l_est)
```

```
[x_est, y_est, z_est] = list(A*vector([xt_est,yt_est,zt_est]))
```

```
x_est1 = x_est(x = eps)
y_est1 = y_est(x = eps)
z_est1 = z_est(x = eps)
l_est1 = l_est(x = eps)
print(float(f1(x_est1, y_est1, z_est1, l_est1)))
print(float(f2(x_est1, y_est1, z_est1, l_est1)))
print(float(f3(x_est1, y_est1, z_est1, l_est1)))
```

```
1.2391393119925986e-10
-5.102544882403431e-13
-4.7314531147663847e-07
```

Figure 9.4: approximate solutions to the original problem still have low error

Chapter 10

Example - Discrete Elastic Rod

We consider again the elastic rod, however now a discretisation, so that the spaces X and Y become $\mathbb{R}^n$, and we can apply computational methods.

10.1 The Model

We consider a rod of consisting of $n+1$ connected unit length components. Each forming angle θ_i with the horizontal. The leftmost component's left endpoint is pinned to the origin, and the rightmost component's right endpoint is free to move in the horizontal direction, but pinned in the vertical. See 10.1 for an example.

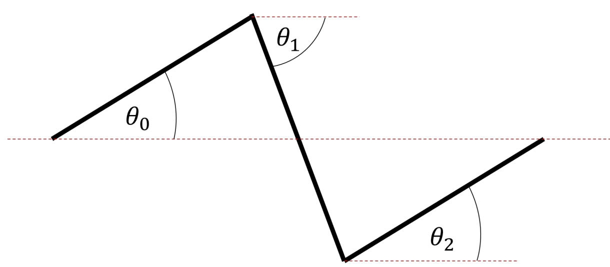


Figure 10.1: model in the case $n = 2$

From the right endpoint's vertical constraint we conclude that:

$$\sum_{i=0}^{n} \sin\theta_i = 0 \tag{10.1}$$

We suppose there is a spring at each join, that resists angular perturbation from rest, where all join angles are π radians. In 10.1 there would be springs at θ_1 and θ_2. In total there will be n springs. Using parallel line angle identities we find that the joint angle at the θ_ith joint is $\pi + \theta_i - \theta_{i-1}$, and so we then compute the i-th spring energy to be:

$$e_i = (\pi - (\pi + \theta_i - \theta_{i-1}))^2 = (\theta_i - \theta_{i-1})^2$$

We also consider horizontal force λ, the energy induced by this force is the work done, which is equal to λH, where H is the horizontal perturbation of the endpoint. We compute H:

$$H = n + 1 - \sum_{i=0}^{n} \cos \theta_i$$

And thus define E

$$E(\lambda, \theta_0, \ldots, \theta_{n-1}) = -H\lambda + \sum_{i=1}^{n} e_i$$

The negative sign here means that we trade work done by the force for angular perturbations in the springs. We see that although there are $n + 1$ angles, we can only control the first n, since the last angle will be uniquely determined to ensure that the right endpoint has zero vertical displacement. The final angle θ_n can be expressed in terms of $\theta_0, \ldots, \theta_{n-1}$ via 10.1, since $\sin \theta_n = -\sum_{i=0}^{n-1} \sin \theta_i$. We thus define:

$$f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

$$f_i(\lambda, \theta_0, \ldots, \theta_{n-1}) = \frac{d}{d\theta_{i-1}} E(\lambda, \theta_0, \ldots, \theta_{n-1})$$

Zeros of f, correspond to configurations that will remain at rest. We can show that $\theta_0 = \cdots = \theta_{i-1} = 0$ is always a solution to f, regardless of the value of λ.

A function was written in SageMath to produce such f, given n, see the ElasticRodCompBifs.ipynb notebook in (Holmes 2022).

Note

We are mainly concerned with the bifurcation properties of this equation, not its physical feasibility, and thus omit rigorous discussions of the derivation.

10.2 $n = 2$

When the dimensions are low the computation time is not a hindrance. For this case we can fully evaluate the bifurcation properties of the model. The computations can be seen in the notebooks ElasticRodCompBifs.ipynb (Holmes 2022).

10.2.1 The Equations

For this example we have:

$$f : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$$

$$
\begin{aligned}
f_1(\theta_0, \theta_1, \lambda) = &- \left(\frac{(\sin(\theta_0) + \sin(\theta_1)) \cos(\theta_0)}{\sqrt{-(\sin(\theta_0) + \sin(\theta_1))^2 + 1}} + \sin(\theta_0) \right) \lambda \\
&+ \frac{2(\theta_1 - \arcsin(-\sin(\theta_0) - \sin(\theta_1))) \cos(\theta_0)}{\sqrt{-(\sin(\theta_0) + \sin(\theta_1))^2 + 1}} \\
&+ 2\theta_0 - 2\theta_1 \\
f_2(\theta_0, \theta_1, \lambda) = &- \left(\frac{(\sin(\theta_0) + \sin(\theta_1)) \cos(\theta_1)}{\sqrt{-(\sin(\theta_0) + \sin(\theta_1))^2 + 1}} + \sin(\theta_1) \right) \lambda \\
&+ 2(\theta_1 - \arcsin(-\sin(\theta_0) - \sin(\theta_1))) \left(\frac{\cos(\theta_1)}{\sqrt{-(\sin(\theta_0) + \sin(\theta_1))^2 + 1}} + 1 \right) \\
&- 2\theta_0 + 2\theta_1
\end{aligned}
$$

10.2.2 The Necessary Condition

We compute the Jacobian with respect to $X = \mathbb{R}^2$ at $\theta_0 = 0$, $\theta_1 = 0$, leaving λ free.

$$
J = \begin{pmatrix} -2\lambda + 4 & -\lambda + 2 \\ -\lambda + 2 & -2\lambda + 10 \end{pmatrix}
$$

Which has determinant

$$
\det J = -(\lambda - 2)^2 + 4(\lambda - 2)(\lambda - 5)
$$

As a polynomial we can solve for λ and see that $\lambda = 2$, $\lambda = 6$ are the only possible bifurcation points.

From the corresponding matrices, we can extract the kernel direction, which we often call ξ_0, we can then produce visualisations of the potential bifurcations, see 12.1, 12.2

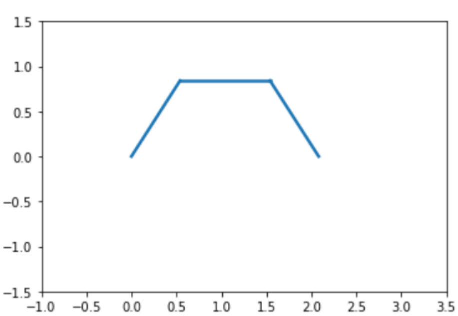


Figure 10.2: $\lambda = 2$ potential bifurcation

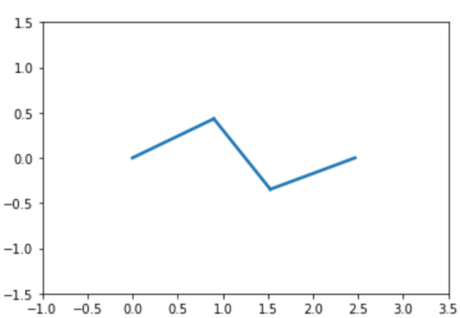


Figure 10.3: $\lambda = 6$ potential bifurcation

10.2.3 Computing h

First we regularise the problems to $\tilde{f}$, using the process from 9.3, also shifting so that $\lambda = 0$ at the potential bifurcation point. We then apply the implicit function theorem code to find a quintic approximation of the map ϕ. Beyond quintic order computational overhead becomes unnecessarily large.

Then following the method used in 9.6, we can obtain terms up to the quintic of the Taylor expansion of h:

$$h_{\lambda=2} = -\frac{2}{27}\,\lambda^4\theta_0 + \frac{7}{27}\,\lambda^2{\theta_0}^3 - \frac{4}{45}\,{\theta_0}^5 - \frac{1}{9}\,\lambda^3\theta_0 + \frac{5}{9}\,\lambda{\theta_0}^3 - \frac{1}{6}\,\lambda^2\theta_0 + \frac{2}{3}\,{\theta_0}^3 - 2\,\lambda\theta_0$$

$$h_{\lambda=6} = \frac{882}{78125}\,\lambda^4\theta_0 + \frac{57}{125}\,\lambda^2{\theta_0}^3 + \frac{9}{125}\,{\theta_0}^5 - \frac{63}{3125}\,\lambda^3\theta_0 + \frac{61}{125}\,\lambda{\theta_0}^3 + \frac{9}{250}\,\lambda^2\theta_0 + \frac{18}{5}\,{\theta_0}^3 - \frac{6}{5}\,\lambda\theta_0$$

Here θ_0 denotes the kernel direction. Thus by corollary 9.5.2 we conclude that both $\lambda = 2$ and $\lambda = 6$ are bifurcation points. Since the $\lambda\theta_0$ have non-zero coefficients for both h.

If we plot the zero sets of these functions we see that both are "pitchfork" bifurcations, which we expected, since each solution has a corresponding solution after reflection in the horizontal.

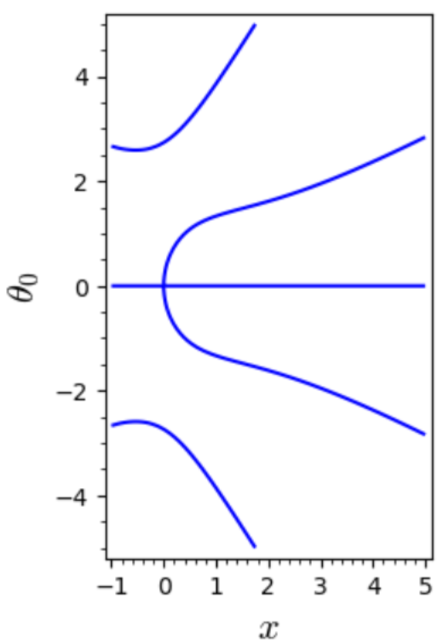


Figure 10.4: zero set of h when $\lambda = 2$

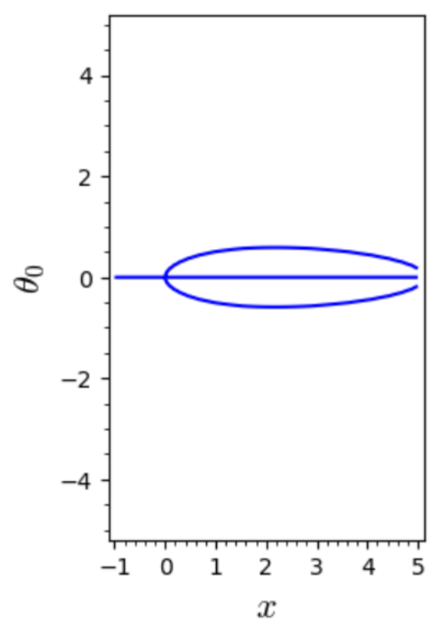


Figure 10.5: zero set of h when $\lambda = 6$

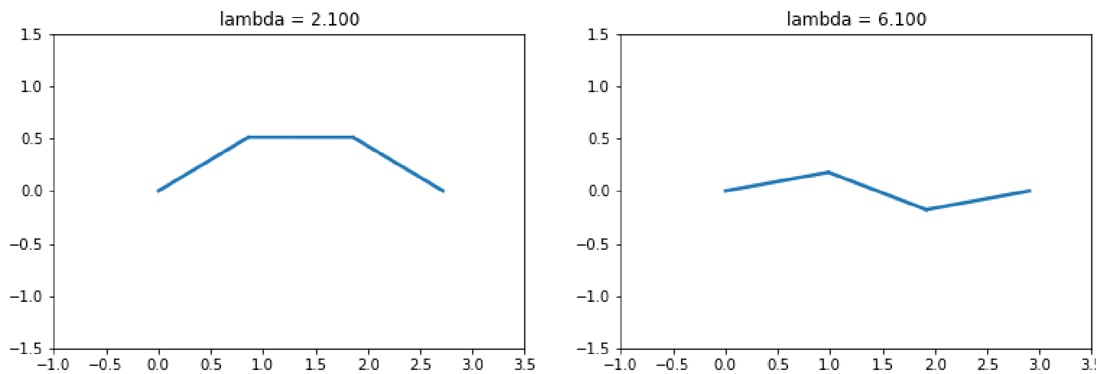
The higher order terms contribute to the changes in shape to these curves. However the theory we have developed is only concerned with the local structure. When $\lambda = 2$ we see extra solutions, however since these do not intersect the line of trivial solutions, our theory cannot establish if they correspond to solutions of the problem before approximations are taken.

10.2.4 Approximating Non-Trivial Solutions

Now we repeat the process of 9.6.2, except now solving h for θ_0 in terms of λ. For example when $\lambda = 2$, the upper part of the pitchfork has curve (found using SageMath solve functionality):

$$\theta_0 = \frac{1}{2}\sqrt{\frac{35}{6}\lambda^2 + \frac{25}{2}\lambda - \frac{1}{6}\sqrt{745\lambda^4 + 4530\lambda^3 + 10845\lambda^2 + 540\lambda + 8100} + 15}$$

We can then use ϕ to generate solutions of $\tilde{f}$ from these, which we then map to solutions of f using the A matrix. This allows us to specify the force λ and see what the geometry of the structure would be:



10.3 $n = 7$

We now consider a model with 8 component parts, so that the function generated is:

$$f : \mathbb{R} \times \mathbb{R}^7 \to \mathbb{R}^7 \tag{10.2}$$

The computations are in ElasticRodMaxLen.ipynb (Holmes 2022), for this example we still are able to compute a Jacobian as in 10.2.2, which we find has determinant:

$$\det J = -8\lambda^7 + 224\lambda^6 - 2496\lambda^5 + 14080\lambda^4 - 42240\lambda^3 + 64512\lambda^2 - 43008\lambda + 8192$$

We can use SageMath solve functionality to find the roots of this polynomial in closed form, these are the potential bifurcation points.

Computing the corresponding ξ_0 vectors, we see in 10.7 that all the potential bifurcation points look like discrete sine waves, mimicking the solutions set structure seen in chapter 7. For this example the kernels of the Jacobian for each λ were found using numerical methods, due to the difficulty of working with large symbolic matrices.

The presence of square roots in all but one of the potential bifurcation points introduces large computational overhead, since the methods used are symbolic. however for the $\lambda = 4$ case we can apply the implicit function theorem code, and reduce our investigations to $h : \mathbb{R}^2 \to \mathbb{R}$. First we regularise and obtain $\tilde{f}$, which has Jacobian:

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{10.3}$$

we are able to compute a Taylor expansion of $\phi : \mathbb{R}^2 \to \mathbb{R}^6$ to cubic order. Then we substitute this ϕ expansion into the cubic Taylor expansions of the last component of $\tilde{f}$. From 10.3, this last component corresponds to the degenerate image direction. Thus obtaining a cubic Taylor expansion for h, finding that:

$$h(\lambda, \theta_0) = \frac{12}{343}\,\lambda^2\theta_0 + \frac{16}{21}\,{\theta_0}^3 - \frac{8}{7}\,\lambda\theta_0 \tag{10.4}$$

Thus confirming that $\lambda = 4$ is indeed a bifurcation point, applying corollary 9.5.2.

Solving h and then using ϕ and A to map the solutions to solutions of f, we can visualise this bifurcation:

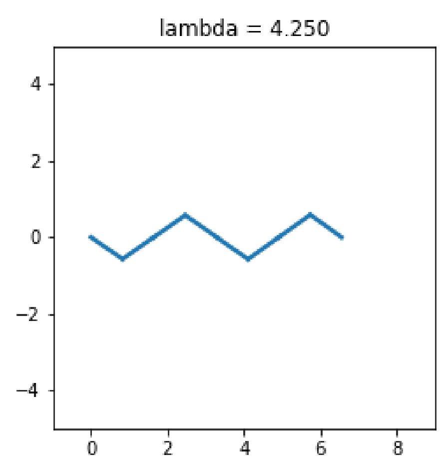


Figure 10.6: Approximate non-trivial solution in the case $n = 7$

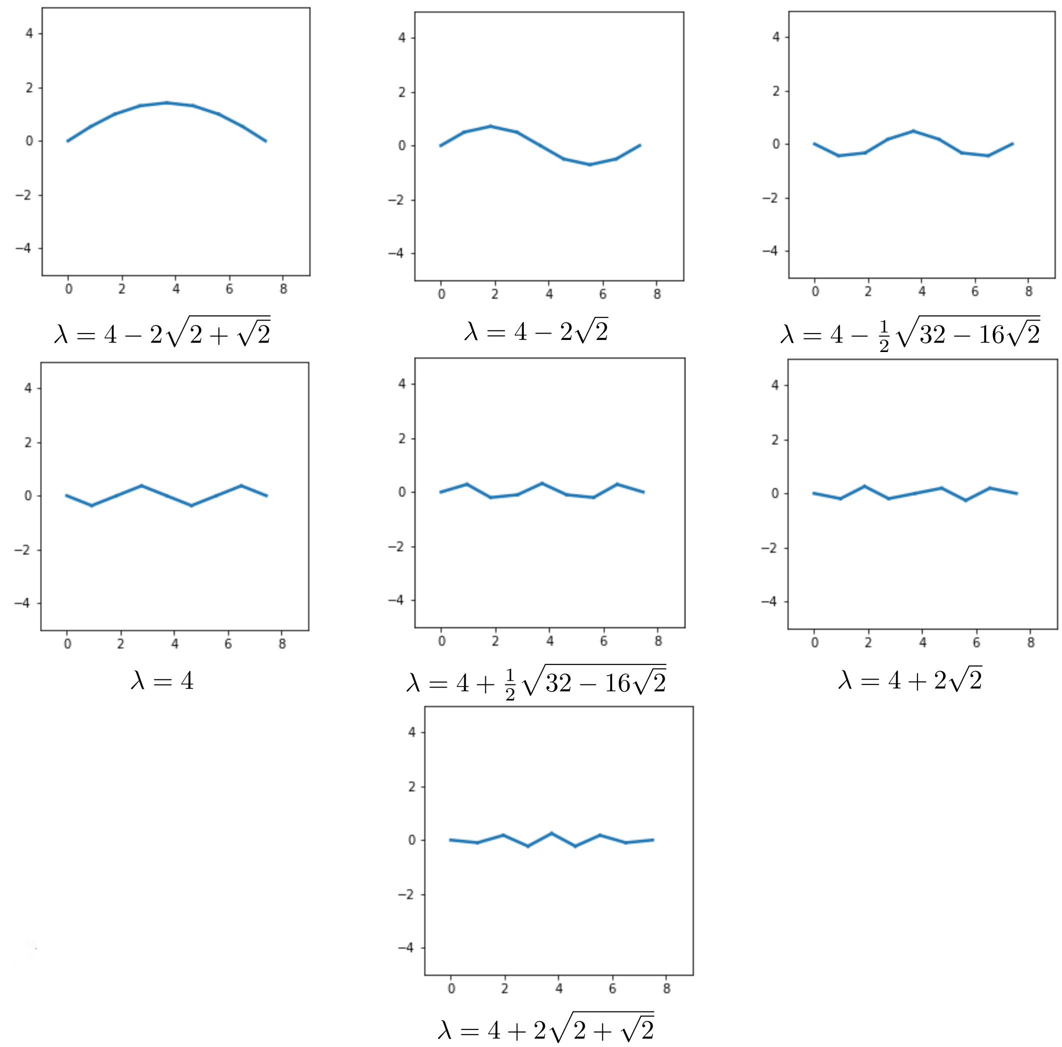


Figure 10.7: Potential bifurcations, visualisations of vectors ξ_0, for which $J\xi_0 = 0$, for a fixed λ, arranged in order of increasing λ

10.4 Conclusion

We see that using the methods introduced in chapter 9, we are able to investigate discretised problems and replicate the conclusions found from the analysis in function spaces.

Despite computational cost being high, when the symbolic expressions are simple, the code can handle up to 7 dimensional functions, reducing them to much simpler problems. Allowing us to apply the Crandall-Rabinowitz theorem and rigorously prove a bifurcation point has been located.

Chapter 11

Imperfect Bifurcations - Example

In this chapter we investigate problems that have so called "imperfect bifurcations". Now supposing we can reduce the problem to a quasi-h function:

$$h : \mathbb{R}^3 \to \mathbb{R}$$

Which has solutions generally lying on a 2-dimensional manifold in $\mathbb{R}^3$. We consider "slices" of this solution set which look like traditional bifurcations, then investigate the solution set structure in the full neighbourhood. The theory of imperfect bifurcations is studied in depth in the paper "Imperfect transcritical and pitchfork bifurcations" Ping Liu 2007. In this section we upgrade the model for the elastic rod that led to the boundary value problem 7.3, 7.4; to include the effects of gravity.

We now consider the equation:

$$\phi''(x) + \lambda \sin(\phi(x)) + g\rho(x)\cos(\phi(x)) = 0 \tag{11.1}$$

With the same boundary value conditions:

$$\phi'(0) = \phi'(L) = 0 \tag{11.2}$$

Where $\rho(x)$ is a density function such that $\int_0^L \rho(x)\phi_0 > 0$

Remark

We pay little attention to the physical veracity of this equation, since we are mainly interested in the bifurcation theory. The cosine term broadly translates to "the acceleration of the angle of the rod at a point is proportional to the horizontal component of the angle", i.e when $\phi(x) = 0$, the rod is flat at x, and gravity induces a downwards acceleration of the angle, whereas when $\phi(x) = \frac{\pi}{2}$, gravity does affect the acceleration. In the next chapter we will see a similar turn appears in a discretised version.

We will see that the imperfection induced by $\rho(x)$ is essential for later analysis. This problem is a slight variation on that studied in the Ping Liu 2007 paper.

Like Ping Liu 2007 we keep the same boundary value conditions.

11.1 The New Problem

We now have a problem of the form:

$$F : X \times \mathbb{R} \times \mathbb{R} \to Y$$

Where we recall that:

$$\begin{aligned} X &= \{\phi \in C^2([0,L]) \ \mid \ \phi'(0) = \phi'(L) = 0\} \\ Y &= C([0,L]) \end{aligned}$$

When we fix $g = 0$, the problem reduces to as discussed in chapter 7. However for $g \neq 0$, we see that $\phi = 0$ is no longer even a trivial solution to 11.1. The results in Ping Liu 2007 provide a method to characterise the solution set around a bifurcation point of the $g = 0$ problem, allowing for small changes in all directions: X, λ and g. We will see that these points are imperfect bifurcations, which often arise when a breaking of symmetry occurs, in this case, the vertical symmetry is violated by introducing a directed gravitational force.

We consider the structure of the solution set around the point:

$$(0, \lambda_0, g_0) = \left(0, \left(\frac{\pi}{L}\right)^2, 0\right)$$

Which we know is bifurcation in the slice of solutions where $g = 0$, it can be further shown that this is a "pitchfork" bifurcation, we will aim to show that the solution set behaves as seen in 11.1. Breaking into two components, one monotone and the other with a turning point.

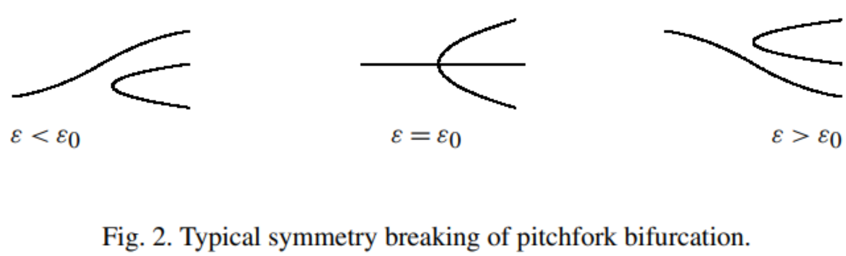


Figure 11.1: Taken from Ping Liu 2007, in our example $\epsilon = g$, $\epsilon_0 = 0$

11.2 Preliminary Computations

Before we begin applying the theorems from Ping Liu 2007 we first establish some derivatives we will require. It is helpful to recall the initial terms in the Taylor expansions of the sine and cosine function at zero:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \qquad \cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots \tag{11.3}$$

Some derivatives we have already computed, or as F1a - are trivial. Shorthand is introduced, care being taken to distinguish when we fix the value of ϕ so that we can use the Taylor expansion of sine from 11.3.

$$\begin{aligned}
F &= F(\phi, \lambda, g) = \phi'' + \lambda \sin(\phi) + g\rho(x)\cos(\phi) \\
F_X &= D_X F_{(0,\lambda,g)}(\phi) = \phi'' + \lambda\phi \\
F_\lambda &= D_\lambda F_{(\phi,\lambda,g)}(\lambda) = \sin(\phi) \\
F_{\lambda X} &= D^2_{\lambda,X} F_{(0,\lambda,g)}(\lambda, \phi) = \phi \\
F_g &= D_g F_{(\phi,\lambda,g)}(1) = \rho(x)\cos(\phi)
\end{aligned}$$

We now just use subscript g, λ, instead of $\mathbb{F}$ or $\mathbb{R}$, since there are two scalar dimensions now.

We also require expressions for the second and third derivatives with respect to X of F. First note that since ϕ'' is a linear map, it has zero second and higher derivatives, much as a matrix multiplication is zero in the second derivative, for the sine and cosine terms we can refer to their Taylor expansions, recalling that derivatives must be symmetric maps.

$$\begin{aligned}
F_{XX} &= D^2_{X,X} F_{(0,\lambda,g)}(\phi, \psi) = -\frac{g\rho(x)}{2!}\phi\psi \\
F_{XXX} &= D^3_{X,X,X} F_{(0,\lambda,g)}(\phi, \psi, \omega) = -\frac{\lambda}{3!}\phi\psi\omega
\end{aligned}$$

11.3 Conditions

We now state some conditions that will be required, using the naming conventions from Ping Liu 2007

$$\dim\ker D_X F_{(0,\lambda_0,g)} = \operatorname{Codim}\operatorname{Im} D_X F_{(0,\lambda_0,g)} = 1 \tag{F1a}$$

$$\ker D_X F_{(0,\lambda_0,g)} = \langle \phi_0 \rangle \tag{F1b}$$

$$D_\lambda F_{(0,\lambda_0,g)} \in \operatorname{Im} D_X F_{(0,\lambda_0,g)} \tag{F2'}$$

$$D^2_{\lambda,X} F_{(0,\lambda_0,g)}(1, \phi_0) \notin \operatorname{Im} D_X F_{(0,\lambda_0,g)} \tag{F3}$$

$$D^2_{X,X} F_{(0,\lambda_0,g_0)}(\phi_0, \phi_0) \in \operatorname{Im} D_X F_{(0,\lambda_0,g_0)} \tag{F4'}$$

$$D_g F_{(0,\lambda_0,g_0)}(1) \notin \operatorname{Im} D_X F_{(0,\lambda_0,g_0)} \tag{F5}$$

In the paper they also define a function ψ_2, which is used in another condition:

$$D^3_{X,X,X} F_{(0,\lambda_0,g_0)}(\phi_0, \phi_0, \phi_0) + 3D^2_{X,X\,(0,\lambda_0,g_0)}(\psi_2, \phi_0) \notin \operatorname{Im} D_X F_{(0,\lambda_0,g_0)} \tag{F6}$$

We will see that in this example the definition of ψ_2 is irrelevant.

11.4 Theorem

We now state a theorem that can be used to verify the local solutions of a function behave as in figure 11.1, it combines theorems 3.1, 3.2 and 4.1 from Ping Liu 2007. In our statement we assume that the known functional solution u_0 is the zero function.

Theorem

Given $F \in C^3(X \times \mathbb{R} \times \mathbb{R}, Y)$; points $\lambda_0, g_0 \in \mathbb{R}$ such that $f(0, \lambda_0, g_0) = 0$, and a function $\phi_0 \in X$ such that $D_X F_{(0,\lambda_0,g_0)}(\phi_0) = 0$,

such that $F(0, \lambda, g_0) = 0$ for $|\lambda - \lambda_0| < \delta$, for some $\delta > 0$,

then provided that (F1), (F2'), (F3), (F4'), (F5) hold, and also that:

$$D^2_{\lambda,X} F_{(0,\lambda_0,g_0)}(1, \phi_0) + D^2_{X,X} F_{(0,\lambda_0,g_0)}(v_1, \phi_0) \notin \operatorname{Im} D_X F_{(0,\lambda_0,g_0)} \tag{11.4}$$

where v_1 is the solution of:

$$D_X F_{(0,\lambda_0,g_0)}(v_1) + D_\lambda F_{(0,\lambda_0,g_0)}(1) = 0 \tag{11.5}$$

then taking the unique $\psi_2 \in X_3$ such that:

$$D_X F_{(0,\lambda_0,g_0)}(\psi_2) + D^2_{X,X} F_{(0,\lambda_0,g_0)}(\phi_0, \phi_0) = 0 \tag{11.6}$$

where $X = \langle \phi_0 \rangle \oplus X_3$

if we have that (F6) holds, then the solution set around $(0, \lambda_0, g_0)$ takes the form of an imperfect pitchfork bifurcation.

This means up to symmetry in the λ and g directions, we have that for a neighbourhood N of $(0, \lambda_0, g_0)$:

(A) For the $g = g_0$ slice there is a pitchfork bifurcation:

$$F^{-1}(0) \cap N = \{(0, \lambda, g_0) : |\lambda - \lambda_0| \leq \delta\} \cup \Sigma_0 \qquad \Sigma_0 = \left\{ \left(\bar{x}(t), \bar{\lambda}(t), 0\right), |t| \leq \eta \right\} \tag{11.7}$$

where $\bar{\lambda}(0) = \lambda_0,\ \bar{\lambda}'(0) = 0,\ \bar{\lambda}'' > 0$

(B) For fixed $g \in (g_0 - \rho_1, g_0) \cup (g_0, g_0 + \rho_1),\ \rho_1 > 0$ the slice is:

$$F^{-1}(0) \cap N = \Sigma_g^+ \cup \Sigma_g^- \tag{11.8}$$

where $\Sigma_g^+ = \left\{ \left(\bar{x}_+(t), \bar{\lambda}_+(t), 0\right), |t| \leq \right\},\ \bar{\lambda}'_+(0) = 0,\ \bar{\lambda}''_+ > 0$

and Σ_g^- is a monotone curve.

Proof

Omitted. See Ping Liu 2007.

Notes

In the full theorem statement of Ping Liu 2007 there are also constraints on the signs of three inner products, these merely fix an orientation, so we omit their discussion. The last two inner products' signs determine the direction of the fork when $g = 0$, see expression (4.3) in the proof in Ping Liu 2007. The first ensures that we can arrange for the bifurcation diagram to look like 11.1, as opposed to transposing the first and third diagram. This is discussed in Shi 1999.

Ping Liu 2007 also concerns whether solutions are degenerate or not, finding that $\left(\bar{x}_+(t), \bar{\lambda}_+(t), 0\right)$ is the unique degenerate solution on Σ_g^+,and that Σ_g^- has no degenerate solutions. We do not discuss degeneracy, but results on it can be found in Ping Liu 2007 and Shi 1999.

11.5 Applying to the Problem

We now apply theorem 11.4 to problem 11.1, $F(\phi, \lambda, g) = \phi''(x) + \lambda \sin(\phi(x)) + g\cos(\phi(x))$ at the point $(0, \lambda_0, g_0) = \left(0, \left(\frac{\pi}{L}\right)^2, 0\right)$, where $\phi_0 = \cos\left(\frac{\pi x}{L}\right)$. First verifying conditions (F1), (F2'), (F3), (F4'), (F5). We use the preliminary computations from 11.2.

Recall:

$$\dim\ker D_X F_{(0,\lambda_0,g)} = \operatorname{Codim} \operatorname{Im} D_X F_{(0,\lambda_0,g)} = 1 \tag{F1a}$$
$$\ker D_X F_{(0,\lambda_0,g)} = \langle \phi_0 \rangle \tag{F1b}$$
$$D_\lambda F_{(0,\lambda_0,g)} \in \operatorname{Im} D_X F_{(0,\lambda_0,g)} \tag{F2'}$$
$$D^2_{\lambda,X} F_{(0,\lambda_0,g)}(1, \phi_0) \notin \operatorname{Im} D_X F_{(0,\lambda_0,g)} \tag{F3}$$
$$D^2_{X,X} F_{(0,\lambda_0,g_0)}(\phi_0, \phi_0) \in \operatorname{Im} D_X F_{(0,\lambda_0,g_0)} \tag{F4'}$$
$$D_g F_{(0,\lambda_0,g_0)}(1) \notin \operatorname{Im} D_X F_{(0,\lambda_0,g_0)} \tag{F5}$$

Since $D_X F_{(0,\lambda_0,g)} = \phi'' + \lambda\phi$ is the same as from chapter 7, we see that (F1a) and (F1b) hold. Similarly (F3) is just transversality, which is also true since $D^2_{\lambda,X} F_{(0,\lambda_0,g)}(1, \phi_0) = \phi$ is the same as from chapter 7 and thus the same arguments can be applied.

For (F2') we see that $D_\lambda F_{(0,\lambda_0,g)}(1) = \sin(0) = 0$ and thus trivially lies in the image of $D_X F_{(0,\lambda_0,g)}$.

For (F4') we see that $D^2_{X,X} F_{(0,\lambda_0,g_0)}(\phi_0, \phi_0) = -\frac{g_0\rho(x)}{2!}\phi_0^2 = -\frac{0}{2!}\phi_0^2 = 0$ which also trivially lies in the image of $D_X F_{(0,\lambda_0,g)}$.

For (F5) see that $D_g F_{(0,\lambda_0,g_0)}(1) = \rho(x)\cos(0) = \rho(x)$, validating the necessity of the introduction of an imperfection, since if the rod had uniform density, we would be able to solve the boundary value problem $D_X F_{(0,\lambda_0,g)} = \phi'' + \lambda\phi = 1$, and thus (F5) would be violated. However since we specified that $\int_0^L \rho(x)\phi_0 > 0$ we can use the same argument from 7.10 to conclude that (F5) holds.

We now consider 11.4, observing that regardless of the function v_1, we have that:

$$D^2_{X,X} F_{(0,\lambda_0,g_0)}(v_1, \phi_0) = \frac{g_0 \rho(x)}{2!} v_1 \phi_0 = 0 \tag{11.9}$$

since $g_0 = 0$, thus the condition 11.4 reduces to the transversality requirement for $g = g_0$, which we know holds by (F3)

So all that remains to check is (F6):

$$D^3_{X,X,X} F_{(0,\lambda_0,g_0)}(\phi_0, \phi_0, \phi_0) + 3D^2_{X,X\,(0,\lambda_0,g_0)}(\psi_2, \phi_0) \notin \operatorname{Im} D_X F_{(0,\lambda_0,g_0)} \tag{F6}$$

again as for 11.4, the second term in the sum is zero, so we see that independant of ψ_2 this reduces to showing that:

$$D^3_{X,X,X} F_{(0,\lambda_0,g_0)}(\phi_0, \phi_0, \phi_0) \notin \operatorname{Im} D_X F_{(0,\lambda_0,g_0)} \tag{11.10}$$

Where $D^3_{X,X,X} F_{(0,\lambda_0,g_0)}(\phi_0, \phi_0, \phi_0) = -\frac{\lambda_0}{3!}\phi_0^3$ Which if we check the necessary condition 7.10

$$\int_0^L -\frac{\lambda_0}{3!}\phi_0^3(s)\phi_0(s)ds = \int_0^L -\frac{\lambda_0}{3!}\phi_0^4(s) < 0 \tag{11.11}$$

thus since $\lambda_0 \neq 0$, $D^3_{X,X,X} F_{(0,\lambda_0,g_0)}(\phi_0, \phi_0, \phi_0)$ cannot lie in the image of $D_X F_{(0,\lambda_0,g_0)}$, and thus (F6) holds.

Therefore we may conclude that the theorem applies and the solutions set structure around $\left(0, \left(\frac{\pi}{L}\right)^2, 0\right)$ has the qualitative structure seen in 9.2, up to orientation.

Chapter 12

Imperfect Bifurcations - Computational Example

We now return to the discretised problem, now incorporating the effect of gravity, to see if the solutions set structure investigated in chapter 11 can be replicated. We retain the geometry of 7.1, but update the expression E from before, recalling that in the model from 10 we had:

$$E_{old} = -H\lambda + \sum_{i=1}^{n} e_i$$

We now suppose that each spring in the structure has weight 1, and add the change in gravitational potential energy of each spring. We see that the j-th spring has vertical displacement V_j where:

$$V_j = \sum_{i=0}^{j} \sin\theta_i$$

For $j = 1, \ldots, n$. Thus we find a new expression:

$$E(\lambda, g, \theta_0, \ldots, \theta_{n-1}) = -H\lambda + \sum_{i=1}^{n} gV_i + \sum_{i=1}^{n} e_i$$

We have a positive sign for the gravity terms since a positive V_j value is directed against the direction of gravity. Recall that:

$$\sin\theta_n = -\sum_{i=0}^{n-1} \sin\theta_i$$

and so we only need to consider n parameters: $\theta_0, \ldots, \theta_{n-1}$, after making the appropriate substitutions. We then define our new function f:

$$f : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

$$f_i(\lambda, g, \theta_0, \ldots, \theta_{n-1}) = \frac{d}{d\theta_{i-1}} E(\lambda, g, \theta_0, \ldots, \theta_{n-1})$$

A function was written in SageMath to produce such f, when n is specified. We can show that in the $g = 0$ case it the same as in chapter 10, and thus has a line of trivial solutions.

Remark

We no longer insist on an imperfect weight distribution of the rod, thus we would struggle to prove rigorously that the solution set structure behaves a certain way. However computational methods will allow us to explore what impact introducing gravity has to each bifurcation point, while preserving symmetry of weight distribution.

12.1 Applying the Implicit Function Theorem

Given a point λ_0 that is a bifurcation for the reduced $g = 0$ problem, we let $(\lambda_0, 0)$ denote it's embedding into the solution space of the new problem. We know at $(\lambda_0, 0)$, the derivative of f with respect to $\mathbb{R}^n$ is the same as for for the reduced $g = 0$ problem, and thus non-invertible, with one dimensional kernel. But what about the derivatives with respect to g and λ?

Since we still have the line of trivial solutions for $g = \theta_0 = \cdots = \theta_{n-1} = 0$, we know that the derivative with respect to λ is zero, however we can make no such deductions about the derivative with respect to g.

Thankfully the implicit function theorem allows us to simply include g as another dimension in the non-invertible part of the derivative.

We can produce a regularised function $\tilde{f}$ as before, using the method from 9.3 where:

$$\tilde{f}(\lambda, g, x) = B \circ f(\lambda, g, \cdot) \circ A(x)$$

Such that the derivative with respect to x has block form as in 9.3.1, here x denotes the variables $\theta_0, \ldots, \theta_{n-1}$ from the $\mathbb{R}^n$ dimensions. We may then write:

$$\tilde{f} : \mathbb{R}_\lambda \times \mathbb{R}_g \times \mathbb{R}_{\xi_0} \times \mathbb{R}^{n-1} \to \mathbb{R}^n$$

We can now apply the implicit function theorem to force the values of the inputs for the $\mathbb{R}^{n-1}$ dimensions given the others, solving the first $n-1$ codomain dimensions. The only dimension that remains to be solved is the degenerate image dimensions, thus after making substitutions as before, we reduce the problem to that of solving:

$$h : \mathbb{R}_\lambda \times \mathbb{R}_g \times \mathbb{R}_{\xi_0} \to \mathbb{R}$$

The computations for these examples can be found in the notebooks ElasticRodGravity2.ipynb, ElasticRodGravity6.ipynb (Holmes 2022) for $\lambda = 2$, $\lambda = 6$ respectively.

12.2 Example $n = 2$

In this case we find that:

$$f_1(\lambda, g, \theta_0, \theta_1) = -\left(\frac{(\sin(\theta_0) + \sin(\theta_1))\cos(\theta_0)}{\sqrt{-(\sin(\theta_0) + \sin(\theta_1))^2 + 1}} + \sin(\theta_0)\right)\lambda + 2\,g\cos(\theta_0)$$
$$+ \frac{2\,(\theta_1 - \arcsin(-\sin(\theta_0) - \sin(\theta_1)))\cos(\theta_0)}{\sqrt{-(\sin(\theta_0) + \sin(\theta_1))^2 + 1}} + 2\,\theta_0 - 2\,\theta_1$$

$$f_2(\lambda, g, \theta_0, \theta_1) = -\left(\frac{(\sin(\theta_0) + \sin(\theta_1))\cos(\theta_1)}{\sqrt{-(\sin(\theta_0) + \sin(\theta_1))^2 + 1}} + \sin(\theta_1)\right)\lambda$$
$$+ 2\,(\theta_1 - \arcsin(-\sin(\theta_0) - \sin(\theta_1)))\left(\frac{\cos(\theta_1)}{\sqrt{-(\sin(\theta_0) + \sin(\theta_1))^2 + 1}} + 1\right)$$
$$+ g\cos(\theta_1) - 2\,\theta_0 + 2\,\theta_1$$

The presence of the $g\cos(\theta_i)$ terms helps motivate the cos term introduced in 11.1.

We recall the example from 10.2, where we found bifurcation points at $\lambda = 2$ and $\lambda = 6$. In the new model these correspond to points at $(\lambda, g) = (2, 0)$, and $(\lambda, g) = (6, 0)$.

We will investigate both points, recalling the geometry they correspond to in figures 12.1, 12.2

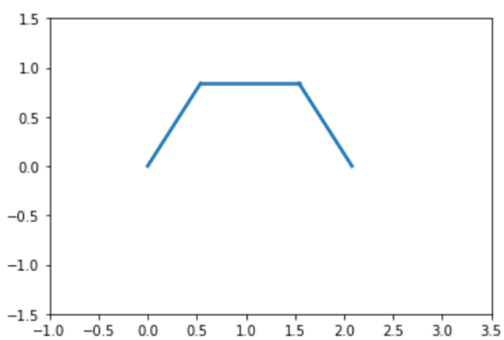


Figure 12.1: $\lambda = 2$ bifurcation when $g = 0$

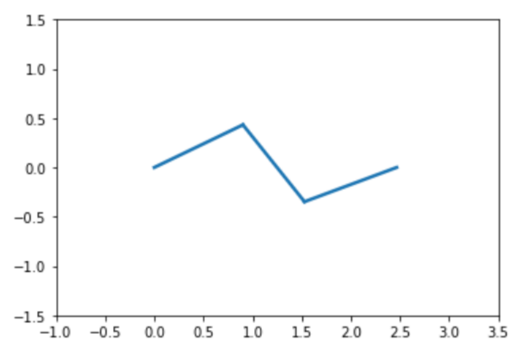


Figure 12.2: $\lambda = 6$ bifurcation when $g = 0$

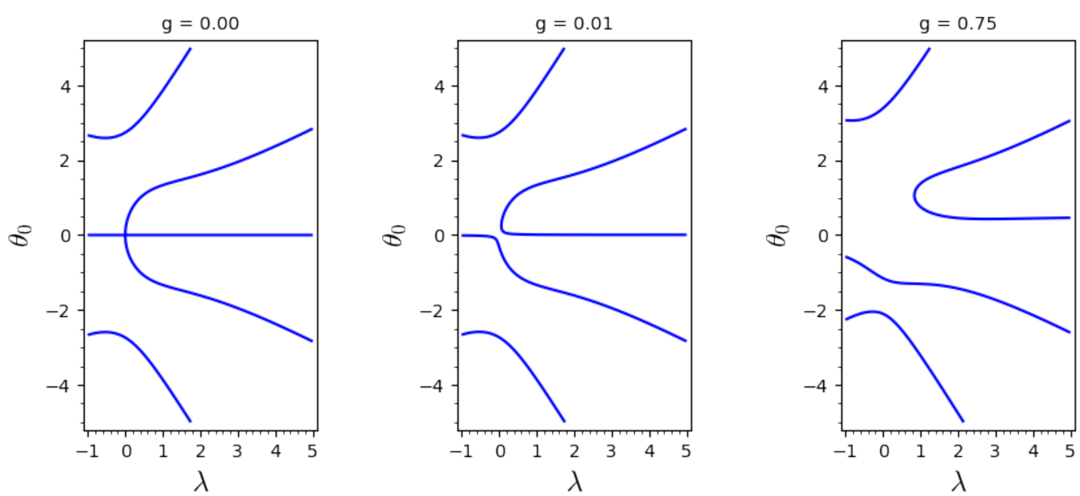
12.2.1 Computing h

Following the process outlined in 12.1, after converting to $\tilde{f}$, we compute the functions $h : \mathbb{R}^3 \to \mathbb{R}$, where the ξ_0 variable is called θ_0, up to quintic order, finding that:

$$\begin{aligned}
h_{\lambda=2} = {} & \frac{529}{699840}\, g^5 - \frac{173}{5832}\, g^3\lambda^2 + \frac{4}{81}\, g\lambda^4 - \frac{55}{5184}\, g^4\theta_0 + \frac{121}{648}\, g^2\lambda^2\theta_0 - \frac{2}{27}\, \lambda^4\theta_0 \\
& + \frac{115}{1944}\, g^3{\theta_0}^2 - \frac{43}{108}\, g\lambda^2{\theta_0}^2 - \frac{127}{648}\, g^2{\theta_0}^3 + \frac{7}{27}\, \lambda^2{\theta_0}^3 + \frac{19}{72}\, g{\theta_0}^4 - \frac{4}{45}\, {\theta_0}^5 \\
& - \frac{71}{3888}\, g^3\lambda + \frac{2}{27}\, g\lambda^3 + \frac{31}{216}\, g^2\lambda\theta_0 - \frac{1}{9}\, \lambda^3\theta_0 - \frac{13}{36}\, g\lambda{\theta_0}^2 + \frac{5}{9}\, \lambda{\theta_0}^3 \\
& - \frac{1}{162}\, g^3 + \frac{1}{9}\, g\lambda^2 + \frac{1}{12}\, g^2\theta_0 - \frac{1}{6}\, \lambda^2\theta_0 - \frac{7}{6}\, g{\theta_0}^2 + \frac{2}{3}\, {\theta_0}^3 + \frac{1}{6}\, g\lambda - 2\,\lambda\theta_0 + 2\,g \\
h_{\lambda=6} = {} & -\frac{11267}{75000000}\, g^5 + \frac{20507}{3125000}\, g^3\lambda^2 - \frac{4116}{390625}\, g\lambda^4 - \frac{106607}{15000000}\, g^4\theta_0 + \frac{11539}{156250}\, g^2\lambda^2\theta_0 \\
& + \frac{882}{78125}\, \lambda^4\theta_0 + \frac{10411}{125000}\, g^3{\theta_0}^2 - \frac{33531}{62500}\, g\lambda^2{\theta_0}^2 + \frac{37073}{75000}\, g^2{\theta_0}^3 + \frac{57}{125}\, \lambda^2{\theta_0}^3 - \frac{19683}{5000}\, g{\theta_0}^4 \\
& + \frac{9}{125}\, {\theta_0}^5 - \frac{1207}{250000}\, g^3\lambda + \frac{294}{15625}\, g\lambda^3 - \frac{1467}{25000}\, g^2\lambda\theta_0 - \frac{63}{3125}\, \lambda^3\theta_0 + \frac{1551}{2500}\, g\lambda{\theta_0}^2 \\
& + \frac{61}{125}\, \lambda{\theta_0}^3 + \frac{1}{1250}\, g^3 - \frac{21}{625}\, g\lambda^2 + \frac{1}{500}\, g^2\theta_0 + \frac{9}{250}\, \lambda^2\theta_0 - \frac{27}{50}\, g{\theta_0}^2 + \frac{18}{5}\, {\theta_0}^3 + \frac{3}{50}\, g\lambda - \frac{6}{5}\, \lambda\theta_0
\end{aligned}$$

When we set $g = 0$, these correspond to the expansions we saw 10.2.3. We also see that when $\lambda = 6$, there is no g term at the end, whereas there is when $\lambda = 2$, indicating some potential qualitative difference in the solution sets.

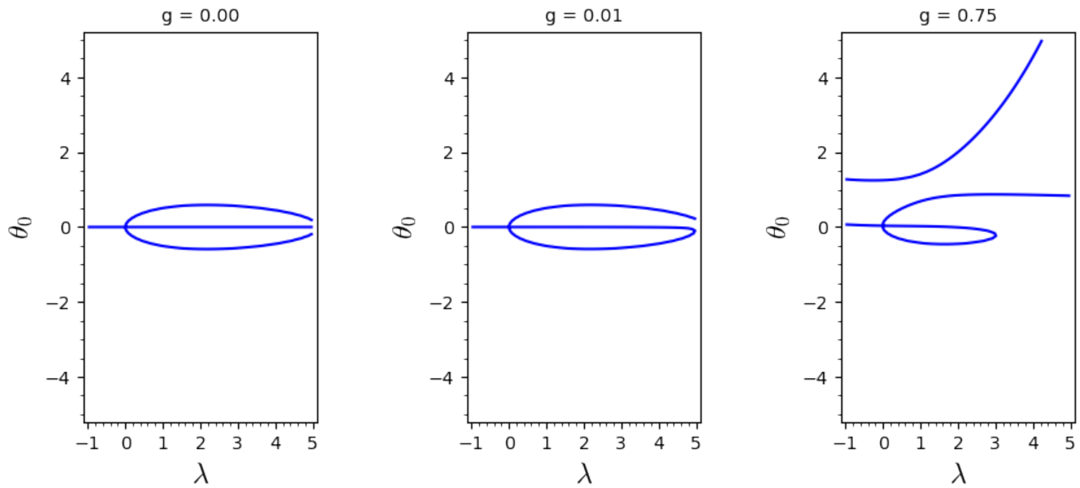
We can plot the solution set of these polynomials for fixed g to see if they have the same structure as seen in theorem 11.4 - having a component with a turning point and another that is monotone. In the diagrams we the x-axis is shifted so the bifurcation point is at the origin. For now we suppose that the solutions of these polynomials approximate non-trivial solutions of the full function without justification.



pitchfork bifurcation — adding gravity breaks symmetry — behaves as in theorem 11.4

Figure 12.3: Bifurcation diagrams for fixed values of g, when $\lambda = 2$

We see that when $\lambda = 2$, ignoring the solutions in the left corners, the solution set behaves as in theorem 11.4. There is one component with a turning point, and another monotone component. The monotone component corresponds to the "sagging" under gravity solutions, whereas the other component contains solutions where the rod bends upwards. We may expect the lower portion of the turning solutions to be unstable, since they seems to be "locked", much as the trivial solutions behave when $g = 0$.



pitchfork bifurcation — adding gravity preserves gravity — potential novel behaviour

Figure 12.4: Bifurcation diagrams for fixed values of g, when $\lambda = 6$

However when $\lambda = 6$ we see that locally the bifurcation is preserved. This bending mode leaves the centre of mass of the rod invariant, since as one spring rises, the other falls. Here we see an example of how 11.4 can fail, since a level of symmetry seems to have been preserved. As the looping behaviour exhibited may be unique to the lower order Taylor approximations so we do not draw conclusions beyond the local behaviour here.

12.3 Conclusion

Here we see the utility of Taylor expansions when trying to explore local behaviour of solutions sets. Considering these finite dimensional polynomial problems helps us classify the structure in general - thanks to the power of the implicit function theorem we can focus only on the degenerate aspects of a problem.

By computing higher degree expansions we can extract more information about the structure of the manifold of solutions, and potentially build insights to lay the foundations for further results.

Appendix A

Appendix

A.1 Lemma

Using the notation from the Crandall-Rabinowitz theorem we have that:

$$D_{\langle \xi_0 \rangle} h_{(\lambda_0, 0)} = 0 \tag{A.1}$$

Proof

Holding $\lambda = \lambda_0$ fixed we can view h as:

$$h(\lambda_0, \cdot) : \langle \xi_0 \rangle \to Y \tag{A.2}$$

$$h(\lambda_0, \cdot)(\xi) = P[f(\lambda_0, \cdot) \circ (\xi + \phi(\lambda_0, \xi))] \tag{A.3}$$

So using that chain rule inside the square brackets $(A.3 : \langle \xi_0 \rangle \to X \to Y)$ we see that:

$$\begin{aligned} D_{\langle \xi_0 \rangle} h_{(\lambda_0, 0)} &= D_{\langle \xi_0 \rangle} h(\lambda_0, \cdot)_{\xi = 0} && \text{(A.4)} \\ &= P \cdot D_X f_{(\lambda_0, 0 + \phi(\lambda_0, 0))} \cdot D_{\langle \xi_0 \rangle} (\xi + \phi(\lambda_0, \xi))_{\xi = 0} && \text{(A.5)} \\ &= PL(I_{\langle \xi_0 \rangle} + D_{\langle \xi_0 \rangle} \phi_{(\lambda_0, 0)}) && \text{(A.6)} \end{aligned}$$

Finally observing that $\ker P = \operatorname{Im} L$ we yield:

$$D_{\langle \xi_0 \rangle} h_{(\lambda_0, 0)} = 0 \tag{A.7}$$

$\square$

A.2 Lemma

Using the notation of the Crandall-Rabinowitz theorem we have that

$$D_{\langle \xi_0 \rangle} \phi_{(\lambda_0, 0)} = 0 \tag{A.8}$$

Proof

Recall from the proof of Lyapunov-Schmidt that ϕ solves the projected f, for all λ, ξ:

$$(I - P) f(\lambda, 0 + \xi + \phi(\lambda, \xi)) = 0 \tag{A.9}$$

Since this is constant when we differentiate with respect to $\langle \xi_0 \rangle$ at $(\lambda_0, 0)$ we yield:

$$(I - P) D_{\langle \xi_0 \rangle} f_{(\lambda_0, 0)} (I_{\langle \xi_0 \rangle} + D_{\langle \xi_0 \rangle} \phi_{(\lambda_0, 0)}) = 0 \tag{A.10}$$

But see that $D_{\langle \xi_0 \rangle} f_{(\lambda_0, 0)} = L$, which $I_{\langle \xi_0 \rangle}$ maps into the kernel of. Thus A.10 reduces to:

$$(I - P) L D_{\langle \xi_0 \rangle} \phi_{(\lambda_0, 0)} = D_{\langle \xi_0 \rangle} \phi_{(\lambda_0, 0)} = 0 \tag{A.11}$$

But then we observe that $D_{\langle \xi_0 \rangle} \phi_{(\lambda_0, 0)} \in (L)(\langle \xi_0 \rangle, W$, where W is the complement of $\ker L$. Thus we conclude that

$$D_{\langle \xi_0 \rangle} \phi_{(\lambda_0, 0)} = 0 \tag{A.12}$$

□

A.3 Lemma

Using the notation and assumptions from the Crandall-Rabinowitz theorem we have that:

$$D^2_{\mathbb{F}, \langle \xi_0 \rangle} h_{\lambda_0, 0}(1, \xi_0) \neq 0 \tag{A.13}$$

Proof

We begin by considering:

$$\begin{aligned}
D^2_{\mathbb{F},\langle\xi_0\rangle} h_{\lambda_0,0}(1,\xi_0) &= D_{\mathbb{F}}[D_{\langle\xi_0\rangle} h_{(\lambda,0)}(\xi_0)]_{\lambda_0}(1) && \text{(A.14)}\\
&= D_{\mathbb{F}}[P\cdot[D_X f_{(\lambda,0+\phi(\lambda,0))}\cdot D_{\langle\xi_0\rangle}(\xi+\phi(\lambda,\xi))_{\xi=0}](\xi_0)]_{\lambda_0}(1) && \text{(A.15)}\\
&= PD_{\mathbb{F}}[[D_X f_{(\lambda,0)}\cdot D_{\langle\xi_0\rangle}(\xi+\phi(\lambda,\xi))_0](\xi_0)]_{\lambda_0}(1) && \text{(A.16)}\\
&= PD_{\mathbb{F}}[[D_X f_{(\lambda,0)}\cdot (I_{\langle\xi_0\rangle}+D_{\langle\xi_0\rangle}\phi_{(\lambda,0)})(\xi_0)]_{\lambda_0}(1) && \text{(A.17)}\\
&= PD_{\mathbb{F}}[[D_X f_{(\lambda,0)}\cdot (\xi_0+D_{\langle\xi_0\rangle}\phi_{(\lambda,0)}(\xi_0))]_{\lambda_0}(1) && \text{(A.18)}
\end{aligned}$$

Computing the derivative in A.18 is tricky, since the nature of relationship between the parts is somewhat obscure, let:

$$B:\mathbb{F}\to\mathcal{L}(X,Y) \qquad \text{(A.19)}$$

$$B(\lambda)=D_X f_{(\lambda,0)} \qquad \text{(A.20)}$$

$$g:\mathbb{F}\to W\subset X \qquad \text{(A.21)}$$

$$g(\lambda)=\xi_0+D_{\langle\xi_0\rangle}\phi_{(\lambda,0)}(\xi_0) \qquad \text{(A.22)}$$

Then we see the derivative in A.18 that we must compute is:

$$D_{\mathbb{F}}[B(\lambda)(g(\lambda))]_{\lambda_0}(1)=D_{\mathbb{F}}[(B(\lambda)\cdot g(\lambda)]_{\lambda_0}(1) \qquad \text{(A.23)}$$

Since evaluation of $B(\lambda)$ is merely matrix-vector multiplication, thus we can apply the product rule to yield:

$$\begin{aligned}
&D_{\mathbb{F}}(B(\lambda)\cdot g(\lambda)]_{\lambda_0}(1) && \text{(A.24)}\\
&\qquad = D_{\mathbb{F}}B_{\lambda_0}(1)\cdot g(\lambda_0)+B(\lambda_0)\cdot D_{\mathbb{F}}g_{\lambda_0}(1) && \text{(A.25)}
\end{aligned}$$

We compute:

$$D_{\mathbb{F}}B_{\lambda_0}=D_{\mathbb{F}}D_X f_{(\lambda_0,0)}=D^2_{\mathbb{F},X}f_{(\lambda_0,0)} \qquad \text{(A.26)}$$

and,

$$D_{\mathbb{F}}g_{\lambda_0}=D_{\mathbb{F}}[\xi_0+D_{\langle\xi_0\rangle}\phi_{(\lambda_0,0)}(\xi_0)]=D^2_{\mathbb{F},\langle\xi_0\rangle}\phi_{(\lambda_0,0)}(\cdot,\xi_0) \qquad \text{(A.27)}$$

And thus combining we conclude that:

$$\begin{aligned}
D^2_{\mathbb{F},\langle\xi_0\rangle}h_{\lambda_0,0}(1,\xi_0) &= PD^2_{\mathbb{F},X}f_{(\lambda_0,0)}(1,\xi_0+D_{\langle\xi_0\rangle}\phi_{(\lambda_0,0)}(\xi_0)) && \text{(A.28)}\\
&\quad + PD_X f_{(\lambda_0,0)}\cdot D^2_{\mathbb{F},\langle\xi_0\rangle}\phi_{(\lambda_0,0)}(1,\xi_0) && \text{(A.29)}
\end{aligned}$$

Then from A.2 we saw that:

$$D_{\langle \xi_0 \rangle} \phi_{(\lambda_0, 0)} = 0 \tag{A.30}$$

Also we observe that the second term is zero by the choice of kernel of P, we arrive at:

$$D^2_{\mathbb{F}, \langle \xi_0 \rangle} h_{\lambda_0, 0}(1, \xi_0) = P D^2_{\mathbb{F}, X} f_{(\lambda_0, 0)}(1, \xi_0) \neq 0 \tag{A.31}$$

Thus proving the result using the transversality assumption.

□

Appendix B

Examples

In this chapter we consider examples of the Crandall-Rabinowitz theorem.

We now apply the theory developed in chapter 5 to some examples, first where the solution set can be easily visualised.

B.0.1 Example 1 : $m = 1$

We start by considering an example in the case where $X = Y = \mathbb{R}$ and $\mathbb{F} = \mathbb{R}$:

$$f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

$$f(\lambda, x) = (x^2 + e^\lambda - 4)(x - \sin(\lambda))x$$

This function has closed form solutions:

$$[x = \pm\sqrt{-e^\lambda + 4},\ x = \sin(\lambda),\ x = 0]$$

With so few dimensions we may easily visualise the solution set of f, as seen in figure B.1.

Note that according to the formal definition of a bifurcation we only are interested in self intersections of the zero set that occur on the line $x = 0$.

The Necessary Condition

First we inspect the first derivative of f, since unless this is zero we stand no chance of finding a bifurcation point, seeing that:

$$D_X f_{(\lambda,0)} = \partial_x f(0, \lambda) = -(e^\lambda - 4)\sin(\lambda)$$

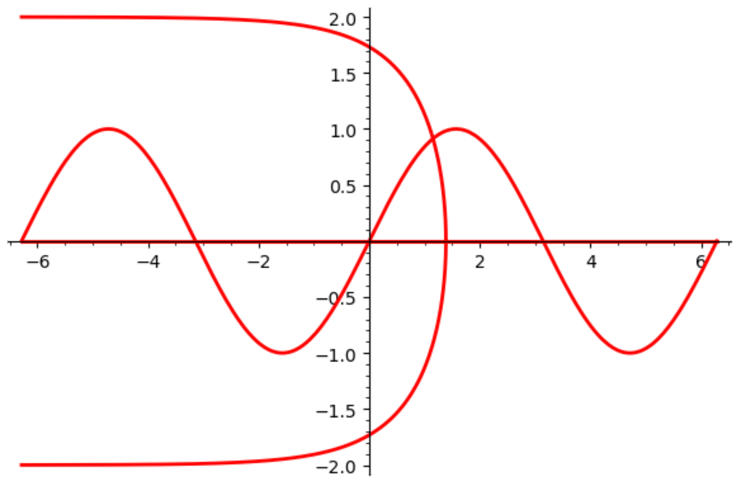


Figure B.1: Solution set of f

Where non-invertibility now merely equates to this being zero, thus we conclude that the only possible values from which the solution set can bifurcate are:

$$\lambda = \ln 4$$
$$\lambda = k\pi \quad k \in \mathbb{Z}$$

Observe that each time $L = \partial_x f(\lambda_0, 0)$ has non-trivial kernel, this means that L is the zero operator, thus we have that $\ker L = \mathbb{R}$ and $\text{range}(L) = \{0\}$. So we may take $\xi_0 = 1$ each time so that $\ker L = \langle \xi_0 \rangle = \langle 1 \rangle = \mathbb{R}$

Applying Crandall-Rabinowitz

Now the necessary condition has been checked we will try to apply the Crandall-Rabinowitz theorem, which hinges on the evaluation of the expression:

$$D^2_{\mathbb{F},X} f_{(\lambda_0,0)}(1, \xi_0) \tag{B.1}$$

In this case we have that $D^2_{\mathbb{F},X} f_{(\lambda_0,0)} \in \mathcal{L}(\mathbb{R}, \mathcal{L}(\mathbb{R}, \mathbb{R}))$, and thus like the first derivative collapsed to just being able to be viewed as a scalar. We have that:

$$D^2_{\mathbb{F},X} f_{(\lambda,x)}(1, \xi_0) = \partial_\lambda \partial_x f \tag{B.2}$$
$$= -2\,x^2 \cos(\lambda) - \left(x^2 + e^\lambda - 4\right) \cos(\lambda) + (x - \sin(\lambda))e^\lambda + xe^\lambda \tag{B.3}$$

We require that $D^2_{\mathbb{F},X} f_{(\lambda,x)}(1, \xi_0) \notin \operatorname{Im} L$ for (λ, x) to be a bifurcation. Since L is the zero operator at our points of consideration, this reduces to

$$D^2_{\mathbb{F},X} f_{(\lambda,x)}(1, \xi_0) \neq 0 \tag{B.4}$$

Lets first check $(\ln 4, 0)$:

$$D^2_{\mathbb{F},X} f_{(\ln 4,0)}(1, \xi_0) = -4 \sin(2 \log(2)) \neq 0$$

since $2\log 2$ is not a multiple of π. Thus $(\ln 4, 0)$ is verified to be a bifurcation point.

Next consider $(k\pi, 0)$:

$$\begin{aligned} D^2_{\mathbb{F},X} f_{(k\pi,0)}(1, \xi_0) &= -\left(e^{(\pi k)} - 4\right) \cos(\pi k) - e^{(\pi k)} \sin(\pi k) \\ &= -\left(e^{(\pi k)} - 4\right) \cos(\pi k) \\ &= (-1)^{k+1} \left(e^{(\pi k)} - 4\right) \\ &\neq 0 \qquad\qquad\qquad\qquad \forall k \in \mathbb{Z} \end{aligned}$$

And thus $(k\pi, 0)$ is a bifurcation point for all $k \in \mathbb{Z}$

All the points where degeneracy occurred are bifurcations, and combining the conclusions of the necessary condition and sufficient Crandall-Rabinowitz theorem, we have established the qualitative behaviour of the solution set about the entire trivial line of solutions.

It should be noted that this is not always possible in general, we may meet the necessary condition but fail the transversality condition.

B.0.2 Example 2 : m = 3

Now we consider a multidimensional example, which will require the full linear algebraic implementation of the Crandall-Rabinowitz theorem. First we construct the function we must solve:

$$f : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$$

Such functions are composed of 3 components:

$$f(\lambda, x, y, z) = \begin{bmatrix} f_1(\lambda, x, y, z) \\ f_2(\lambda, x, y, z) \\ f_3(\lambda, x, y, z) \end{bmatrix} \in \mathbb{R}^3 \tag{B.5}$$

In this example the functions f_1 and f_2 will be independant of the z variable, thus omitting this variable we may plot their individual solution sets:

First define f_1:

$$f_1(\lambda, x, y, z) = -y \cos\left(\frac{1}{3}\sqrt{\pi^2 + 9\,x^2 + 9\,y^2}\right) + x \sin\left(\frac{1}{3}\sqrt{\pi^2 + 9\,x^2 + 9\,y^2}\right)$$

This has a spiraling solution set, as pictured in figure B.2.

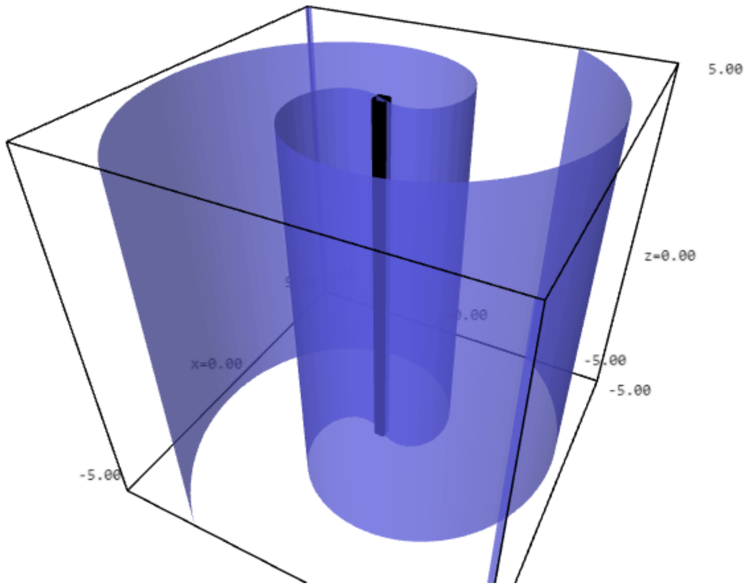


Figure B.2: Solution set of f_1, the black line is the trivial solutions

Next we define f_2:

$$f_2(\lambda, x, y, z) = x \cos(\lambda) + y \sin(\lambda) \tag{B.6}$$

Which has a corkscrew solution set, as seen in B.3

For a point (λ, x, y, z) to be a solution of f we must have that it solves both f_1 and f_2. This means it is in the intersection of these two solution manifolds. We see there are non-trivial such intersections in figure B.4.

Finally we define:

$$f_3(\lambda, x, y, z) = y^3 + x^2 + \lambda z + z^2 + 4\,z$$

It will turn out that this component plays no role in the bifurcation theory at the point we will be considering, and can be thought of as an analogue for the dimensions that get discarded in the general formulation.

Claimed Bifurcation Point

We now claim that $\lambda = -\frac{\pi}{6}$ is a bifurcation point from the line of trivial solutions of the function:

$$f(\lambda, x, y, z) = \begin{bmatrix} -y\cos\left(\frac{1}{3}\sqrt{\pi^2 + 9\,x^2 + 9\,y^2}\right) + x\sin\left(\frac{1}{3}\sqrt{\pi^2 + 9\,x^2 + 9\,y^2}\right) \\ x\cos(\lambda) + y\sin(\lambda) \\ y^3 + x^2 + \lambda z + z^2 + 4\,z \end{bmatrix}$$

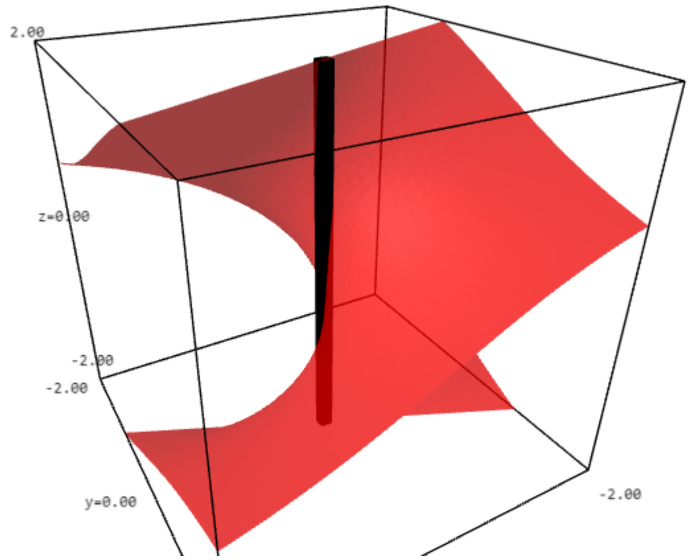


Figure B.3: Solution set of f_2, the black line is the trivial solutions

For brevity we will denote $(\lambda, 0, 0, 0) = (\lambda, 0)$, where 0 on the right hand side represents the zero vector in $\mathbb{R}^3$.

The Necessary Condition

First we must compute the first derivative of f at this point, recall that:

$$L = D_X f_{(-\frac{\pi}{6}, 0)} \in (\mathbb{R}^3, \mathbb{R}^3)$$

Where we may canonically identify L with the Jacobian matrix of f, evaluated at $(\lambda, 0)$, which evaluates to:

$$\begin{pmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 1 \\ \cos(\lambda) & \sin(\lambda) & 1 \\ 0 & 0 & \lambda + 4 \end{pmatrix} \tag{B.7}$$

Where we see at the point $(-\frac{\pi}{6}, 0)$ in question we have a dependency between the first two rows:

$$\begin{pmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 1 \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 1 \\ 0 & 0 & -\frac{1}{6}\pi + 4 \end{pmatrix} \tag{B.8}$$

We see that (now writing row vectors for conciseness):

$$\ker L = \langle \xi_0 \rangle = \langle \left(1, \sqrt{3}, 0\right) \rangle$$

Thus L has a non-trivial kernel and the necessary condition for $(-\frac{\pi}{6}, 0)$ to be a bifurcation point is met.

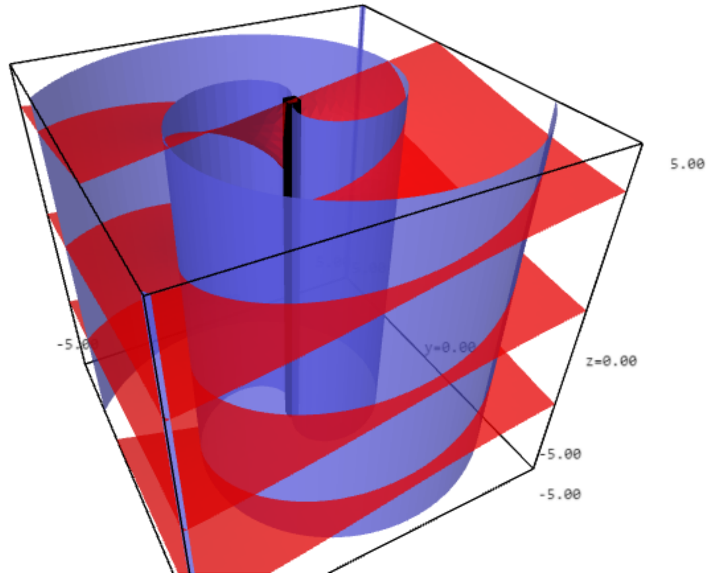


Figure B.4: Solutions to f must have λ, x, y components in the intersection

The Sufficient Condition

Now we aim to show that $(-\frac{\pi}{6}, 0)$ is indeed a bifurcation point, by verifying the sufficient condition from the Crandall-Rabinowitz theorem, we must evaluate:

$$D^2_{\mathbb{F},X} f_{(-\frac{\pi}{6},0)}(1, \xi_0)$$

In 5.5 we saw that this evaluation is equivalent to differentiating the expression:

$$D_X f_{(\lambda,0)}(\xi_0)$$

with respect to λ and evaluating at $(\lambda_0, 0) = (-\frac{\pi}{6}, 0)$

We have that:

$$D_X f_{(\lambda,0)}(\xi_0) = \begin{pmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 1 \\ \cos(\lambda) & \sin(\lambda) & 1 \\ 0 & 0 & \lambda + 4 \end{pmatrix} \begin{bmatrix} 1 \\ \sqrt{3} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \cos(\lambda) + \sqrt{3}\sin(\lambda) \\ 0 \end{bmatrix}$$

Differentiating with respect to λ:

$$D^2_{\mathbb{F},X} f_{(\lambda,0)}(1, \xi_0) = \Big(0, -\sin(\lambda) + \sqrt{3}\cos(\lambda), 0\Big) \tag{B.9}$$

So,

$$D^2_{\mathbb{F},X} f_{(-\frac{\pi}{6},0)}(1, \xi_0) = (0, 2, 0) \tag{B.10}$$

So we can confirm that $\lambda_0 = -\frac{\pi}{6}$ is a bifurcation point if we have that this vector does not lie in the range of L.

What is the range of L? We must compute the image of the matrix in B.8, which is found to be:

$$\operatorname{Im} L = \left\langle \begin{bmatrix} 1 \\ -\frac{1}{3}\sqrt{3} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle$$

So we see that the zeros in the first and third entry of B.10 ensure that:

$$D^2_{\mathbb{F},X} f_{(-\frac{\pi}{6},0)}(1, \xi_0) \notin \operatorname{Im} L$$

and thus the transversality condition is met and we conclude that $-\frac{\pi}{6}$ is indeed a bifurcation point of f.

Bibliography

Buffoni, Boris and John Toland (2003). *Analytic Theory of Global Bifurcation.* 41 William Street, Princeton, New Jersey 08540: Princeton University Press.

Poincare (1895). "Analysis Situs". In: 1.2, pp. 1–123.

Kielhöfer, Hansjörg (2012). *Bifurcation Theory.* 233 Spring Street, New York, NY 10013, USA: Springer.

The Sage Developers (2021). *SageMath, the Sage Mathematics Software System (Version 9.2).* `https://www.sagemath.org`.

Holmes, Matthew (Apr. 2022). *Computational Bifurcations.* Version 1.0.0. URL: `https://github.com/Matthew-Holmes/Bifurcations`.

Ping Liu Junping Shi, Yuwen Wang (2007). "Imperfect transcritical and pitchfork bifurcations". In: *Journal of Functional Analysis* 251.2, pp. 573–600.

Shi, Junping (1999). "Persistence and Bifurcation of Degenerate Solutions". In: *Journal of Functional Analysis* 169.1, pp. 494–531.