Numerical analysis of homoclinic orbits emanating from a Takens–Bogdanov point

W.-J. BEYN

Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany

[Received 23 September 1992 and in revised form 10 February 1993]

It is known that branches of homoclinic orbits emanate from a singular point of a dynamical system with a double zero eigenvalue (Takens–Bogdanov point). We develop a robust numerical method for starting the computation of homoclinic branches near such a point. It is shown that this starting procedure relates to branch switching. In particular, for a certain transformed problem the homoclinic predictor is guaranteed to converge to the true orbit under a Newton iteration.

1. Introduction

One of the important discoveries of dynamical systems theory says that global bifurcations (such as homoclinic or torus bifurcations) appear locally in the neighbourhood of singular stationary points. We find these phenomena near singular points of codimension at least 2, i.e. stationary points which occur generically in dynamical systems with at least two parameters

\[ \dot{u} = f(u, \lambda), \quad u \in \mathbb{R}^n, \quad \lambda \in \mathbb{R}^2. \]  

(1.1)

The most prominent examples are

(i) the formation of saddle–saddle homoclinic orbits close to a Takens–Bogdanov singularity, i.e. a stationary point with a double zero eigenvalue of the Jacobian (see Takens, 1974; Bogdanov, 1975; Arnold, 1988; Guckenheimer & Holmes, 1983);

(ii) the occurrence of saddle–focus homoclinic orbits of Shilnikov type and of torus bifurcations close to a Gavrilov–Guckenheimer singular point, i.e. a stationary point with a zero and two imaginary eigenvalues (Langford, 1979; Guckenheimer, 1984; Guckenheimer & Holmes, 1983; Arnold, 1988; Kirk, 1991; Gaspard, 1992);

(iii) focus–focus homoclinic orbits created at a double Hopf point (Morozov, 1982; Arnold, 1988).

In some sense these global bifurcations render the analysis of phase portraits near singular points more and more difficult. However, if such an analysis is possible, it provides an interesting insight into the interplay of bifurcations.

From a numerical point of view this effect shows that computed branches of singular stationary points and global bifurcations are connected to each other and that this should have some impact on continuation codes such as AUTO (Doedel & Kernévez, 1986) or LOCBIF (Khibnik, 1990; Khibnik et al 1992). In particular, we
could use codimension 2 singular points to start the numerical continuation of global bifurcation curves.

The purpose of this paper is to carry out such a procedure in full algorithmic detail for the branch of homoclinic orbits emanating from a TB-point (Takens–Bogdanov).

For the sake of convenience let us mention the two-dimensional model example of a TB-point (cf. Guckenheimer & Holmes, 1983; Arnold, 1988; Beyn, 1991)

\[ \dot{u}_1 = u_2, \quad \dot{u}_2 = \lambda_1 + \lambda_2 u_2 - u_1^2 + u_1 u_2. \]  

(1.2)

Its bifurcation diagram is well-known (cf. Guckenheimer & Holmes, 1983; Arnold, 1988) and is shown in Fig. 1. There exists a branch of homoclinic orbits which, in the parameter plane, is approximately given by \( \lambda_1 = (\frac{1}{4} \lambda_2)^2, \lambda_2 < 0. \)

In a generic sense the system (1.2) represents the normal form of a system near
a TB-point. It is this type of TB-point which we will treat in the current paper. However, if the problem has a symmetry, other normal forms become appropriate. For example, if the symmetry forces the existence of a trivial solution for all parameter values, a suitable normal form is (cf. Carr, 1981; Hale, 1983)

\[
\dot{u}_1 = u_2, \quad \dot{u}_2 = \lambda_1 u_1 + \lambda_2 u_2 - u_1^3 + u_1 u_2.
\]  

(1.3)

Here the critical lines in Fig. 1 become transversal, and it is not too difficult to adapt our approach to this situation.

The idea of starting homoclinic branches near TB-points in planar systems with the help of Melnikov’s method was first used in Rodriguez Luis et al (1990). In this paper we will treat the general n-dimensional case. Moreover, in our development of the overall algorithm (summarized in Appendix A) we emphasize the following features.

1. All linear systems which have to be solved during the computations exhibit the same type of matrix. These matrices result from bordering the (possibly singular) Jacobian \( f_u \) with one or two vectors. There are well-known techniques for such systems which employ black box solvers for \( f_u \) and hence fully exploit its possible structure such as sparsity (see Chan, 1984; Moore, 1987; Govaerts, 1991; Govaerts & Pryce, 1991).

2. From the centre manifold reduction and the normal form transformations we try to extract the minimum number of coefficients which are needed for the approximate homoclinic orbits.

3. Melnikov’s original method (Melnikov, 1963; Guckenheimer & Holmes, 1983) is used in the version of Hale (1983), who reformulates it as a Liapunov–Schmidt reduction in function spaces. In this way, starting homoclinic orbits at a TB-point may be viewed as branch switching and we use this viewpoint to establish convergence of Newton’s method close to the predicted homoclinic orbit (§5). It turns out that the convergence cones from Jepson and Decker (1986) have to be refined.

In §2 we discuss the numerical computation of TB-points using the approach of Griewank and Reddien (1989) (see Roose, 1987; Pönisch, 1991; for related methods). The resulting transversality conditions and further algorithmic details prove to be useful for the next steps in §3 and §4. There we perform the centre manifold reduction and the unavoidable scalings and normal form transformations.

Finally, in §6 we apply the above procedure to a few examples. The above algorithm is utilized to start a branch of homoclinic orbits and then to continue it by the method described in Beyn (1990a).

Our examples are all low dimensional though the above algorithm is well-suited for large scale problems (1.1) (see feature 1 above). The reason is that the initial approximate homoclinic may be obtained cheaply, but the correction to the true homoclinic involves the solution of a large boundary value problem. In Bai et al (1992) a typical case of this type is treated where (1.1) is obtained by the method of lines from a spectral approximation to a PDE.

In concluding, let us remark that our numerical results suggest that the theorem
on convergence in §5 should carry over from the reduced transformed system to the original system (1.1). But we are not aware of any rigorous result yet.

2. Defining equation for \( \tau \beta \)-points

In this section we analyze a defining equation for \( \tau \beta \)-points based on the approach of Griewank and Reddien (1989). In addition to their paper some more details will be worked out which will prove to be useful in the following sections. Other direct methods for \( \tau \beta \)-points have been proposed in Roose (1987) and Pönisch (1991).

We consider an \( n \)-dimensional dynamical system with two parameters

\[
\begin{align*}
\dot{u} &= f(u, \lambda), \quad u(t) \in \mathbb{R}^n, \\
\lambda &= (\lambda_1, \lambda_2) \in \mathbb{R}^2
\end{align*}
\]

(2.1)

where \( f \in C^k(\mathbb{R}^{n+2}, \mathbb{R}^n), \ k \geq 3. \)

A stationary point \((u^0, \lambda^0)\) of (2.1) is called a \( \tau \beta \)-point, if

\[
J_0 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}
\]

is the only entry in the Jordan normal form of

\[
f_u^0 = \frac{\partial f}{\partial u}(u^0, \lambda^0)
\]

which belongs to the zero eigenvalue. Of course, the double zero eigenvalue is not sufficient to guarantee a specific singularity (up to flow equivalence) and an unfolding picture as in Fig. 1. But for simplicity we refrain from calling such a point a 'potential \( \tau \beta \)-point' as has been done for Hopf points in Keller and Jepson (1984).

The extra conditions on the higher-order Taylor terms usually fall into two categories (cf. Arnold, 1988; Khibnik, 1990):

- **non-degeneracy conditions** for the higher derivatives of \( f \) with respect to the state variable \( u \);
- **transversality conditions** for the higher derivatives of \( f \) with respect to \( u \) and \( \lambda \).

Violating one of the non-degeneracy conditions increases the codimension of the singularity while violating a transversality condition means that the given parameters may not exhibit the standard unfolding picture (see (1.3) for a typical case).

In what follows we will investigate the relation between the extra conditions and the regularity of the defining equation.

Let us assume that we are given vectors \( b_0, c_0 \in \mathbb{R}^n \) such that the \((n + 1) \times (n + 1)\) matrix (where \( T \) represents transpose)

\[
A(u, \lambda) = \begin{pmatrix} f_u^0(u, \lambda) & b_0 \\ c_0^T & 0 \end{pmatrix}
\]

(2.2)

is non-singular for \((u, \lambda)\) in some domain \( \Omega \subset \mathbb{R}^{n+2} \). Of course, this requires the rank of \( f_u \) to be at least \( n - 1 \) in \( \Omega \).
Then we calculate a TB-point by solving the following defining equation

\[ S(u, \lambda) = \begin{pmatrix} f(u, \lambda) \\ g(u, \lambda) \\ h(u, \lambda) \end{pmatrix} = 0. \]  

(2.3)

Here the functions \( g, h \in C^{k-1}(\Omega, \mathbb{R}) \) are implicitly defined through

\[ \begin{align*}
A \left( \begin{array}{c} v \\ g \\ h \end{array} \right) &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\
A \left( \begin{array}{c} w \\ h \\ v \end{array} \right) &= \begin{pmatrix} v \\ 0 \end{array} \end{align*} \]

for \( (u, \lambda) \in \Omega. \)  

(2.4)

The \( C^{k-1} \) functions \( v \) and \( w \) are useful since they give us the generalized eigenvectors at the TB-point and since they help us to calculate the derivatives of \( g \) and \( h \) according to the following lemma (see Griewank & Reddien, 1989; Beyn, 1991).

**Lemma 2.1** In addition to (2.4) define the functions \( \tilde{g}, \tilde{h} \in C^{k-1}(\Omega, \mathbb{R}) \) and \( \Psi, \zeta \in C^{k-1}(\Omega, \mathbb{R}^n) \) by the adjoint equations

\[ \begin{align*}
(\Psi^T, \tilde{g})A &= (0, 1), \\
(\zeta^T, \tilde{h})A &= (\Psi^T, 0). \\
\end{align*} \]

Then the following relations hold for all \( z = (u, \lambda) \in \Omega \)

\[ \begin{align*}
g &= \tilde{g} = -\Psi^T f u, \\
h &= \tilde{h} = \Psi^T v \\
g_z &= -\Psi^T f u_z v, \\
h_z &= -\Psi^T f u_z w - \zeta^T f u_z v. \\
\end{align*} \]

(2.5)

(2.6)

(2.7)

A proof of this lemma may be found in Griewank and Reddien (1989).

During the Newton process for equation (2.3) we use formulae (2.6) and (2.7) in order to evaluate

\[ S = \begin{pmatrix} f \\ g \\ h \end{pmatrix} \quad \text{and} \quad S' = \begin{pmatrix} f_u & f_\lambda \\ g_u & g_\lambda \\ h_u & h_\lambda \end{pmatrix}. \]

The matrix \( S' \) is obtained by bordering \( f_u \) with two rows and two columns. One linear system with \( S' \) can be reduced to 4 linear systems with the matrix \( A \) from (2.2) and a small \( 3 \times 3 \) system (see Griewank & Reddien 1989; Beyn 1991). If, in addition, the second derivatives in (2.7) are approximated by forward differences of \( f_z \), then one Newton step for (2.3) requires 5 linear systems with \( A \), two with \( A^T \) and 4 evaluations of \( f_z = (f_u, f_\lambda). \)

Initial approximations for the TB-point are usually obtained when following a branch of turning points or Hopf points (Roose, 1987; Pönisch, 1991; Spence et al 1989).

Obviously, any TB-point \( (u^0, \lambda^0) \in \Omega \) is a solution of (2.3), but the converse only holds if we assume the eigenvalue zero to be of multiplicity at most 2.

Now we analyze the non-singularity of \( S'((u^0, \lambda^0) = S'' \) at a TB-point. In contrast to Roose (1987) and Pönisch (1991) we do not assume that \( (u^0, \lambda^0) \) is a quadratic turning point, i.e. \( f^0 u v^0 u^0 \notin R(f^0 u) \). Such a non-degeneracy condition will only be used for the homoclinic orbits in Section 4.

Let the superscript ‘0’ always denote evaluation at \( (u^0, \lambda^0) \). From (2.4) and
(2.5) we have
\[ f_u^0 v^0 = 0, \quad f_u^0 w^0 = v^0, \quad \Psi^{0T} f_u^0 = 0, \quad \zeta^{0T} f_u^0 = \Psi^{0T}. \] (2.8)
Moreover,
\[ \Psi^{0T} v^0 = 0 \quad \text{and} \quad \alpha := \Psi^{0T} w^0 = \zeta^{0T} v^0 \neq 0. \] (2.8)
There is no practical loss of generality in assuming biorthogonality
\[ \Psi^{0T} v^0 = 0, \quad \Psi^{0T} w^0 = \zeta^{0T} v^0 = 1. \] (2.9)
This can always be achieved by replacing \( \Psi^0 \) by \( \alpha^{-1} \Psi^0 \) and \( \zeta^0 \) by \( \alpha^{-1}(\zeta^0 - \alpha^{-1}(\zeta^{0T} w^0) \Psi^0) \). Equivalently, we may view this process as an \textit{a posteriori} normalization of the vectors \( b_0, c_0 \) from (2.2).
Let us put \( f_u^0 \) into block diagonal form
\[ f_u^0 = \Phi \begin{bmatrix} H & 0 \\ 0 & J_0 \end{bmatrix}, \quad \Phi = (R \ v^0 \ w^0) \in \mathbb{R}^{n \times n} \] (2.10)
where \( H \in \mathbb{R}^{n-2, n-2} \) is non-singular and the columns of \( R \) form a basis of the invariant subspace \( \{ \Psi^0, \zeta^0 \} \). The matrices \( H \) and \( R \) are introduced for theoretical reasons only. They need not be computed for the construction of the homoclinic orbits, see Section 4 and Appendix A. This is particularly important for large sparse matrices \( f_u^0 \).
Using (2.10) we put (2.1) into linear normal form (including parameters) by the transformation
\[ \begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} u_0 \\ \lambda_0 \end{pmatrix} + \Gamma \begin{pmatrix} x \\ \mu \end{pmatrix}, \quad \Gamma = \begin{bmatrix} \Phi & D \\ 0 & B \end{bmatrix} \in \mathbb{R}^{n+2, n+2}. \] (2.11)
The matrices \( D \in \mathbb{R}^{n-2} \) and \( B \in \mathbb{R}^{2, 2} \) will be determined in such a way that the transformed system
\[ \dot{x} = \Phi^{-1} f \begin{pmatrix} u_0 \\ \lambda_0 \end{pmatrix} + \Gamma \begin{pmatrix} x \\ \mu \end{pmatrix} =: \tilde{f}(x, \mu) \] (2.12)
satisfies
\[ \tilde{f}'(0, 0) = \begin{pmatrix} H & 0 \\ 0 & J_0 \end{pmatrix}. \] (2.13)
From (2.10) we find that this holds iff
\[ (w^0 \ 0) = \Phi \begin{bmatrix} 0 \\ J_0 \end{bmatrix} = f_u^0 D + f_\lambda^0 B. \] (2.14)
In order to achieve this, we assume the transversality condition
\[ 0 \neq \Psi^{0T} f_\lambda^0 =: (\beta_1, \beta_2). \] (T1)
Clearly, this is equivalent to rank \((f_u^0, f_\lambda^0) = n \) and hence also a necessary condition for the non-singularity of \( S^{0T} \). By (T1) there is a non-singular matrix \( B \) such that
\[ C = (w^0 \ 0) - f_\lambda^0 B \] has both columns in the range of \( f_u^0 \). For example, we may take \( B \), to be
orthogonal

\[ B = (\beta_1^2 + \beta_2^2)^{-1} \begin{pmatrix} \beta_1 & -\beta_2 \\ \beta_2 & \beta_1 \end{pmatrix}. \]

The two columns of \( D \) can then be computed in a numerically stable way from the system

\[ A^0 \begin{pmatrix} D \\ d \end{pmatrix} = \begin{pmatrix} C \\ 0 \end{pmatrix}. \]

(2.15)

By the construction of \( C \) we have \( d = 0 \). In practice we can use \( \|d\| \) to check the accuracy. Using the linear transformation (2.11) for the defining operator \( S \) from (2.3) yields

\[ T(x, \mu) = S \left( \begin{pmatrix} u^0 \\ \lambda^0 \end{pmatrix} + \Gamma (x) \right) \]

and by (2.12), (2.13)

\[ T'(0, 0) = S^0 \Gamma = \begin{pmatrix} \Phi & 0 \\ 0 & l_2 \end{pmatrix} \begin{pmatrix} \tilde{r}(0, 0) \\ g \end{pmatrix} \Gamma. \]

(2.16)

From the structure of \( \tilde{r}'(0, 0) \) we find that \( S^0 \) is non-singular if and only if the \( 2 \times 2 \) minor of \( \begin{pmatrix} g_2^0 \\ h_2^0 \end{pmatrix} \Gamma \) formed by the last and fourth but last column is non-zero, i.e.

\[ 0 \neq \Delta := \det \begin{pmatrix} g_2^0 v^0 & g_2^0 \delta_2 \\ h_2^0 v^0 & h_2^0 \delta_2 \end{pmatrix}, \quad \begin{pmatrix} D \\ B \end{pmatrix} = (\delta_1, \delta_2) \in \mathbb{R}^{n+2,2}. \]

(2.17)

By Lemma 2.1 we can express \( \Delta \) in terms of the original function \( f \). We introduce the coefficients

\[ Q_{1,11} = s^{00} f^{00} v^0 w^0, \quad Q_{2,11} = \psi^{00} f^{00} v^0 w^0, \quad Q_{2,12} = \psi^{00} f^{00} v^0 w^0, \quad Q_{1,14} = s^{00} f^{00} v^0 \delta_2, \quad Q_{1,14} = s^{00} f^{00} v^0 \delta_2, \quad Q_{2,24} = \psi^{00} f^{00} w^0 \delta_2. \]

(2.18)

The index notation for the \( Qs \) will be made clear in the next section where they appear in the quadratic terms of the centre manifold reduction. (2.17) may now be written as the transversality condition

\[ 0 \neq \Delta = Q_{2,11}(Q_{1,14} + Q_{2,24}) - Q_{2,14}(Q_{1,11} + Q_{2,12}) \]

(T2)

and we can summarize the preceding discussion in the following theorem.

**Theorem 2.2** Let \( (u^0, \lambda^0) \) be a TB-point of (2.1) in some domain \( \Omega \subset \mathbb{R}^{n+2} \) where the matrices (2.2) are non-singular. Then \( (u^0, \lambda^0) \) is a regular solution of the defining equation (2.3) if and only if the transversality conditions (T1) and (T2) are satisfied.

A few remarks concerning (T2) seem to be in order. If we truncate \( S \) to

\[ \bar{S}(u, \lambda) = \begin{pmatrix} f \\ g \end{pmatrix}(u, \lambda), \]
then (T1) and (T2) imply \( \text{rank} (S'^n) = n + 1 \). Hence there exists a branch \( z(s) = (u(s), \lambda(s)) \) of solutions to \( S = 0 \) such that \( z(0) = (u^0, \lambda^0) \). Moreover, by differentiating \( S(z(s)) = 0 \) and using (2.16) we obtain

\[
z'(0) = \gamma \left[ g_x^0 \sigma_2 \begin{pmatrix} u^0 \\ 0 \end{pmatrix} - (g_x^0 v^0) \delta_2 \right]
\]

for some \( \gamma \neq 0 \). (2.19)

The points \( z(s) \) are singular solutions of (2.1) with the left and right eigenvectors of \( f_u \) being \( \Psi(z(s)) \) and \( v(z(s)) \), see (2.4), (2.5). The transversality condition used in Roose (1987) and Pönisch (1991) is

\[
0 \neq \frac{d}{ds} [\Psi(z(s))^T v(z(s))]_{s=0}.
\]

(2.20)

With the help of Lemma 2.1 and (2.19) we find

\[
\frac{d}{ds} [\Psi(z(s))^T v(z(s))]_{s=0} = \frac{d}{ds} h(z(s))|_{s=0} = h_x^0 z'(0) = \gamma \Delta
\]

and hence (2.20) and (T2) are equivalent. In Roose (1987), Spence et al (1989), and Pönisch (1991) the additional non-degeneracy condition \( Q_{2,11} \neq 0 \) is imposed, so that \( z(s) \) becomes a branch of quadratic turning points. A typical example where \( Q_{2,11} \) vanishes, but Theorem 2.2 is still applicable is

\[
\dot{u}_1 = u_2, \quad \dot{u}_2 = \lambda_1 + \lambda_2 u_1 + u_1 u_2.
\]

This is of course a singularity of codimension higher than two, but the defining equation (2.3) does not take into account these dynamical features. However, for the construction of the homoclinic orbits in the next section we will need the non-degeneracy condition \( Q_{2,11} \neq 0 \).

3. The centre manifold reduction and scaling

Since the centre manifold reduction is a fairly standard technique (see e.g. Carr, 1981; Guckenheimer & Holmes, 1983) we emphasize here the computational aspects needed for the overall algorithm.

Let us assume that we have computed a TB-point \( (u^0, \lambda^0) \) as a regular solution of (2.3) and that \( f_x^0 \) has no other eigenvalues on the imaginary axis except zero. The matrix \( H \) in (2.10) is then hyperbolic. Let us write the transformed system (2.12) as

\[
\dot{x} = f(x, \mu), \quad \dot{\mu} = 0
\]

(3.1)

where \( x^T = (\eta, \xi) \in \mathbb{R}^{n-2} \times \mathbb{R}^2, \mu \in \mathbb{R}^2 \). Then there exists a locally invariant centre manifold for (3.1)

\[
W_{\text{loc}} = \{(F(\xi, \mu), \xi, \mu) : (\xi, \mu) \in U(0) \subset \mathbb{R}^4\}
\]

where \( F \in C^{k-1} (U(0), \mathbb{R}^{n-2}) \) and

\[
F(0, 0) = 0, \quad F'(0, 0) = 0.
\]

(3.2)
The reduced system within $W^s_{\text{loc}}$ is
\[
\dot{\xi} = \left(\frac{\dot{f}_n}{\dot{f}_n}\right)(F(\xi, \mu), \xi, \mu) = r(\xi, \mu).
\] (3.3)

By (2.9), (2.12), (2.13) and (3.2) the right hand side has a Taylor expansion
\[
r(\xi, \mu) = \left(\frac{\dot{f}_n}{\dot{f}_n}\right)(0, \xi, \mu) + O_3 = \left(\frac{\xi_2}{\mu_1}\right) + \frac{1}{2}Q(\frac{\xi}{\mu})^2 + O_3
\]
where $O_3$ denotes third-order terms and the quadratic form $Q: \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}^2$ is given by
\[
Q(\frac{\xi}{\mu})^2 = \left(\xi^0 \psi^0_{00} \right) f^{\alpha_0} \left[\begin{array}{cc}
v^0 & w^0 \\
0 & B
\end{array}\right] \xi^0 \mu.
\] (3.4)

If we consider $\mu_1, \mu_2$ as the third and fourth component of $\left(\frac{\xi}{\mu}\right)$ then the indices in (2.18) are consistent with (3.4). In addition to the coefficients from (2.18) we also need
\[
Q_{2.44} = \psi^{0T} f^{\alpha_0} \delta_2^2.
\] (3.5)

All the other quadratic and higher-order coefficients will not be used later on, so we work with the system
\[
\dot{\xi}_1 = \xi_2 + \frac{1}{2}Q_{1.11} \xi_1^2 + Q_{1.14} \xi_1 \mu_2 + \cdots
\]
\[
\dot{\xi}_2 = \mu_1 + \frac{1}{2}Q_{2.11} \xi_1^2 + Q_{2.12} \xi_2 + Q_{2.14} \xi_1 \mu_2 + Q_{2.24} \xi_2 \mu_2 + \frac{1}{2}Q_{2.44} \mu_2^2 + \cdots
\] (3.6)

The next two steps for simplifying this system usually are normal form transformation of the quadratic terms and scaling (cf. Wiggins, 1990, Chapters 3.1 and 4.9). We put these two steps together and try to minimize the number of quadratic terms in the normal form transformation. The variables $(\xi_1, \xi_2, \mu_1, \mu_2)$ are replaced by $(x, y, \epsilon, \tau)$ as follows
\[
\mu_1 = (\alpha_0 + \alpha_1 \tau^2) \epsilon^4, \quad \mu_2 = \tau \epsilon^2
\] (3.7a)
\[
\dot{\xi}_1(t) = \alpha_2 \epsilon^2 (x(\alpha_3 \epsilon t) + \alpha_4 \tau), \quad \dot{\xi}_2(t) = \alpha_5 \epsilon^3 y(\alpha_3 \epsilon t).
\] (3.7b)

Here the constants $\alpha_i$ are yet to be determined. The terms which involve $\alpha_1, \alpha_4$ are necessary for the normal form and in case $\alpha_1 = \alpha_4 = 0$ we perform a well-known scaling (cf. Guckenheimer & Holmes, 1983; Wiggins, 1990). A straightforward but tedious calculation yields the transformed system
\[
\begin{pmatrix}
\dot{x} \\
\dot{y}
\end{pmatrix} = g(x, y, \epsilon, \tau) = \begin{pmatrix} y \\ x^2 - 4 \end{pmatrix} + \epsilon \begin{pmatrix}
\frac{1}{2}a_1x^2 + a_2xy + a_3y^2 \\
b_1xy + b_2xy
\end{pmatrix} + O(\epsilon^2)
\] (3.8)
by setting
\[
\alpha_0 = -(1/2)Q_{2.11}, \quad \alpha_1 = \alpha_0(Q_{2.11}Q_{2.44} + Q_{2.14}^2), \quad \alpha_2 = -\alpha_0, \quad \alpha_3 = \frac{1}{4}, \quad \alpha_4 = -2Q_{2.14}, \quad \alpha_5 = -\frac{1}{2} \alpha_0.
\]

The coefficients in the quadratic terms of (3.8)—except for $a_3$, which is not
needed later on—are found to be
\begin{equation}
\begin{align*}
a_1 &= \bar{q}Q_{1,11}, \\
b_1 &= Q_{2,12}, \\
a_2 &= 2\bar{q}(Q_{1,14}Q_{2,11} - Q_{1,11}Q_{2,14}), \\
b_2 &= 2\bar{q}(Q_{2,24}Q_{2,11} - Q_{2,14}Q_{2,12}).
\end{align*}
\end{equation}
We notice that it is this step which requires the quadratic turning point condition
\begin{equation}
Q_{2,11} = \Psi^0 \int_0^t v^n v^n \neq 0
\end{equation}
which we were able to avoid in the previous section. Moreover, the transversality condition (T2) now translates into
\begin{equation}
a_2 + b_2 \neq 0.
\end{equation}
For \( \epsilon = 0 \) the system (3.8) is Hamiltonian and has the well-known homoclinic orbit (cf. Guckenheimer & Holmes, 1983).
\begin{equation}
(x(t), y(t)) = 2\left(1 - 3 \text{sech}^2(t), 6 \text{sech}^2(t) \tanh(t)\right), \quad t \in \mathbb{R}.
\end{equation}
This explicit formula provides the root to all approximate homoclinic orbits for the original system (2.1).

4. Melnikov's method and branch switching

The persistence of homoclinic orbits for the perturbed system (3.8) with \( \epsilon \neq 0 \) is usually analyzed by Melnikov's method (Melnikov, 1963; Guckenheimer & Holmes, 1983; Wiggins, 1990). This method is oriented at the geometric behaviour of stable and unstable manifolds in phase space. Here we prefer to use Hale's approach (Hale, 1983). He shows that Melnikov's method can be viewed as a Liapunov-Schmidt reduction of a suitable operator equation in function spaces. In fact, in the case of equation (3.8) a simple bifurcation point occurs. Therefore, Melnikov's method relates to branch switching and it is this viewpoint which will be used in Section 5 to establish convergence of our numerical method.

We introduce the Banach spaces
\begin{equation}
X_0 = \left\{ z \in C(\mathbb{R}, \mathbb{R}^2) : \lim_{t \to \pm \infty} z(t) \text{ and } \lim_{t \to \pm \infty} z(t) \text{ exist} \right\}
\end{equation}
with norm
\begin{equation}
\|z\|_0 = \sup \{ \|z(t)\| : t \in \mathbb{R} \}, \quad \| \cdot \| \text{ some norm in } \mathbb{R}^m
\end{equation}
and
\begin{equation}
X_1 = \left\{ z \in C^1(\mathbb{R}, \mathbb{R}^2) : z, \dot{z} \in X_0 \right\}, \quad \|z\|_1 = \|z\|_0 + \|\dot{z}\|_0.
\end{equation}
With \( w = (z, \epsilon) = (x, y, \epsilon) \in X_1 \times \mathbb{R} \) let us write (3.8) as
\begin{equation}
F(w, \tau) = 0
\end{equation}
where \( F : X_1 \times \mathbb{R} \times \mathbb{R} \to X_0 \times \mathbb{R} \) is defined by
\begin{equation}
F(w, \tau) = \begin{pmatrix} \dot{z} - g(z, \epsilon, \tau) \\ y(0) \end{pmatrix}.
\end{equation}
The condition \( y(0) = 0 \) is used to fix the phase of the homoclinic orbits.
A careful inspection of the transformation (3.7b) shows that $F$ is only defined for arguments $(w, \tau) = (z, \varepsilon, \tau)$ satisfying $|\varepsilon| \leq \varepsilon_0$, $|\tau| \leq C\varepsilon^{-2}$, $||z|| \leq C\varepsilon^{-2}$ for some $\varepsilon_0 > 0$. This will be sufficient for our purposes and we will not always mention it explicitly.

Setting $\tilde{w} = (\tilde{x}, \tilde{y}, 0)$ with $(\tilde{x}, \tilde{y})$ from (3.12) we find

$$F(\tilde{w}, \tau) = 0 \quad \text{for all } \tau.$$ 

Hence we have a trivial branch $(\tilde{w}, \tau)$ of homoclinic orbits and we look for values of $\tau$ at which bifurcation occurs.

**Theorem 4.1** Consider the two-parameter system (3.8) with $g$ in $C^2$ and assume $a_2 + b_2 \neq 0$ as in (3.11). Then equation (4.1) has a unique simple bifurcation point (in the sense of Crandall & Rabinowitz, 1971) at

$$
\tau_0 = \frac{10a_1 + b_1}{7a_2 + b_2}.
$$

The emanating $C^1$-branch can be parametrized by $\varepsilon$

$$
(w(\varepsilon), \tau(\varepsilon)) = (z(\varepsilon), \varepsilon, \tau(\varepsilon)) \in X \times \mathbb{R}^2.
$$

It has tangent

$$
(z'(0), 1, \tau'(0)) = (z_0, 1, \tau'(0))
$$

where $z_0 = (x_0, y_0)$ is the unique solution in $X_1$ of the linear system

$$
\begin{align*}
\dot{x} - y &= \frac{1}{2}a_1 \dot{x}^2 + a_2 x \dot{x} + a_3 x^2 \\
\dot{y} - 2\dot{x} &= b_1 \dot{x} + b_2 \dot{y}
\end{align*}
$$

and $y(0) = 0$.

**Remark** By our construction we only know that both limits

$$
\lim_{\varepsilon \to \infty} z(t, \varepsilon) = z_+ \quad \text{and} \quad \lim_{\varepsilon \to -\infty} z(t, \varepsilon) = z_-
$$

exist.

These points must be stationary points of (3.8) at $\tau = \tau(\varepsilon)$ and lie close to the stationary point $(2, 0)$ of the unperturbed system ($\varepsilon = 0$). Since $(2, 0)$ is hyperbolic, the implicit function theorem guarantees unique nearby stationary points of the perturbed system which then have to coincide with $z_-$ and $z_+$. Therefore, the orbits $z(\varepsilon)$ are in fact homoclinic.

**Proof.** Consider the linearization of $F$ with respect to $w \in X \times \mathbb{R}$ about the trivial solution $(\tilde{w}, \tau)$.

$$
F_w(\tilde{w}, \tau)w = \begin{pmatrix} Lz - g_\varepsilon(\tilde{z}, 0, \tau) \varepsilon \\ y(0) \end{pmatrix}, \quad w = (z, \varepsilon) = (x, y, \varepsilon) \in X \times \mathbb{R}
$$

where $L: X_1 \to X_0$ is given by (cf. (3.8))

$$
Lz = \begin{pmatrix} \dot{x} - y \\ \dot{y} - 2\dot{x} \end{pmatrix}, \quad z = (x, y).
$$

We now need a few facts from the Fredholm theory of linear differential operators $Lz = \dot{z} - A(\varepsilon)z$ which have the property that $\lim_{\varepsilon \to \varepsilon_0} A(\varepsilon) = A(\varepsilon)$
exists and has no eigenvalues on the imaginary axis (see Beyn, 1990a, Lemma 2.2, and also Palmer, 1984, Lemma 4.2; Hale, 1983).

(i) \( L : X_1 \rightarrow X_0 \) is Fredholm of index 0;
(ii) \( \dim N(L) = \dim N(L^*) \) where \( L^*z = \dot{z} + A(t)^Tz \);
(iii) \( z \in R(L) \iff \int_{-\infty}^{\infty} \Psi^T(t)z(t) \, dt = 0 \quad \forall \Psi \in N(L^*) \).

For the special case (4.7) we have
\[
N(L) = \text{span} \{\dot{z}\}, \quad N(L^*) = \text{span} \{(-\dot{y}, \dot{x})\}. \tag{4.8}
\]

Notice that \( \dot{z} \in N(L) \) follows by differentiating (3.8) with respect to \( t \). Furthermore, the second fundamental solution of \( L \) cannot be in \( X_1 \), for otherwise the Wronskian determinant tends to zero as \( t \rightarrow \infty \) in contradiction to Liouville's theorem. The representation of \( N(L^*) \) then follows easily from (ii).

Using (i) and the bordering lemma (Beyn, 1990a, Lemma 2.3) we find that \( F_w(\tilde{w}, \tau) : X_1 \times \mathbb{R} \rightarrow X_0 \times \mathbb{R} \) also has Fredholm index 0. Since \( \dot{\tilde{y}}(0) \neq 0 \) the only way that \( F_w(\tilde{w}, \tau) \) can have a non-trivial null space is in case
\[
g_e(\tilde{z}, 0, \tau) \in R(L).
\]
By (iii), (4.8) and (3.12) this is equivalent to a vanishing Melnikov integral
\[
0 = \int_{-\infty}^{\infty} (-\dot{\tilde{y}}, \dot{\tilde{x}})g_e(\tilde{z}, 0, \tau) \, dt
= \int_{-\infty}^{\infty} -\ddot{x}((a_1 \tilde{x}^2 + a_2 \tau \tilde{x} + a_3 \tau^2) + \tilde{x}b_1 \tilde{x} + b_2 \tau \tilde{x} \, dt
= \int_{-\infty}^{\infty} (a_1 + b_1)\tilde{x}^2 + \tau(a_2 + b_2)\tilde{x}^2 \, dt = 192(-a_1 + b_1) + \tau/5(a_2 + b_2).
\]

This is satisfied at \( \tau = \tau_0 \) and we obtain
\[
N(F_w(\tilde{w}, \tau_0)) = \text{span} \{(z_0, 1)\}, \quad R(F_w(\tilde{w}, \tau_0)) = R(L) \times \mathbb{R} \tag{4.9}
\]
where \( z_0 \) is the unique solution of (4.5).

The final condition for bifurcation from \( (\tilde{w}, \tau_0) \) is (cf. Crandall & Rabinowitz, 1971)
\[
F_w(\tilde{w}, \tau_0)\begin{pmatrix} z_0 \\ 1 \end{pmatrix} \notin R(F_w(\tilde{w}, \tau_0)).
\]
Using (4.9) and once more (iii) and (4.8) this turns out to be equivalent to
\[
0 \neq \int_{-\infty}^{\infty} (-\dot{\tilde{y}}, \dot{\tilde{x}})g_e(\tilde{z}, 0, \tau_0) \, dt = \int_{-\infty}^{\infty} -\ddot{x}(a_2 \tilde{x} + 2a_3 \tau_0) + \dot{x}b_2 \dot{x} \, dt = 192(a_2 + b_2)
\]
which is true by assumption.

Finally, (4.9) shows that we may parametrize the non-trivial branch such that it has tangent \((z_0, 1, \tau'(0))\) in \( X_1 \times \mathbb{R}^2 \).

In Theorem 4.1 we have analyzed the homoclinic orbits of the system (3.8) without using the transformations which led to this system. In fact, the special
transformation (3.7) induces a certain symmetry in (3.8) which yields $\tau'(0) = 0$, i.e. a vertical bifurcation. This idea was pointed out to me by M. Stiefenhofer (1991).

**Proposition 4.2** Assume that the system (3.8) is obtained from (3.3) with $r$ in $C^3$ by the transformation (3.7) and let the assumptions of Theorem 4.1 hold. Then the non-trivial branch of homoclinic orbits

$$(x(t, \varepsilon), y(t, \varepsilon), \varepsilon, \tau(\varepsilon)) = (x(t, \varepsilon), y(t, \varepsilon), \varepsilon, \tau(\varepsilon)), \quad t \in \mathbb{R}, |\varepsilon| < \varepsilon_0$$

satisfies

$$(x(-t, \varepsilon) = x(t, -\varepsilon), \quad y(-t, \varepsilon) = -y(t, -\varepsilon), \quad \tau(\varepsilon) = \tau(-\varepsilon). \quad (4.10)$$

In particular $\tau'(0) = 0$.

**Proof.** Let $\Xi(t, \xi_0, \mu)$ and $\Phi(t, \eta_0, \varepsilon, \tau)$ denote the $t$-flows of (3.3) and (3.8), respectively. We set

$$D_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad D_\varepsilon = \begin{pmatrix} \alpha_2 \varepsilon^2 & 0 \\ 0 & \alpha_4 \varepsilon^3 \end{pmatrix} \quad \text{for } \varepsilon \neq 0.$$

From (3.7) we find the relation

$$\Phi(s, \eta_0, \varepsilon, \tau) = D_\varepsilon^{-1}\Xi[(s/\alpha_5 \varepsilon), D_\varepsilon(\eta_0 + \alpha_4 \tau(1, 0)), (\alpha_0 + \alpha_1 \tau^2)\varepsilon^4, \varepsilon^2]$$

and hence

$$\Phi(s, D_0 \eta_0, -\varepsilon, \tau) = D_0 \Phi(-s, \eta_0, \varepsilon, \tau), \quad \varepsilon \neq 0.$$

Therefore, if $(x(t, \varepsilon), y(t, \varepsilon), \varepsilon, \tau(\varepsilon))$ is a homoclinic solution of (4.1) then so is $(x(-t, \varepsilon), -y(-t, \varepsilon), -\varepsilon, \tau(\varepsilon))$. Moreover, $\bar{x}$ is an even function and $\bar{y}$ is odd (see (3.12)), so both solutions are close to the bifurcation point $(\bar{w}, \tau_0)$ in $X_1 \times \mathbb{R}^2$. Our assertion (4.10) now follows from the uniqueness of the non-trivial branch. \qed

By Theorem 4.1 and Proposition 4.2 we see that starting homoclinic orbits for the system (3.8) is a matter of branch switching. The tangent approximation is given by

$$(z, \varepsilon, \tau) = (\bar{z} + \varepsilon z_0, \varepsilon, \tau_0). \quad (4.11)$$

Our aim is to fix a small $\varepsilon > 0$ and then try to solve (4.1) for $(z, \tau)$.

Actually, we were unable to solve (4.5) explicitly for $z_0$, so a simpler alternative is the non-tangent predictor

$$(z, \varepsilon, \tau) = (\bar{z}, \varepsilon, \tau_0). \quad (4.12)$$

In the next section we will show that this initial approximation is sufficient to guarantee convergence of a Newton or chord method for (4.1) (with $\varepsilon > 0$ fixed). For the tangent approximation (4.11) this follows from the theory of convergence cones near bifurcation in Keller (1977) and Jepson and Decker (1986).

For the numerical implementation we actually have to transform the predictors (4.11) and (4.12) back to the original equation (2.1). This involves the non-linear transformation (3.7), the centre manifold reduction and the linear transformation (2.11). A summary of the resulting algorithm will be given in Appendix A.
a homoclinic orbit of (2.1) has been found, we can use standard algorithms (see Doedel & Friedman, 1989; Beyn, 1990a; Rodriguez Luis et al 1990; Kuznetsov, 1990; Friedman & Doedel, 1992) to continue branches of homoclinics \((u(s), \lambda(s))\) in the two-parameter system (2.1).

We notice that we have not been able to prove convergence of the Newton process which starts at the predictor obtained by back-substitution of (4.12) (or (4.11)) in (2.1). Here, several technical difficulties arise. First, the non-linear transformation (3.7) does not commute with the Newton iteration. Second, it is not obvious how a convergent Newton process for a homoclinic of the reduced system (3.3) carries over to the full system (2.1). Clearly, the numerical experiments in Section 6 support a positive result and we expect some future progress in this direction. Generally speaking, it is of great interest to have theorems which transfer convergent iterations from a centre manifold reduction to the original system.

5. Convergence of Newton’s method near the singularity

In this section we show that the non-tangent predictor (4.12) is sufficient to generate a convergent Newton process for the operator equation (4.1). We use the fact that the bifurcating branch can be parametrized by \(\epsilon\) (Theorem 4.1) and rewrite (4.1) as

\[
H(v, \epsilon) = 0
\]

where \(v = (z, \tau) = (x, y, \tau) \in X_1 \times \mathbb{R}\) and

\[
H(v, \epsilon) = (\dot{z} - g(z, \epsilon, \tau), y(0)) \in X_0 \times \mathbb{R}.
\]

Equation (5.1) now has a simple bifurcation point at \(v_0 = (\bar{z}, \tau_0)\), \(\epsilon = 0\). For fixed \(\epsilon \neq 0\) we consider the iteration

\[
v_{n+1} = T(v_n, \epsilon), \quad n = 0, 1, 2, \ldots
\]

where either

\[
T(v, \epsilon) = v - H_0(v_0, \epsilon)^{-1}H(v, \epsilon)
\]

is the chord operator or

\[
T(v, \epsilon) = v - H_0(v, \epsilon)^{-1}H(v, \epsilon)
\]

is the Newton operator.

We employ the following abstract theorem on convergence wedges for these iterations. Its proof will be given in Appendix B. We use the superscript '0' to denote evaluation at \((v_0, 0)\).

**Theorem 5.1** Given \(H \in C^3(V \times \mathbb{R}, W)\) with Banach spaces \(V, W\) and a point \((v_0, 0) \in V \times \mathbb{R}\) with the following properties.

(i) \(H(v_0, 0) = 0\).

(ii) The operator \(H_0^0\) is Fredholm of index 0 with \(N(H_0^0) = \text{span} \{\varphi\}\) for
some $\varphi \neq 0$. Take splittings $V = N(H_0^u) \oplus V_1$, $W = R(H_0^u) \oplus W_1$ and let $P : V \to N(H_0^u)$, $Q : W \to W_1$ be the corresponding projectors.

(iii) $QH_0^0 \varphi \neq 0$.

(iv) $QH_0^e = 0$.

(v) $QH_0^v v = 0$ for all $v \in V$.

Then, for any $\alpha, \frac{1}{2} < \alpha < 1$, there exist $\epsilon_0 = \epsilon_0(\alpha) > 0$, $c_0 = c_0(\alpha) > 0$ such that

$$H(v, \epsilon) = 0, \quad 0 < \epsilon \leq \epsilon_0$$

has a unique solution $v(\epsilon)$ in

$$K_\epsilon = \{ v \in V : ||P(v - v_0)|| \leq \epsilon^\alpha, ||(I - P)(v - v_0)|| \leq c_0 \epsilon \}.$$

Moreover, the chord and the Newton iteration started at $v_0$ (or somewhere else in $K_\epsilon$) converge to $v(\epsilon)$ for any $0 < \epsilon \leq \epsilon_0$.

REMARKS

1. The assumptions of Theorem 5.1 imply that $(v_0, 0)$ is a simple bifurcation point. By (i), (ii) and (iv) we have $H_0^v \in R(H_0^u)$ and

$$N(H_0^u) = \text{span}\{\varphi, 0\}, \quad (\xi, 1)\}$$

where $H_0^0 \xi = -H_0^v$. The quadratic form induced by $H^0$ on $N(H_0^u)$ is given by

$$q(x, y) = \frac{1}{2} \alpha x^2 + \beta xy + \frac{1}{2} \gamma y^2$$

where $\alpha = QH_0^0 v \varphi^2 = 0$ by (v), $\beta = QH_0^v \varphi \neq 0$ by (iii) and $\gamma = QH^0(\xi, 1)^2$. Therefore, the tangents to the bifurcating branches are

$$(\varphi, 0) \quad \text{and} \quad (\xi, 1) - \frac{\gamma}{2\beta} (\varphi, 0).$$

In general, $\gamma$ and $\xi$ do not vanish, so that $(0, 1)$ is not tangent to a branch. Nevertheless, the initial value

$$(v_0, \epsilon) = (v_0, 0) + \epsilon(0, 1)$$

leads to a convergent iteration. The main reason for this fact is assumption (v) which requires more than just $QH_0^v \varphi^2 = 0$.

2. Gluing the balls $K_\epsilon$ together gives us the domain of convergence

$$W = \{ (v, \epsilon) : v \in K_\epsilon, 0 < \epsilon \leq \epsilon_0 \}$$

which is illustrated in Fig. 2. It has the shape of a wedge.

Such a wedge seems neither to be covered by the convergence cones for parameter-dependent problems (Keller, 1977; Jepson & Decker, 1986) nor by the wedges for the parameter-independent case in Decker et al (1983).

Let us apply this theorem to equations (5.1), (5.2). We simply have to reinterpret the results from Section 4 with the roles of $\tau$ and $\epsilon$ interchanged.
From the proof of Theorem 4.1 we obtain
\[ H(v_0, 0) = (\ddot{x} - g(\bar{z}, 0, \tau_0), \bar{y}(0)) = 0 \]
\[ H^0_\nu(z, \tau) = (Lz, y(0)) \quad \text{with } L \text{ from (4.7)}, \]
\[ N(H^0_\nu) = \text{span} \{ \varphi \} \quad \text{with } \varphi = (0, 0, 1) \in X_1 \times \mathbb{R}, \]
\[ R(H^0_\nu) = \{(z, s) \in X_0 \times \mathbb{R} : \Psi(z, s) = 0\}, \quad \Psi(z, s) = \int_{-\infty}^{\infty} (-\dot{y}, \dot{x})z \, dt, \]
\[ H^0_{\nu e} \varphi \notin R(H^0_\nu) \quad \text{since} \]
\[ \Psi(H^0_{\nu e} \varphi) = \int_{-\infty}^{\infty} \dot{x}(a_2 + 2a_3 \tau_0) - \dot{\bar{z}}b_2 \dot{x} \, dt = -(a_2 + b_2) \int_{-\infty}^{\infty} \dot{x}^2 \, dt \neq 0, \]
\[ \Psi(H(v_0, \epsilon)) = O(\epsilon^2) - \epsilon \int_{-\infty}^{\infty} (-\ddot{y}, \ddot{x})g_\epsilon(\bar{z}, 0, \tau_0) \, dt = O(\epsilon^2), \]
\[ H^0_{\nu v} \varphi \nu = \left( -g_{	au z}(\bar{z}, 0, \tau_0)z - g_{\tau r}(\bar{z}, 0, \tau_0)\tau, 0 \right) = 0 \quad \text{for all } \nu = (z, \tau) \in X_1 \in \mathbb{R}. \]

Thus we have shown the following.

**Corollary 5.2** For the two-parameter system (3.8) assume that \( g \) is in \( C^3 \) and let \( a_2 + b_2 \neq 0 \). Then for any \( \epsilon > 0 \) sufficiently small, Newton's method for solutions \( (x, y, \tau) \in X_1 \times \mathbb{R} \) of the system
\[
\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = g(x, y, \epsilon, \tau), \quad y(0) = 0 \tag{5.6}
\]
converges to the non-trivial homoclinic orbits \((x(\epsilon), y(\epsilon), \tau(\epsilon))\) of Theorem 4.1 when started at \((\bar{x}, \bar{y}, \tau_0)\).
Of course, Newton's method in the Banach space \( X \times \mathbb{R} \) is not a practical method. Also, the phase condition \( y(0) = 0 \) can be replaced by a more efficient integral condition. Further details on the numerical implementation are contained in Appendix A and the next section.

6. Examples and numerical calculations

Our algorithm gives us an explicit approximation of homoclinic orbits close to the TB-point \((\lambda^0, \lambda^0)\) (see Appendix A). This approximation depends on a parameter \( \varepsilon \) as follows

\[
\tilde{u}(t) = u^0 + \xi_1(t, \varepsilon) v + \xi_2(t, \varepsilon) w + D \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad D \in \mathbb{R}^{n, 2}, \quad v, w \in \mathbb{R}^n \tag{6.1a}
\]

\[
\tilde{\lambda} = \lambda^0 + B \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad B \in \mathbb{R}^{2, 2} \tag{6.1b}
\]

\[
\xi_1(t, \varepsilon) = (\varepsilon^2/2) \left[ 1 - 3 \text{sech}^2 \left( (\varepsilon/2)t \right) - Q_{2, 14} \right] \tag{6.1c}
\]

\[
\xi_2(t, \varepsilon) = \left( 3 \varepsilon^2 / Q_{2, 11} \right) \text{sech}^2 \left( (\varepsilon/2)t \right) \tanh (\varepsilon/2 t). \tag{6.1d}
\]

We then fix one parameter, say \( \lambda_1 = \tilde{\lambda} \), and solve a well-posed boundary value problem for \((u, \lambda_2) \in X \times \mathbb{R}\) (cf. Doedel & Friedman, 1989; Beyn, 1990a 1990b; and Section 5).

\[
\dot{u} = f(u, \tilde{\lambda}_1, \lambda_2), \quad \dot{\lambda}_2 = 0 \quad \text{for } t \in \mathbb{R} \tag{6.2a}
\]

\[
\sigma(u) = 0 \tag{6.2b}
\]

where \( \sigma : X \rightarrow \mathbb{R} \) defines a phase condition. For example, the phase condition \( y(0) = 0 \), which was suitable for the \((x, y)\) system (5.1), (5.2), now transforms into (cf. (6.1), (2.8))

\[
\sigma(u) = \Psi^T \begin{pmatrix} u(0) - u^0 - D \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \end{pmatrix} = 0. \tag{6.3}
\]

A more reliable choice is an integral condition (see Doedel & Kernévez, 1986)

\[
\sigma(u) = \int_{-\infty}^{\infty} \dot{u}(t)^T (u(t) - \tilde{u}(t)) \, dt = 0. \tag{6.4}
\]

In the following examples we have concentrated on the local behaviour of the starting algorithm, i.e. the initial homoclinic orbit \((u, \lambda_2)\) was computed and it was then continued by slightly varying \( \lambda_1 \).

Of course, for the global bifurcation diagram one should apply the usual continuation idea of stepping forward tangentially in the \( \lambda \)-plane (or the \( \mu \)-plane with \( \mu \) as in (6.1b)) and then correcting in the normal direction.

As in Beyn (1990a) we replaced (6.2) by a boundary value problem on a finite interval \( J = [T_-, T_+] \) namely

\[
\dot{u} = f(u, \tilde{\lambda}_1, \lambda_2), \quad \dot{\lambda}_2 = 0 \quad \text{in } J \tag{6.5a}
\]

\[
\sigma_\varepsilon(u) = 0 \tag{6.5b}
\]

\[
B_-(u(T_-), \lambda_2) = 0, \quad B_+(u(T_+), \lambda_2) = 0. \tag{6.5c}
\]
Here, $\sigma_r$ denotes a finite phase condition, e.g. the point condition (6.3) or the truncated integral condition

$$\sigma_r(u) = \int_{T_r}^{T_*} \dot{u}(t)^T (u(t) - \bar{u}(t)) \, dt.$$  

(6.6)

In (6.5c) we employ projection boundary conditions (see Beyn, 1990a)

$$B_-(v, \lambda_2) = L_u(\lambda_2) (v - u(\lambda_2)) \in \mathbb{R}^n,$$

$$B_+(v, \lambda_2) = L_v(\lambda_2) (v - u(\lambda_2)) \in \mathbb{R}^n.$$  

where $v(\lambda_2)$ are the saddles of

$$\dot{u} = f(u, \bar{x}, \lambda_2)$$  

(6.7)

close to $\bar{u}(-\infty) = \bar{u}(+\infty)$. Moreover, the rows of $L_u(\lambda_2) \in \mathbb{R}^{n \times n}$ (resp. $L_v(\lambda_2) \in \mathbb{R}^{n \times n}$) form a basis of the stable (resp. unstable) subspace of $f_u'(v(\lambda_2), \bar{x}, \lambda_2)$. The choice of $T_-, T_+$ was guided by the behaviour of $\xi_1(t, \epsilon)$, which is mainly responsible for the amplitude of the initial homoclinic orbit. To be more specific, $T_+, T_- \in [0, 1]$ were determined from

$$|\xi_1(\infty, \epsilon) - \xi_1(T_*, \epsilon)| = \frac{3\epsilon^2}{|Q_{2,11}|} \text{sech}^2 \left( \frac{\epsilon}{2} T_+ \right) = \delta_0$$

where $\delta_0$ is a prescribed small quantity.

For all calculations shown below we used the integral phase condition (6.6) and some standard transformations in order to write (6.5) as a two-point boundary value problem on $[0, 1]$ (see Beyn, 1990a). The solver DO2RAF from the NAG-library, Oxford was applied.

**EXAMPLE 1** $\dot{u}_1 = u_2$, $\dot{u}_2 = \lambda_1 + \lambda_2 u_2 - u_1^2 + u_1 u_2$.

This is the model example from (1.2). With $\epsilon = 0.1$, $\delta_0 = 10^{-4}$ and $\bar{x}_1 = 0.25 \times 10^{-4}$ the b.v.p. (6.5) could be solved easily. Then $\lambda_1$ was increased and a branch of homoclinic orbits was calculated up to $\lambda_1 = 0.35 \times 10^{-2}$. Figure 3 shows the evolving homoclinics in phase space. During the continuation the adaptation

![Fig. 3. Branch of homoclinic orbits in phase space for Example 1 ($\lambda_1 = 0.25 \times 10^{-4}, ..., 0.35 \times 10^{-2}$).](image)
strategy for $[T_-, T_+]$, as developed in Beyn (1990b), was used. Figures 4(a), (b) show the time diagrams for the $u_1$- and $u_2$-components while Fig. 5 displays the variation of $\lambda_2$, $T_-$, $T_+$ with increasing $\lambda_1$.

**Example 2** (Nerve excitation (Chay & Keizer, 1983; Rinzel, 1985)).

This two-dimensional system models the time course of voltage oscillations ($V$) across the membrane of pancreatic $\beta$-cells. It is a modification by Chay and Keizer (1983) of the Hodgkin–Huxley model, further simplified by Rinzel (1985). The parameters are $\lambda_1 = C_a \, [\mu M]$ (intracellular calcium concentration) and $\lambda_2 = T[\circ C]$ (the temperature).

The dimensionless equations are

$$C_M \dot{V} = 2\bar{g}_{Ca,HH} m^3_n(V) h_m(V)(C_a - V) + (\bar{g}_{K,HH} n^4 + \bar{g}_{K,ca} \frac{C_a}{1 + C_a})(V_k - V) + \bar{g}_L(V_L - V)$$

$$\dot{n} = \Phi(\alpha_n(V)(1 - n) - \beta_n(V)n)$$

![Figure 4](image1.png)

**Fig. 4.** (a), (b): Time diagrams of $u_1$, $u_2$ for the homoclinic orbits from Fig. 3.

![Figure 5](image2.png)

**Fig. 5.** Dependence of $\ln|T_+|$, $\ln(T_-)$, $\ln|\lambda_2|$ on the continuation parameter $\lambda_1$ for Example 1.
where
\[ C_M = 1, \quad \dot{g}_{C_a,HH} = 6.5, \quad \dot{g}_{K,HH} = 12, \quad \dot{g}_K,ca = 0.09, \quad \dot{g}_L = 0.04, \]
\[ V_{C_a} = 100, \quad V_K = -75, \quad V_L = -40, \quad \Phi = 3^{(T-6.3)^{10}}, \]
\[ m_V(V) = \frac{\alpha_m(V)}{\alpha_m(V) + \beta_m(V)}, \quad h_s(V) = \frac{\alpha_s(V)}{\alpha_s(V) + \beta_s(V)}. \]
\[ \alpha_m(V) = \tilde{\epsilon}(\frac{-25 - V}{10}), \quad \alpha_n(V) = \tilde{\epsilon}(\frac{-20 - V}{10}), \quad \tilde{\epsilon}(x) = \frac{x}{e^x - 1}, \]
\[ \alpha_s(V) = 0.07 \exp\left(\frac{-50 - V}{20}\right), \quad \beta_s(V) = \left(1 + \exp\left(\frac{-20 - V}{10}\right)\right)^{-1} \]
\[ \beta_m(V) = 4 \exp\left(\frac{-50 - V}{18}\right), \quad \beta_n(V) = 0.125 \exp\left(\frac{-30 - V}{80}\right). \]

We found a TBP-point at \((\lambda_1, \lambda_2) = (1.88, 14.94), (V, n) = (-34.21, 0.26)\). We did some experiments on the quality of the predicted homoclinic (6.1) with varying values of \(\epsilon\) (see also Rodriguez Luis et al, 1990). Figure 6 compares the predicted values \(\lambda_2 = \lambda_2(\epsilon)\) with the computed ones \(\lambda_2(\epsilon)\). Note that \(\lambda_2(\epsilon)\) is obtained by solving (6.5) with \(\lambda_1 = \lambda_1(\epsilon)\) fixed. The predictor was found to be reliable up to \(\epsilon = 0.3\). For this value, Fig. 7 compares the predicted voltage function with its finally corrected form.

**Example 3** (CO-oxidation (Bykov et al, 1978; Khibnik, 1990))

\[
\begin{align*}
\dot{x} &= 2k_1 z^2 - 2k_3 x^2 - k_3 x y, \\
\dot{y} &= k_2 z - k_2 x y - k_3 x y, \\
\dot{z} &= k_4 \left(z - \frac{k_4}{k_4} s\right).
\end{align*}
\]

**Fig. 6.** Comparison of parameter values \(\lambda_2(\epsilon)\), predicted by the starting procedure, with computed values \(\lambda_2(\epsilon)\) (Example 2).
The underlying reaction scheme is displayed in Bykov et al (1978) and notice that a factor 2 is missing in front of $k_1 z^2$ in Khibnik (1990). Most of the reaction constants above were fixed (see Khibnik, 1990)

$$k_1 = 2.5, \quad k_{-1} = 1, \quad k_3 = 10, \quad k_{-2} = 0.1, \quad k_4 = 0.0675$$

and a TB-point was calculated at

$$(x, y, s) = (0.1159, 0.31547, 0.28844)$$

$$\lambda_1 = k_2 = 1.4176, \quad \lambda_2 = k_{-4}/k_4 = 0.9714.$$
The emanating homoclinic orbits for a few values of $\epsilon$ are shown in Fig. 8.

In this 3-dimensional example we also measured the deviation of the actual homoclinic orbit $\tilde{u}(t, \epsilon)$ for $\lambda_1 = \lambda_1(\epsilon)$ from the linear approximation to the centre manifold (see Fig. 9) by calculating

$$d(\epsilon) = \sup_{t \in [\tau_-, \tau_+]} \left| \gamma^T \left( \tilde{u}(t, \epsilon) - u^0 - D \left( \mu_1 \right) \right) \right|.$$

Here $\gamma \in \mathbb{R}^3$ spans $\{v, w\}^\perp$ (see (6.1a)) and it has to be a left eigenvector corresponding to the non-trivial eigenvalue of $f_u(u^0, \lambda^0)$.

**Appendix A. Summary of the algorithm**

### A.1 Computation of a TB-point

Let $(u^0, \lambda^0)$ be a solution of the defining equation (2.3) obtained by a Newton iteration. As in (2.2) let

$$A^0 = \begin{pmatrix} f_u(u^0, \lambda^0) & b_0 \\ c_0^T \\ 0 \end{pmatrix} \in \mathbb{R}^{n+1, n+1}.$$

Generally, we use the upper index '0' to denote evaluation at $(u_0, \lambda_0)$. However, we try to suppress this index whenever it does not affect the readability. Compute $v, w, \Psi, \xi \in \mathbb{R}^n$ from the linear systems:

$$A^0(v) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad A^0(w) = \begin{pmatrix} v \\ 0 \end{pmatrix}, \quad A^0r(\Psi) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad A^0r(\xi) = \begin{pmatrix} \Psi \\ 0 \end{pmatrix}.$$

Normalize

$$\alpha \leftarrow \Psi^Tw; \quad \xi \leftarrow \frac{1}{\alpha} \left( \xi - \frac{1}{\alpha} (\xi^Tw) \Psi \right); \quad \Psi \leftarrow \frac{1}{\alpha} \Psi.$$
A.2 Linear normal form

\[ \beta_i \leftarrow \Psi^T f_i^0, \quad i = 1, 2, \]

\[ B \leftarrow (\beta_1^2 + \beta_2^2)^{-1} \begin{pmatrix} \beta_1 & -\beta_2 \\ \beta_2 & \beta_1 \end{pmatrix} \in \mathbb{R}^{2,2} \]

\[ C \leftarrow (w \ 0) - f_2^0 B \in \mathbb{R}^{n,2}. \]

Solve the two linear systems

\[ A^0 \begin{pmatrix} \frac{D}{d \tau} \\ 0 \end{pmatrix} = \begin{pmatrix} C \\ 0 \end{pmatrix}. \]

Let \( \delta_2 \in \mathbb{R}^{n+2} \) be the second column of \( \begin{pmatrix} D \\ B \end{pmatrix} \).

A.3 Coefficients of quadratic terms (equations (2.18), (3.5))

We assume that \( f_2 = (f_u, f_s) \in \mathbb{R}^{n,n+2}, \ z = (u, \lambda) \) can be evaluated explicitly and we choose a suitable stepsize \( s > 0 \) for numerical differentiation

\[ u_1 \leftarrow (1/s)(f_u(u^0 + sv, \lambda^0) - f_u^0)v; \]

\[ u_4 \leftarrow (1/s)(f_u(u^0 + sv, \lambda^0) - f_u^0)\delta_2; \]

\[ Q_{1,11} \leftarrow \xi^T u_1; \quad Q_{2,11} \leftarrow \Psi^T u_1; \quad Q_{1,14} \leftarrow \xi^T u_4; \quad Q_{2,14} \leftarrow \Psi^T u_4; \]

\[ Q_{2,12} \leftarrow (1/s)\Psi^T (f_u(u^0 + sv, \lambda^0) - f_u^0)w; \]

\[ Q_{2,24} \leftarrow (1/s)\Psi^T [f_u((u^0, \lambda^0) + s\delta_2) - f_u^0]w; \]

\[ Q_{2,44} \leftarrow (1/s)\Psi^T [f_s((u^0, \lambda^0) + s\delta_2) - f_s^0]\delta_2. \]

Notice that we need only 3 evaluations of \( f_z \) for this step. Also, we have not used the equalities \( f_u^0 v = 0, f_s^0 w = v \) since they will only be true approximately in practice.

A.4 The approximate homoclinic of the reduced system (3.3)

\[ \Delta \leftarrow Q_{2,11}(Q_{1,14} + Q_{2,24}) - Q_{2,14}(Q_{1,11} + Q_{2,12}) \]

\[ \tau_0 \leftarrow \frac{5}{7\Delta} (Q_{1,11} + Q_{2,12}); \quad \sigma_0 \leftarrow \frac{1}{2Q_{2,11}} \left((Q_{1,14}^2 - Q_{2,11}Q_{2,44})\tau_0^2 - 1\right); \]

choose an \( \varepsilon > 0 \) and let

\[ \mu_1 \leftarrow \sigma_0 \varepsilon^4; \quad \mu_2 \leftarrow \tau_0 \varepsilon^2 \]

\[ \xi_1(t) \leftarrow \frac{\varepsilon^2}{Q_{2,11}} \left(1 - 3 \text{sech}^2\left(\frac{\varepsilon}{2} t\right) - Q_{2,14}\tau_0\right) \]

\[ \xi_2(t) \leftarrow \frac{3\varepsilon^3}{Q_{2,11}} \text{sech}^2\left(\frac{\varepsilon}{2} t\right) \tanh\left(\frac{\varepsilon}{2} t\right) \].
A.5 Approximate homoclinic for the original system (2.1)

\[ \ddot{u}(t) = u^0 + \xi_1(t)u + \xi_2(t)w + D \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \]

\[ \lambda = \lambda^0 + B \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}. \]

Here we trace back the centre manifold reduction from (3.1) to (3.3) and the linear transformation (2.11).

Appendix B. Proof of Theorem 5.1

First, we consider the chord operator

\[ T(v, \epsilon) = v - H_\epsilon(v_0, \epsilon)^{-1}H(v, \epsilon) \]

and show that \( T(\cdot, \epsilon) \) is a contraction on \( K_\epsilon \) in a suitable norm. Later on, we will indicate the necessary changes for Newton's method. It will be convenient to work with generalized norms (or distances) \( |\cdot| : V \to \mathbb{R}^m \), i.e.

\[ |v| \geq 0, \quad |v| = 0 \iff v = 0, \quad |v_1 + v_2| \leq |v_1| + |v_2| \]

holds with the natural partial ordering of \( \mathbb{R}^m \) (cf. Bohl, 1978). Each norm \( ||\cdot||_* \) in \( \mathbb{R}^m \) then leads to a norm \( ||\cdot|| \) in \( V \) via

\[ ||v|| = |||v|||_*. \]

and any two norms of this type are equivalent. Without loss of generality we will therefore work in \( V \) with the norm

\[ ||v|| = \sum_{i=1}^m |v_i|. \]

Our choice for the proof of Theorem 5.1 is \( m = 2 \) with

\[ |v| = \left( \frac{||Pv||}{||(I-P)v||} \right), \quad v \in V; \]

\[ |w| = \left( \frac{||Qw||}{||(I-Q)w||} \right), \quad w \in W. \quad (B1) \]

We may then write

\[ K_\epsilon = \{ v \in V : |v - v_0| \leq r_\epsilon \}, \quad r_\epsilon = \left( \frac{\epsilon^\alpha}{c_0 \epsilon} \right). \]

We use the following version of the contraction mapping theorem (cf. Bohl, 1978).

**Theorem B1** Let \( (V, |\cdot|) \) be a Banach space with a generalized norm \( |\cdot| \). Given a ball

\[ B = \{ v \in V : |v - v_0| \leq r \}, \quad \text{where } r \in \mathbb{R}^m, \quad r > 0, \]

\[ ||v|| = \sum_{i=1}^m |v_i|. \]
an operator $T : B \rightarrow V$ and a non-negative matrix $L \in \mathbb{R}^{m \times m}$ with the following properties:

$$|T(v) - T(y)| \leq L |v - y| \quad \text{for all } v, y \in B \quad (B2)$$

$$|T(v_0) - v_0| < (I - L)r. \quad (B3)$$

Then $T$ has a unique fixed point $v$ in $B$ and for any $y_0 \in B$ the sequence $y_{n+1} = T(y_n)$ converges to $v$. Moreover,

$$|v - y| \leq (I - L)^{-1} |v - T(v) - (y - T(y))| \quad \forall v, y \in B. \quad (B4)$$

**Remark** By (B3) and $r > 0$ we have that $I - L$ is an $M$-matrix. Moreover, it is easy to see from (B2) and (B3) that $T$ maps $B$ into itself and is an ordinary contraction with respect to the norm

$$||v||_r = \max \left\{ \left| \frac{v_i}{r_i} \right| : i = 1, \ldots, m \right\}. \quad (B5)$$

Our second aid is a careful estimate of the inverses $H_v(v_0, \epsilon)^{-1}$. It can be extracted from Jepson and Decker (1986), but for the sake of completeness we will give a short proof in terms of generalized norms later on.

**Lemma B2** Under the assumptions of Theorem 5.1 there exists an $\epsilon_0 > 0$ and a $C^* > 0$ such that $H_v(v_0, \epsilon)$ is a homeomorphism for all $0 < \epsilon \leq \epsilon_0$. The following estimate holds with respect to the generalized norms from (B1)

$$|H_v(v_0, \epsilon)^{-1}v| \leq C^* \begin{bmatrix} \epsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix} |v| \quad \forall v \in V. \quad (B6)$$

The estimate (B5) shows that the singularity of $H_v(v_0, \epsilon)$ near $\epsilon = 0$ only involves the critical subspaces $R(P)$ and $R(Q)$. We apply Theorem B1 to the chord operator $T(\cdot, \epsilon)$ with the ball $B = K_\epsilon$. In what follows, $C$ will denote a generic constant independent of $\epsilon$.

From our assumptions (i), (iv) and Lemma B2 we obtain

$$|T(v_0, \epsilon) - v_0| = |H_v(v_0, \epsilon)^{-1}H(v_0, \epsilon)| \leq C \begin{bmatrix} \epsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \epsilon^2 \\ \epsilon \end{bmatrix} \leq C \left( \begin{bmatrix} \epsilon \end{bmatrix} \right).$$

With this constant $C$ we now fix $c_0$ in $K_\epsilon$ and $\epsilon_1 > 0$ through

$$c_0 = 2C \quad \text{and} \quad c_0 \epsilon \leq \epsilon^a \quad \text{for } 0 < \epsilon \leq \epsilon_1. \quad (B7)$$

Then the following holds

$$|T(v_0, \epsilon) - v_0| \leq \frac{1}{2} r_\epsilon \quad \text{for } 0 < \epsilon \leq \epsilon_1. \quad (B8)$$

Next we consider the Lipschitz estimate for $v, y \in K_\epsilon$

$$|T(v, \epsilon) - T(y, \epsilon)| = |H_v(v_0, \epsilon)^{-1}[H_v(v_0, \epsilon)(v - y) - (H_v(v, \epsilon) - H(y, \epsilon))]|$$

$$\leq C^* \begin{bmatrix} \epsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix} \int_0^1 [H_v(v_0, \epsilon) - H_v(y + \epsilon(v - y), \epsilon)] dt (v - y)$$

$$= C^* \begin{bmatrix} \epsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix} |z|.$$
where
\[ z = \int_0^1 \int_0^1 H_{vv} \left[ (1 - s)v_0 + s(y + t(v - y)), \varepsilon \right] (y + t(v - y) - v_0)(v - y) \, ds \, dt. \]

The key idea is to use assumption (v) and \( H \in C^3 \) in order to estimate \( Qz \).

\[
\|Qz\| \leq \left\| Q \int_0^1 \int_0^1 \left[ H_{vv} ((1 - s)v_0 + s(y + t(v - y)), \varepsilon) - H_{vv}^0 \right]
\times (y + t(v - y) - v_0)(v - y) \, ds \, dt \right\|
\]

\[
+ \left\| Q \int_0^1 \int_0^1 H_{vv}^0 \, ds(y + t(v - y) - v_0)(v - y) \right\|
\]

\[
\leq C(||v - v_0|| + ||y - v_0|| + \varepsilon(||v - v_0|| + ||y - v_0||) ||v - y||
+ ||QH_{vv}^0 \left( \frac{1}{2} (v + y) - v_0 \right)(v - y) ||
\]

\[
\leq C \varepsilon^2 ||v - y|| + \|QH_{vv}^0 ((1 - P)(\frac{1}{2} (v + y) - v_0))(v - y) ||
\]

\[
\leq C \{ \varepsilon^2 (||P(v - y)|| + ||(1 - P)(v - y)||) + \varepsilon \| ||(1 - P)(v - y)|| \}
\]

\[
\leq C \{ \varepsilon^2 (||P(v - y)|| + \varepsilon \| ||(1 - P)(v - y)||) \}.
\]

Furthermore,

\[
\|(I - Q)z\| \leq C(||v - v_0|| + ||y - v_0||) ||v - y||
\]

\[
\leq C \varepsilon^2 (||P(v - y)|| + ||(1 - P)(v - y)||).
\]

Summarizing these estimates we obtain

\[
|T(v, \varepsilon) - T(y, \varepsilon)| \leq C \left[ \begin{array}{c}
\varepsilon^{-1} \\
1
\end{array} \right] \left[ \begin{array}{c}
\varepsilon^{2 \alpha} \\
\varepsilon^\alpha
\end{array} \right] ||v - y|| \leq L_\varepsilon ||v - y||
\]

where

\[
L_\varepsilon = C \left[ \begin{array}{c}
\varepsilon^{2 \alpha - 1} \\
\varepsilon^\alpha
\end{array} \right].
\]

With this Lipschitz bound we calculate

\[
(I - L_\varepsilon) r_\varepsilon = \left( \begin{array}{c}
\varepsilon^\alpha - C \varepsilon^{2 \alpha - 1} - C_0 \varepsilon \\
C_0 \varepsilon - C \varepsilon^{2 \alpha} - C_0 \varepsilon^{\alpha + 1}
\end{array} \right).
\]

Thus, for \( \varepsilon \) sufficiently small, (B7) yields

\[
(I - L_\varepsilon) r_\varepsilon \geq \frac{3}{2} r_\varepsilon > |T(v_0, \varepsilon) - v_0|.
\]

An application of Theorem B1 finishes the proof.

**Remark** From (B4) and (B5) we obtain

\[
|v - y| \leq C \left[ \begin{array}{c}
1 \\
\varepsilon^\alpha
\end{array} \right] |H_{vv}(v_0, \varepsilon)^{-1}(H(v, \varepsilon) - H(y, \varepsilon))|
\]

\[
\leq C \left[ \begin{array}{c}
\varepsilon^{-1} \\
1
\end{array} \right] |H(v, \varepsilon) - h(y, \varepsilon)|
\]

for \( v, y \in K_\varepsilon \).  \hspace{1cm} (B8)

This is the non-linear generalization of (B5).
Proof of Lemma B2. Let us approximate $H_\varepsilon(v_0, \varepsilon)$ by the linear block-diagonal operator $A(\varepsilon) := \varepsilon Q H_0^0 P + (I - Q) H_0^0 (I - P)$. Expanding $H_\varepsilon(v_0, \varepsilon)$ with respect to $\varepsilon$ it is easily seen that

$$|A(\varepsilon) - H_\varepsilon(v_0, \varepsilon)| \leq C_2 \begin{bmatrix} \varepsilon^2 & \varepsilon \\ \varepsilon & \varepsilon \end{bmatrix} |v| \quad \forall v \in V.$$ 

By assumptions (ii) and (iii) of Theorem 5.1 we have for some $C_1 > 0$

$$C_1 \begin{bmatrix} \varepsilon & 0 \\ 0 & 1 \end{bmatrix} |v| \leq |A(\varepsilon)v| \quad \text{for all } v \in V.$$ 

Both inequalities together imply for $v \in V$

$$|H_\varepsilon(v_0, \varepsilon)v| \geq M(\varepsilon) |v|, \quad M(\varepsilon) = \begin{bmatrix} C_1 \varepsilon - C_2 \varepsilon^2 & -C_2 \varepsilon \\ -C_2 \varepsilon & C_1 - C_2 \varepsilon \end{bmatrix}. \quad (B9)$$

For $\varepsilon$ sufficiently small, $M(\varepsilon)$ is an $M$-matrix and

$$M(\varepsilon)^{-1} \leq C \begin{bmatrix} \varepsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix}. \quad (B10)$$

Therefore, (B9) yields

$$|v| \leq C \begin{bmatrix} \varepsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix} |H_\varepsilon(v_0, \varepsilon)v| \quad \forall v \in V.$$ 

$H_\varepsilon(v_0, \varepsilon)$ is a small perturbation of $H_0^0$, hence it is also of Fredholm index 0. Inequality (B10) then shows that $H_\varepsilon(v_0, \varepsilon)$ is a homeomorphism which satisfies (B5).

Note on Newton’s method

We will only indicate the major changes necessary for the Newton operator

$$T(v, \varepsilon) = v - H_\varepsilon(v, \varepsilon)^{-1} H(v, \varepsilon), \quad v \in K_\varepsilon.$$ 

First, we need an analogue of Lemma B2

$$|H_\varepsilon(y, \varepsilon)^{-1} v| \leq C \begin{bmatrix} \varepsilon^{-1} & 1 \\ \varepsilon^{\sigma^{-1}} & 1 \end{bmatrix} |v| \quad \text{for } y \in K_\varepsilon, \quad v \in V.$$ 

To see this, we replace $v$ by $y + tv$ in (B8) and let $t \to 0$.

For the second step we utilize the known solutions $v(\varepsilon) \in K_\varepsilon$ and show

$$|T(v, \varepsilon) - T(v(\varepsilon), \varepsilon)| \leq Q_\varepsilon |v - v(\varepsilon)|^2, \quad v \in K_\varepsilon, \quad (B11)$$

where $Q_\varepsilon : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ is bilinear and satisfies

$$Q_\varepsilon \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = C \begin{pmatrix} \varepsilon^{\sigma^{-1}} x_1^2 + \varepsilon^{-1} x_2^2 \\ x_1^2 + \varepsilon^{\sigma^{-1}} x_2^2 \end{pmatrix}.$$
Similar to the chord method the crucial estimate uses assumption (v).

\[ \|Q[H(v(e), \varepsilon) - (H(v, \varepsilon) + H_u(v, \varepsilon)(v(\varepsilon) - v))]| \]
\[ \leq \left| Q \int_0^1 (1-t)[H_{uv}(v + t(\varepsilon - v)) - H^0_{uv}(v(\varepsilon) - v)^2 dt \right| \]
\[ + \frac{1}{2} \|H^0_{uv}[(I-P)(v(\varepsilon) - v)]\| \]
\[ \leq C(\varepsilon^\alpha \|P(v(\varepsilon) - v)\|^2 + \|(I-P)(v(\varepsilon) - v)\|^2). \]

From (B11) it is straightforward to choose \( \varepsilon \) sufficiently small so that

\[ |T(v, \varepsilon) - T(v(\varepsilon), \varepsilon)| \leq C \left[ \varepsilon^{2\alpha-1} \frac{1}{\varepsilon^\alpha} \right] |v - v(e)|, \quad v \in K_\varepsilon \]

holds. This proves convergence of Newton's method.

We notice that (B11) still signals some kind of quadratic convergence. However, due to the dependence of \( Q_\varepsilon \) on \( \varepsilon \) we have to take an \( \varepsilon \)-dependent norm for this purpose. For example,

\[ \|v\|_\varepsilon = [\varepsilon^\rho \|Pv\|^2 + (\varepsilon^\rho \|(I-P)v\|)^2]^{1/2}, \]

with \( \rho = \min(2(\alpha - 1), -\frac{3}{2}) \) as a possible choice.

Acknowledgements

I would like to thank the SERC (Great Britain) for supporting my visit to the University of Bath in 1989, during which this research was initiated. I am particularly indebted to A. Spence (University of Bath) for providing this opportunity and for stimulating discussions on this project.

References


