Generalized Lyapunov–Schmidt reduction for parametrized equations at near singular points

Arnold Neumaier

Institut für Mathematik, Universität Wien
Strudlhofgasse 4, A-1090 Wien, Austria
email: neum@cma.univie.ac.at

Abstract.
The Lyapunov-Schmidt reduction is generalized to the case of imperfect singularities. The results presented neither need very precise information about the location of the (near) singularities nor a precise knowledge of (near) null spaces.

Keywords: Lyapunov-Schmidt reduction, unfolding, imperfect singularities

1991 MSC Classification: 58C15

1 Introduction

Let X, Y be (finite- or infinite-dimensional) Banach spaces, let $F : \Lambda \times D \subseteq \mathbb{R}^p \times X \to Y$ be an $k \geq 1$ times continuously differentiable nonlinear mapping such that the partial derivative $F_u(\lambda, u)$ of F with respect to u is a Fredholm operator of index zero. We want to study the solution manifold of $F(\lambda, u) = 0$ in a neighbourhood of an approximate zero $(\lambda_0, u_0) \in \mathrm{int}(\Lambda \times D)$ of F.

In practice, one often is not interested in the manifold itself, but in a few key parameters $\mu_i(\lambda, u)$. Thus we assume that we are also given a vector valued function $\mu : \Lambda \times D \to \mathbb{R}^m$ defining k times continuously differentiable parameter functionals μ_i $(i = 1, \ldots, m)$, which we shall call the *unfolding functionals*, whose behaviour on the solution manifold is of primary interest.

In practice, the unfolding functionals are chosen such that the behaviour of the solution manifold is reflected by the way the μ_i depend on λ. In the most widely studied case $p = m = 1$, the plot of μ against λ is called the *bifurcation diagram*, and μ is chosen intuitively to make the bifurcation diagram a faithful description of the topology of the solution manifold.

The goal of the paper is to show that one can rigorously reduce the problem of describing the solution manifold, including its singular or near singular behaviour, to the solution of a low-dimensional problem $\xi(\lambda, \sigma) = 0$, where $\sigma = \mu(\lambda, u)$, with an appropriately constructed function ξ defined on a subset of $\mathbb{R}^p \times \mathbb{R}^m$. We give a constructively verifiable condition for the unfolding functionals and bounds on the residuals of the approximate solution that, together, guarantee the existence of ξ and of a diffeomorphism between the solution manifold for the original equation and that for the reduced equation.

This generalizes the well-known Lyapunov-Schmidt reduction technique (see e.g., Golubitzky & Schaeffer [5] or Chow & Hale [3]), which is concerned with the special case of a singular point of the manifold, where $F(\lambda_0, u_0) = 0$ and $\det F_u(\lambda_0, u_0) = 0$. In contrast to this classical method, however, our method also works when there is only an imperfect singularity, i.e., when $F_u(\lambda, u)$ becomes ill-conditioned on the manifold but is never singular, corresponding to the case when two solution branches come close but do not touch each other.

This is important in view of the Sard-Smale theorem [8, Theorem 4.18] that states that bifurcation points are unstable under perturbations in the sense that the set of perturbations f of norm one for which $F(\lambda, u) = \varepsilon f$ has no bifurcation point is open and dense in the unit ball of L^2. In contrast to the treatment of imperfect bifurcations in Golubitzky & Schaeffer [4, 5], our neighbourhoods need not be arbitrarily small (i.e., defining germs) but can be quite sizeable, since the assumptions of our main result, Theorem 3.1

contain no tiny quantities.

2 A fixed point formulation

Our reduction procedure is based on the assumption that we have linear mappings $A_0 : X \to Y$, $A_1 : \mathbb{R}^m \to Y$, and $A_2 : X \to \mathbb{R}^m$ such that A_0 approximates $F_u(\lambda_0, u_0)$ for some pair $(\lambda_0, u_0) \in \mathrm{int}(\Lambda \times D)$ with $F(\lambda_0, u_0) \approx 0$, and the linear operator $A : X \times \mathbb{R}^m \to Y \times \mathbb{R}^m$ defined by

$$A\begin{pmatrix} u \\ \xi \end{pmatrix} := \begin{pmatrix} A_0 u + A_1 \xi \\ A_2 u \end{pmatrix} \tag{1}$$

is a bijection bewteen the Banach spaces $X \times \mathbb{R}^m$ and $Y \times \mathbb{R}^m$ (with norms given by $\|\binom{u}{\xi}\| := \|u\| + \|\xi\|$), with a bounded inverse,

$$\|A^{-1}\| \leq \alpha. \tag{2}$$

To see the meaning of this construction, note that for an true or an imperfect bifurcation, the case of main interest, the Frechet derivative $F_u(\lambda_0, u_0)$ is almost singular in the sense that close to it there is a linear operator A_0 with

$$\dim Y/\, \mathrm{range}\, A_0 = \dim \ker A_0 = m_0 > 0. \tag{3}$$

(Here the assumption is used that $F_u(\lambda_0, u_0)$ is a Fredholm operator with index zero.)

2.1 Proposition. *Let $v_1, \ldots, v_{m_0}$ be linearly independent null vectors of A_0, and let $w_1, \ldots, w_{m_0}$ be linearly independent null vectors of the adjoint $A_0^* : Y^* \to X^*$.*

(i) If A is a bijection then $A_2 v_1, \ldots, A_2 v_{m_0}$ are linearly independent and $A_1^* w_1, \ldots, A_1^* w_{m_0}$ are linearly independent.

(ii) If (3) *holds with $m_0 = m$ then, conversely, these conditions imply that A is a bijection.*

Proof. (i) If $\sum \alpha_i A_2 v_i = 0$ then $v = \sum \alpha_i v_i$ satisfies $A\binom{v}{0} = 0$, and since A is a bijection, this forces $v = 0$. Then all $\alpha_i = 0$ since the v_i are linearly independent. Hence the $A_2 v_i$ are linearly dependent. Similarly if $\sum \alpha_i A_1^* w_i = 0$ then $w = \sum \alpha_i w_i$ satisfies $(w^*, 0)A\binom{u}{\xi} = 0$ for all $u \in X, \xi \in \mathbb{R}^n$, and since A is a bijection, this forces $w = 0$. As before, this implies that the $A_1^* w_i$ are linearly independent.

(ii) We need to show that under the stated conditions,

$$A_0 u + A_1 \xi = v, \tag{4}$$

$$A_2 u = \eta. \tag{5}$$

is uniquely solvable for $u \in X,\ \xi \in \mathbb{R}^n$. Now by the Fredholm alternative, $A_0 u = v - A_1 \xi$ is solvable iff $w^*(v - A_1\xi) = 0$ for all $w \in \ker A_0^*$. By assumption, this holds iff

$$w_i A_i^* \xi = w_i^* v \quad \text{for } i = 1, \ldots, m,$$

and this system is uniquely solvable for ξ. The solution u of (4) is then determined up to a vector $v = \sum \alpha_i v_i$ in $\ker A_0$, and by assumption, this vector is uniquely fixed by (5). $\square$ $\square$

If $v_1, \ldots, v_{m_0}$ and $w_1, \ldots, w_{m_0}$ span the invariant subspaces corresponding to the small eigenvalues of $F_u(\lambda_0, u_0)$ and its adjoint, the qualitative conclusion from Proposition 2.1 is that we should choose A_2 and A_1 such that $[A_2 v_1, \ldots, A_2 v_{m_0}]$ and $[A_1^* w_1, \ldots, A_1^* w_{m_0}]$ are $m \times m_0$ matrices of rank m_0, and the rank should be stable under perturbations of the size needed to move the small eigenvalues of $F_u(\lambda_0, u_0)$ to zero.

Thus, in order to make A a bijection we need to choose A_1 such that it extends the range of A_0 to Y, and A_2 such that it shrinks the null space of A_0 to 0. Clearly this can be achieved only if $m \geq m_0$, but then it holds for almost all choices for A_1 and A_2 (with exception of a set of measure zero), and of course this remains true also when A_0 is not exactly singular. In particular, to avoid a large α in (3), the number m of unfolding functionals must be (at least) the algebraic number of small eigenvalues of $F_u(\lambda_0, u_0)$.

Note that the size of the constant α in (2) depends of the closeness of A_1 and A_2 to the exceptional set, whence one should choose A_1 and A_2 far away from the set of exceptions. Since, in the following, $\mu_u(\lambda_0, u_0)$ needs to be approximately equal to A_2, the above discussion implies a condition for the unfolding functionals.

Practitioners usually know from experience which unfolding functionals give revealing bifurcation diagrams and often these satisfy this condition. Sometimes, however, an additional (or a different) unfolding functional must be chosen. This is the case, e.g., when $u_0 = 0$ is a trivial solution, $m_0 = 1$, and $\mu(\lambda, u) = \|u\|^2$, since then $A_2 = \mu_u(\lambda_0, u_0) = 0$ has not the required properties.

From now on, we assume that A is specified by an expression of the form (1) in such a way that (2) hold. For fixed values of $\lambda \in \mathbb{R}^p$ and $\sigma \in \mathbb{R}^m$, we consider the nonlinear mapping $\Phi_{\lambda,\sigma} : D \times \mathbb{R}^m \to X \times \mathbb{R}^m$ defined by

$$\Phi_{\lambda,\sigma}\begin{pmatrix} u \\ \xi \end{pmatrix} := \begin{pmatrix} u \\ \xi \end{pmatrix} - A^{-1}\begin{pmatrix} F(\lambda, u) + A_1\xi \\ \mu(\lambda, u) - \sigma \end{pmatrix}. \tag{6}$$

Clearly, any fixed point of $\Phi_{\lambda,\sigma}$ satisfies the equations

$$F(\lambda, u) + A_1\xi = 0, \quad \mu(\lambda, u) = \sigma, \tag{7}$$

and the solutions with $\xi = 0$ are the solutions of our original problem.

3 The reduction process

3.1 Theorem. *Let δ be a positive number such that the ball*

$$B(u_0, \delta) := \{u \in X \mid \ \|u - u_0\| \le \delta\}$$

is in $\mathrm{int}(D)$*, and for all $u \in B(u_0, \delta)$ we have*

$$\|F_u(\lambda, u) - A_0\| \le \frac{1}{4\alpha}, \quad \|\mu_u(\lambda, u) - A_2\| \le \frac{1}{4\alpha}. \tag{8}$$

If

$$\|F(\lambda, u_0) + A_1\xi_0\| \le \frac{\delta}{4\alpha}, \quad \|\sigma - \mu(\lambda, u_0)\| \le \frac{\delta}{4\alpha} \tag{9}$$

then $\Phi_{\lambda,\sigma}$ is a contraction in

$$B_\delta(u_0, \xi_0) := \left\{ \begin{pmatrix} u \\ \xi \end{pmatrix} \in X \times \mathbb{R}^m \mid \|u - u_0\| + \|\xi - \xi_0\| \le \delta \right\}.$$

Proof. Let $\binom{u}{\xi} \in B_\delta = B_\delta(u_0, \xi_0)$. Then

$$\Phi'_{\lambda,\sigma} = I - A^{-1}\begin{pmatrix} F_u(\lambda, u) & A_1 \\ \mu_u(\lambda, u) & 0 \end{pmatrix} = A^{-1}\begin{pmatrix} A_0 - F_u(\lambda, u) & 0 \\ A_2 - \mu_u(\lambda, u) & 0 \end{pmatrix},$$

so that

$$\|\Phi'_{\lambda,\sigma}\| \le \|A^{-1}\| \left(\|A_0 - F_u(\lambda, u)\| + \|A_2 - \mu_u(\lambda, u)\|\right) \le \frac{1}{2};$$

hence

$$\|\Phi_{\lambda,\sigma}\tbinom{u}{\xi} - \Phi_{\lambda,\sigma}\tbinom{v}{\eta}\| \le \tfrac{1}{2}(\|u - v\| + \|\xi - \eta\|) \quad \text{for } \tbinom{u}{\xi}, \tbinom{v}{\eta} \in B_\delta.$$

Moreover, for the choice $v = u_0$, $\eta = \xi_0$, we find

$$\begin{aligned}
\|\Phi_{\lambda,\sigma}\tbinom{u}{\xi} - \tbinom{u_0}{\xi_0}\| &\leq \tfrac{1}{2}(\|u-u_0\| + \|\xi - \xi_0\|) + \|\Phi_{\lambda,\sigma}\tbinom{u_0}{\xi_0} - \tbinom{u_0}{\xi_0}\| \\
&\leq \tfrac{\delta}{2} + \left\| -A^{-1}\begin{pmatrix} F(\lambda,u_0) + A_1\xi_0 \\ \mu(\lambda,u_0) - \sigma \end{pmatrix} \right\| \\
&\leq \tfrac{\delta}{2} + \|A^{-1}\|(\|F(\lambda,u_0) + A_1\xi_0\| + \|\mu(\lambda,u_0) - \sigma\|) \leq \delta.
\end{aligned}$$

Hence $\Phi_{\lambda,\sigma}$ is a contraction in B_δ.

Note that if we apply the theorem for different values of λ and σ we may take each time different values for u_0 and ξ_0, too. Therefore we may apply the result with suitable functions

$$u_0 = \tilde{u}(\lambda,\sigma), \quad \xi_0 = \tilde{\xi}(\lambda,\sigma).$$

We now denote by Σ a maximal simply connected, open set of $(\lambda,\sigma) \in \Lambda \times \mathbb{R}^m$ such that $\Phi_{\lambda,\sigma}$ is a contraction in some $B_\delta(u_0,\xi_0)$. By Theorem 3.1, we can choose Σ such that $(\lambda_0, \mu(\lambda_0,u_0)) \in \Sigma$ provided that $\|F(\lambda_0,u_0) + A_1\xi_0\|$,$\|F_u(\lambda_0,u_0) - A_0\|$ and $\|\mu_u(\lambda_0,u_0) - A_2\|$ are sufficiently small. By Banach's fixed point theorem, $\Phi_{\lambda,\sigma}$ has, for each pair $(\lambda,\sigma) \in \Sigma$, a *unique* fixed point in B_δ; we denote this fixed point by $\binom{u(\lambda,\sigma)}{\xi(\lambda,\sigma)}$.

By the implicit function theorem, the mappings $u : \Sigma \to X$ and $\xi : \Sigma \to \mathbb{R}^m$ defined in this way are k times continuously differentiable, and they satisfy

$$F(\lambda,u) + A_1\xi = 0, \quad \mu(\lambda,u) = \sigma. \tag{10}$$

Hence the solution manifold of $F(\lambda,u) = 0$ near u_0 is given in an implicit parametrization by

$$M = \left\{ \begin{pmatrix} u(\lambda,\sigma) \\ \lambda \end{pmatrix} \mid (\lambda,\sigma) \in \Sigma,\ \xi(\lambda,\sigma) = 0 \right\},$$

i.e., M is the image under λ of a *reduced manifold* $\Sigma^* \in \Lambda \times \mathbb{R}^m$ given by

$$\Sigma^* = \{(\lambda,\sigma) \in \Sigma \mid \xi(\lambda,\sigma) = 0\}.$$

In particular, in the special case where $F(\lambda_0,u_0) = 0$ we find with $\sigma_0 = \mu(\lambda_0,u_0)$ that

$$u(\lambda_0,\sigma_0) = u_0, \quad \xi(\lambda_0,\sigma_0) = 0, \tag{11}$$

and hence $(\lambda_0,\sigma_0) \in \Sigma^*$. However, a virtue of our construction is that it also works when $F(\lambda_0,u_0)$ does not vanish.

We mow show that the reduced manifold has the required topological property.

3.2 Theorem. *The mapping $\phi : \Sigma \to X \times \Lambda$ defined by*

$$\phi(\lambda, \sigma) = \begin{pmatrix} u(\lambda, \sigma) \\ \lambda \end{pmatrix}$$

is a diffeomorphism from Σ to $\phi(\Sigma)$. In particular, the solution manifold M is diffeomorphic to Σ^*.

Proof. Since Σ is open and simply connected it is sufficient to show that

$$\phi'(\lambda, \sigma) = \begin{pmatrix} u_\lambda(\lambda, \sigma) & u_\sigma(\lambda, \sigma) \\ I & 0 \end{pmatrix}$$

has rank $p+m$. Indeed, suppose that $\phi'(\lambda, \sigma)q = 0$ for some $q = \binom{x}{y} \in \mathbb{R}^{p+m}$. Then $y = 0$ and $u_\sigma(\lambda, \sigma)x = 0$. Differentiating the identity

$$\sigma x = \mu(\lambda, u(\lambda, \sigma))x$$

with respect to σ gives

$$x = \mu_u(\lambda, u(\lambda, \sigma))u_\sigma(\lambda, \sigma)x = 0,$$

hence $q = 0$. This proves the claim.

By differentiating we can see that, by Theorem 3.2, the singular behaviour of M is completely reflected in the reduced manifold Σ^*. In particular, since by construction

$$\mu(\lambda, u(\lambda, \sigma)) = \sigma,$$

the behaviour of the parameter vector $\mu(\lambda, x)$ on the solution manifold is described by the multi-valued function

$$\hat{\mu}(\lambda) = \{\sigma \mid (\lambda, \sigma) \in \Sigma,\ \xi(\lambda, \sigma) = 0\}.$$

Thus, as in classical Lyapunov-Schmidt reduction (see, e.g., Chow & Hale [3]) we have reduced the analysis of singularities of the solution manifold to a low–dimensional problem. However, we need neither precise information about the location of the singularities nor a precise knowledge of null spaces. Moreover, imperfect bifurcations do not require a special treatment, and, in contrast to the treatment of imperfect bifurcations in Golubitzky & Schaeffer [4, 5], our neighbourhoods need not be arbitrarily small (i.e., defining germs); they can be quite sizeable, since the assumptions of Theorem 3.1 contain no tiny quantities. In particular, it can be constructively checked whether a specified naeighborhood satisfies the requirements (8)-(9).

4 Computation of the reduced manifold

We now address the question of how to compute the reduced manifold Σ^*. Values of $u(\lambda,\sigma)$ and $\xi(\sigma)$ can be obtained from the iteration

$$\begin{pmatrix} u_{l+1} \\ \xi_{l+1} \end{pmatrix} := \begin{pmatrix} u_l \\ \xi_l \end{pmatrix} - A^{-1}\begin{pmatrix} F(\lambda,u_l)+A_1\xi_l \\ \mu(\lambda,u_l)-\sigma \end{pmatrix} \quad (l \geq 0), \tag{12}$$

starting with approximations u_1 and ξ_1 from a previous calculation (continuation!). Since $\Phi_{\lambda,\sigma}$ is a contraction, this iteration converges linearly to $\binom{u(\lambda,\sigma)}{\xi(\lambda,\sigma)}$ with local convergence factor

$$\beta_{\lambda,\sigma} := \varrho\left(A^{-1}\begin{pmatrix} F_u(\lambda,u(\lambda,\sigma))-A_0 & 0 \\ \mu_u(\lambda,u(\lambda,\sigma))-A_2 & 0 \end{pmatrix}\right) < 1; \tag{13}$$

here $\varrho(A)$ denotes the spectral radius of an operator A.

Suppose, in particular, that $F(\lambda_0,u_0)=0$ and $A_0=F_u(\lambda_0,u_0)$, $A_2=\mu_u(\lambda_0,u_0)$. Then we see that $\beta_{\lambda,\sigma}\to 0$ for $\lambda\to\lambda_0$ and $\sigma\to\sigma_0$, so that we have fast converence sufficiently close to u_0; and this still holds when $F(\lambda_0,u_0)$ vanishes only approximately. Moreover, the form of the iteration shows that in each step one has to solve system with a *fixed* linear operator A, thus considerably simplifying the numerical calculations. In practice, (12) is calculated using some kind of discretization. Then the convergence speed may be a little slower, but as in defect correction methods (see, e.g., Böhmer & Stetter [2]), the final accuracy mainly depends on the accuracy with which the last residual is computed. In particular, multigrid techniques are applicable.

We can now see the role of the choice of unfolding functionals μ (determining A_2 approximately) and A_1. A good choice has been made if the resulting A is well conditioned, i.e., $\|A^{-1}\|$ is small. Indeed, this leads to a larger contraction domain Σ (Theorem 3.1) and to a better convergence factor (13). Note that the convergence factor can (and should) be monitored in practice; it serves to determine whether a particular choice of σ is close to the boundary of Σ (indicated by slow convergence). If the solution manifold needs to be explored in a region where the current reduction leads to slow convergence, it is advisable to use a new starting point u_0 in the wanted region and compute a new reduction. This amounts to covering the solution manifold of $F(\lambda,u)=0$ by several local patches, each patch being the image of an appropriate reduced manifold.

To find a local approximation of the parameter manifold Σ^* it suffices to compute a reasonably small number of points $u(\lambda,\sigma),\xi(\lambda,\sigma)$. Then we can approximate $\xi(\lambda, \sigma)$ by fitting a low degree polynomial model $\tilde{\xi}(\lambda, \sigma)$ to the computed points to get an approximation

$$\tilde{\Sigma}^* = \{(\lambda, \sigma) \in \Sigma \mid \tilde{\xi}(\lambda, \sigma) = 0\} \tag{14}$$

for the parameter manifold near (λ_0, u_0).

To find the approximate parameter manifold it is now sufficient to solve m low degree equations in $m + p$ variables. This can be done easily for $m \leq 1$. For $m > 1$ one can use continuation methods adapted to solve singular problems (see, e.g., Allgower & Chien [1]); since the computations are low-dimensional and second derivatives are available for the approximation, the work required is low in comparison with the effort needed to compute the points used for the fit. The model (14) can also be used to calculate approximations of singular points, and, used iteratively, to find their precise location. This provides an alternative to methods quoted in Rheinboldt [7].

In principle, all this can be done with rigorous error estimation using Theorem 3.1 and techniques from interval analysis. Details on how this can be done efficiently and automatically, and specific examples, will be discussed in a separate paper [6].

5 Two examples

We illustrate the general theory with bifurcations from the trivial branch $u_0 = 0$ of a nonlinear eigenvalue problem, first in the finite-dimensional case, then for a two-point boundary value problem.

5.1 Example. We consider the nonlinear eigenvalue problem

$$Su = \lambda T(u) \quad (u \in \mathbb{R}^n, \lambda \in \mathbb{R}),$$

where $S \in \mathbb{R}^{n \times n}$, and $T : \mathbb{R}^n \to \mathbb{R}^n$ is $k \geq 2$ times continuously differentiable with

$$T(0) = 0.$$

Then $u_0 = 0$ is a trivial solution, and we look for a bifurcation from this solution. We pick an arbitrary value λ_0 and put

$$A_0 := F_u(\lambda_0, u_0) = S - \lambda_0 T'(0).$$

If λ_0 is close to a simple eigenvalue of the matrix pencil $(S, T'(0))$ then the absolutely smallest eigenvalue of A_0 is simple and the right and left eigenvectors v_1, w_1^* satisfy $w_1^* v_1 \neq 0$. Hence the matrix $A = \begin{pmatrix} A_0 & A_1 \\ A_2 & 0 \end{pmatrix}$ with $A_1 = v_1,\ A_2 = w_1^*\ , m = 1$ is nonsingular. (If λ_0 is close to a multiple eigenvalue we need to take $m > 1$, with an analogous treatment.)

If we define

$$\mu(\lambda, u) = A_2 u,$$

equations (7) can be rewritten as

$$\begin{pmatrix} u \\ \xi \end{pmatrix} = A^{-1} \begin{pmatrix} (\lambda - \lambda_0) T(u) + \lambda_0 \left(T(u) - T'(0) u \right) \\ \sigma \end{pmatrix}. \tag{15}$$

By Theorem 3.1, the solution is k times continuously differentiable in λ, σ. In particular, for $\lambda = \lambda_0 + O(\varepsilon),\ \sigma = O(\varepsilon)$ we have $u = O(\varepsilon)$ and

$$T(u) = T(0) + T'(0)u + d(\varepsilon^2)(\lambda - \lambda_0)T(u) + \lambda_0 \left(T(u) - T'(0)u \right) = O(\varepsilon^2).$$

Hence (15) gives

$$\begin{pmatrix} u \\ \xi \end{pmatrix} = \sigma \begin{pmatrix} u_1 \\ \xi_1 \end{pmatrix} + O(\varepsilon^2),$$

where

$$A \begin{pmatrix} u_1 \\ \xi_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In particular, $u = u_1 \sigma + O(\varepsilon^2)$,

$$T(u) = T'(0)u + \frac{\sigma^2}{2} T''(0) u_1 + O(\varepsilon^3),$$

$$(\lambda - \lambda_0)T(u) + \lambda_0((T(u) - T'(0)u) = (\lambda - \lambda_0)\sigma T'(0)u_1 + \lambda_0 \frac{\sigma^2}{2} T''(0)u_1 + O(\varepsilon^3),$$

and (15) gives

$$\begin{pmatrix} u \\ \xi \end{pmatrix} = \sigma \begin{pmatrix} u_1 \\ \xi_1 \end{pmatrix} + (\lambda - \lambda_0)\sigma \begin{pmatrix} u_2 \\ \xi_2 \end{pmatrix} + \sigma^2 \begin{pmatrix} u_3 \\ \xi_3 \end{pmatrix} + O(\varepsilon^3),$$

where

$$A \begin{pmatrix} u_2 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} T'(0)u_1 \\ 0 \end{pmatrix}, \quad A \begin{pmatrix} u_3 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\lambda_0 T''(0)u_1 \\ 0 \end{pmatrix}.$$

We therefore have

$$\xi(\sigma, \lambda) = \sigma \left(\xi_1 + (\lambda - \lambda_0)\xi_2 + \sigma \xi_3 \right) + O\left(\sigma^2, (\lambda - \lambda_0)^2 \right). \tag{16}$$

Thus by solving three linear systems with the same coefficient matrix A, we obtain a local quadratic approximation for $\xi(\sigma, \lambda)$. By repeating the same process, higher order terms can be found from additional linear systems with coefficient matrix A.

Note that in the special situation treated (bifurcation from a trivial solution), σ must be a factor of $\xi(\sigma,\lambda)$ since, independent of λ, we have $\sigma = 0$ for the trivial solution $u = 0$, $\xi = 0$.

Relation (16) implies that when $|\xi_1| \ll |\xi_3|$, the equation $\xi(\sigma,\lambda) = 0$ has a second solution

$$\sigma = -\xi_1/\xi_3 - (\lambda-\lambda_0)\xi_2/\xi_3 + O\left((\lambda-\lambda_0)^2, (\xi_1/\xi_3)^2\right)$$

corresponding to a nonlinear branch. This branch crosses the trivial branch $\sigma = 0$ at approximately $\hat{\lambda} = \lambda_0 - \xi_1/\xi_2$ (if $|\xi| \ll |\xi_2|$, so that $\hat{\lambda}$ is still in the region where the expansion is valid). On the other hand, if $|\xi_2| \ll |\xi_1| \ll |\xi_3|$, it is possible that this branch actually never crosses although it comes close to the trivial branch and then wanders off again. This corresponds to an imperfect bifurcation near λ_0.

We see from the example that one would like to model $u(\lambda,\sigma)$ and $\xi(\lambda,\sigma)$ at least by a quadratic (and probably even by a cubic) in $\lambda - \lambda_0$ and $\sigma - \sigma_0$. For the case $m = p = 1$ this requires up to 6 (in the cubic case up to 10) terms can that be calculated as in the example by Taylor expansion and the solution of some linear systems with coefficient matrix A.

For operator equations the same approach works, though one now has to solve instead linear operator equations with different right hand sides.

5.2 Example. To illustrate the effect of the choice of A_1 and A_2 on the norm of A^{-1}, we consider the nonlinear eigenvalue problem

$$\begin{aligned} &\ddot{u}(t) - \lambda f\left(u(t)\right) = 0 \quad \text{for } t \in \Omega := [\pi,\pi] \\ &u(-\pi) = u(\pi) = 0, \end{aligned} \tag{17}$$

where $f \in C^1(\mathbb{R})$ satisfies

$$f(0) = 0,\ f'(0) = 1.$$

For every $\lambda \in \mathbb{R}$, (17) has the trivial solution $u_0 = 0$. Nontrivial solutions branch off at the eigenvalues of the linearized problem, here at $\lambda = k^2$ ($k = 1, 2, \ldots$). We look at some arbitrary $\lambda_0 = \omega^2$, ω not an integer.

The problem (17) can be posed as finding zeros of the nonlinear operator $F : \mathbb{R} \times X \to Y$ defined by

$$F(\lambda, u) = \ddot{u} + \lambda f(u),$$

where

$$X = \{u \in C^{2,1}(\Omega) \mid u(-\pi) = u(\pi) = 0\}, \quad Y = C^{0,1}(\Omega).$$

We shall consider the unfolding functional

$$\mu = u(\tau)$$

at a point c to be chosen later, and take $A_0 = F_u(\lambda_0, u_0),\ A_2 = \mu_u(\lambda_0, u_0)$. Thus

$$\begin{aligned} A_0 u &= \bar{u} + \omega^2 u, \\ A_1 \xi &= \xi g, \\ A_2 u &= u(\tau), \end{aligned}$$

with some function $g(t)$ to be chosen later. The equation $A\binom{u}{\xi} = \binom{v}{\eta}$ then becomes

$$\begin{aligned} \ddot{u} + \omega^2 u &= v - \xi g, \\ u(-\pi) = u(\pi) &= 0, \\ u(\tau) &= \sigma. \end{aligned} \tag{18}$$

In this example, the inverse operator can be calculated explicitly using Fourier series. (In general, this will not be the case.) We expand v and g into Fourier series,

$$\begin{aligned} v &= \sum a_k \cos kt + \sum b_k \sin kt, \\ g &= \sum c_k \cos kt + \sum d_k \sin kt, \end{aligned}$$

and find for $\omega \notin \mathbb{Z}$ the following expansion for u:

$$u = \sum (-1)^k \frac{a_k - \xi c_k}{\omega^2 - k^2} \left(\frac{\cos kt}{\cos k\pi} - \frac{\cos \omega t}{\cos \omega\pi} \right) + \sum \frac{b_k - \xi d_k}{\omega^2 - k^2} \sin kt. \tag{19}$$

Now suppose that ω is close to some integer $k_0 \neq 0$. Then the terms in (19) are well-behaved, with exception of the term $\frac{b_{k_0} - \xi d_{k_0}}{\omega^2 - k_0^2} \sin k_0 t$. The unfolding equation $u(\tau) = \sigma$ must render this term harmless, and this is the case

$$|d_{k_0} \sin k_0 \tau| \gg 0, \tag{20}$$

since then the sum in (19) can take the (small) value σ for $t = \tau$ only when $b_{k_0} - \xi d_{k_0}$ is of the order $d_{k_0}(\omega^2 - k_0^2)$. And then $\xi \approx b_{k_0}/d_{k_0}$ is harmless, too. Thus we see that A^{-1} is well-behaved iff (20) holds.

This is a requirement on A_2 since it says that the evaluation $A_2 u = u(\tau)$ must not be taken too close to a multiple of π/k_0, and a requirement on A_1, requiring that the basis function g with $A_1 \xi = \xi g$ contains a significant contribution of $\sin k_0 t$, the eigenfunction of the linearized problem near $\lambda_0 = \omega^2$. (Note that this problem is self-adjoint, and in more general problems, the relevant eigenfunction ist that of the adjoint problem, cf. Proposition 2.1!)

References

[1] E. L. Allgower and C.-S. Chien, Continuation and local perturbation for multiple bifurcations, SIAM J. Sci. Stat. Comput. 7 (1986), 1265–1281.

[2] K. Böhmer and H. J. Stetter (eds.), Defect correction methods. Computing Suppl. 5, Springer, Wien 1984.

[3] S.-N. Chow and J. K. Hale, Methods of bifurcation theory, Springer, New York 1982.

[4] M. Golubitsky and D. Schaeffer, A theory for imperfect bifurcation via singularity theory, Comm. Pure Appl. Math. 32 (1979), 21-98.

[5] M. Golubitsky and D. Schaeffer, Singularities and Groups in Bifurcation Theory I (Applied Mathematical Sciences, Vol 51) Springer, New York 1985.

[6] M.T. Nakao and A. Neumaier, The verification of bifurcation diagrams for nonlinear elliptic equations. In preparation.

[7] W. C. Rheinboldt, Numerical analysis of parametrized nonlinear equations, Wiley, New York 1986.

[8] E. Zeidler, Nonlinear functional analysis and its applications I, Springer, Berlin 1986.