IMA Journal of Numerical Analysis (1994) **14**, 381–410

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

$$\dot{u} = f(u, \lambda), \qquad u \in \mathbb{R}^n, \qquad \lambda \in \mathbb{R}^2. \tag{1.1}$$

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

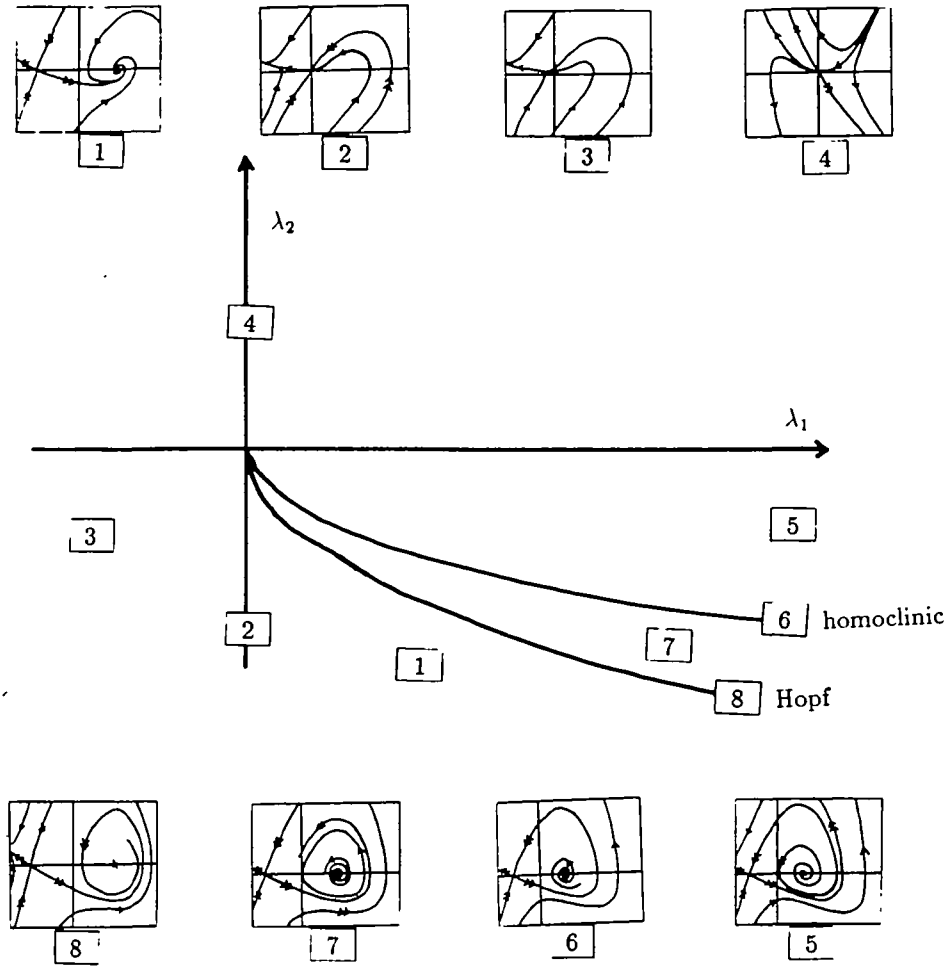

FIG. 1. Unfolding of Takens–Bogdanov point.

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)

$$\dot{u}_1 = u_2, \qquad \dot{u}_2 = \lambda_1 + \lambda_2 u_2 - u_1^2 + u_1 u_2. \tag{1.2}$$

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 $\lambda_1 = (\frac{7}{5}\lambda_2)^2$, $\lambda_2 < 0$.

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)

$$\dot{u}_1 = u_2, \qquad \dot{u}_2 = \lambda_1 u_1 + \lambda_2 u_2 - u_1^2 + u_1 u_2. \tag{1.3}$$

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 view-point 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 (1990*a*).

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 TB-points

In this section we analyze a defining equation for TB-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 TB-points have been proposed in Roose (1987) and Pönisch (1991).

We consider an n-dimensional dynamical system with two parameters

$$\dot{u} = f(u, \lambda), \qquad u(t) \in \mathbb{R}^n, \qquad \lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^2 \tag{2.1}$$

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

A stationary point (u^0, λ^0) of (2.1) is called a TB-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 TB-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 λ.

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)

$$A(u, \lambda) = \begin{pmatrix} f_u(u, \lambda) & b_0 \\ c_0^T & 0 \end{pmatrix} \tag{2.2}$$

is non-singular for (u, λ) 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 Ω.

Then we calculate a TB-point by solving the following defining equation

$$S(u, \lambda) = \begin{pmatrix} f(u, \lambda) \\ g(u, \lambda) \\ h(u, \lambda) \end{pmatrix} = 0. \tag{2.3}$$

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

$$A\begin{pmatrix} v \\ g \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad A\begin{pmatrix} w \\ h \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix} \qquad \text{for } (u, \lambda) \in \Omega. \tag{2.4}$$

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

$$(\Psi^T, \tilde{g})A = (0, 1), \qquad (\zeta^T, \tilde{h})A = (\Psi^T, 0). \tag{2.5}$$

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

$$g = \tilde{g} = -\Psi^T f_u v, \qquad h = \tilde{h} = \Psi^T v \tag{2.6}$$

$$g_z = -\Psi^T f_{uz} v, \qquad h_z = -\Psi^T f_{uz} w - \zeta^T f_{uz} v. \tag{2.7}$$

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

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

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 3×3 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-singulairty of $S'(u^0, \lambda^0) = S'^0$ at a TB-point. In contrast to Roose (1987) and Pönisch (1991) we do not assume that (u^0, λ^0) is a quadratic turning point, i.e. $f^0_{uu} v^0 v^0 \notin R(f^0_u)$. 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, λ^0). From (2.4) and

(2.5) we have

$$f_u^0 v^0 = 0, \qquad f_u^0 w^0 = v^0, \qquad \Psi^{0T} f_u^0 = 0, \qquad \zeta^{0T} f_u^0 = \Psi^{0T}. \tag{2.8}$$

Moreover,

$$\Psi^{0T} v^0 = 0 \qquad \text{and} \qquad \alpha := \Psi^{0T} w^0 = \zeta^{0T} v^0 \neq 0.$$

There is no practical loss of generality in assuming biorthogonality

$$\Psi^{0T} v^0 = \zeta^{0T} w^0 = 0, \qquad \Psi^{0T} w^0 = \zeta^{0T} v^0 = 1. \tag{2.9}$$

This can always be achieved by replacing Ψ^0 by $\alpha^{-1}\Psi^0$ and ζ^0 by $\alpha^{-1}(\zeta^0 - \alpha^{-1}(\zeta^{0T} w^0)\Psi^0)$. Equivalently, we may view this process as an *a posteriori* normalization of the vectors b_0, c_0 from (2.2).

Let us put f_u^0 into block diagonal form

$$f_u^0 \Phi = \Phi \begin{pmatrix} H & 0 \\ 0 & J_0 \end{pmatrix}, \qquad \Phi = (R \;\; v^0 \;\; w^0) \in \mathbb{R}^{n,n} \tag{2.10}$$

where $H \in \mathbb{R}^{n-2,n-2}$ is non-singular and the columns of R form a basis of the invariant subspace $\{\Psi^0, \zeta^0\}^{\perp}$. 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

$$\begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} u^0 \\ \lambda^0 \end{pmatrix} + \Gamma \begin{pmatrix} x \\ \mu \end{pmatrix}, \qquad \Gamma = \begin{pmatrix} \Phi & D \\ 0 & B \end{pmatrix} \in \mathbb{R}^{n+2,n+2}. \tag{2.11}$$

The matrices $D \in \mathbb{R}^{n,2}$ and $B \in \mathbf{R}^{2,2}$ will be determined in such a way that the transformed system

$$\dot{x} = \Phi^{-1} f\left(\begin{pmatrix} u^0 \\ \lambda^0 \end{pmatrix} + \Gamma \begin{pmatrix} x \\ \mu \end{pmatrix}\right) =: \tilde{f}(x, \mu) \tag{2.12}$$

satisfies

$$\tilde{f}'(0, 0) = \begin{pmatrix} H & 0 & 0 \\ 0 & J_0 & J_0^T \end{pmatrix}. \tag{2.13}$$

From (2.10) we find that this holds iff

$$(w^0 \;\; 0) = \Phi \begin{pmatrix} 0 \\ J_0^T \end{pmatrix} = f_u^0 D + f_\lambda^0 B. \tag{2.14}$$

In order to achieve this, we assume the transversality condition

$$0 \neq \Psi^{0T} f_\lambda^0 =: (\beta_1, \beta_2). \tag{T1}$$

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'^0. 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

$$A^0\begin{pmatrix}D\\ d^T\end{pmatrix} = \begin{pmatrix}C\\ 0\end{pmatrix}. \tag{2.15}$$

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

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

and by (2.12), (2.13)

$$T'(0, 0) = S'^0\Gamma = \begin{pmatrix}\Phi & 0\\ 0 & I_2\end{pmatrix}\begin{pmatrix}\tilde{f}'(0, 0)\\ g_z^0\Gamma\\ h_z^0\Gamma\end{pmatrix}. \tag{2.16}$$

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

$$0 \neq \Delta := \det\begin{pmatrix}g_u^0 v^0 & g_z^0\delta_2\\ h_u^0 v^0 & h_z^0\delta_2\end{pmatrix}, \qquad \begin{pmatrix}D\\ B\end{pmatrix} = (\delta_1, \delta_2) \in \mathbb{R}^{n+2,2}. \tag{2.17}$$

By Lemma 2.1 we can express Δ in terms of the original function f. We introduce the coefficients

$$\begin{aligned} Q_{1,11} &= \zeta^{0T} f_{uu}^0 v^0 v^0, & Q_{2,11} &= \Psi^{0T} f_{uu}^0 v^0 v^0\\ Q_{2,12} &= \Psi^{0T} f_{uu}^0 v^0 w^0, & Q_{1,14} &= \zeta^{0T} f_{uz}^0 v^0 \delta_2\\ Q_{2,14} &= \Psi^{0T} f_{uz}^0 v^0 \delta_2, & Q_{2,24} &= \Psi^{0T} f_{uz}^0 w^0 \delta_2. \end{aligned} \tag{2.18}$$

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

$$0 \neq \Delta = Q_{2,11}(Q_{1,14} + Q_{2,24}) - Q_{2,14}(Q_{1,11} + Q_{2,12}) \tag{T2}$$

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

THEOREM 2.2 Let (u^0, λ^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, λ^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 rank $(\bar{S}'^0) = n+1$. Hence there exists a branch $z(s) = (u(s), \lambda(s))$ of solutions to $\bar{S} = 0$ such that $z(0) = (u^0, \lambda^0)$. Moreover, by differentiating $\bar{S}(z(s)) = 0$ and using (2.16) we obtain

$$z'(0) = \gamma\left[g_z^0 \delta_2 \begin{pmatrix} v^0 \\ 0 \end{pmatrix} - (g_u^0 v^0)\delta_2\right] \qquad \text{for some } \gamma \neq 0. \tag{2.19}$$

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

$$0 \neq \frac{\mathrm{d}}{\mathrm{d}s}[\Psi(z(s))^T v(z(s))]|_{s=0}. \tag{2.20}$$

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

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

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

$$\dot{u}_1 = u_2, \qquad \dot{u}_2 = \lambda_1 + \lambda_2 u_1 + u_1 u_2.$$

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, λ^0) as a regular solution of (2.3) and that f_u^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

$$\dot{x} = \bar{f}(x, \mu), \qquad \dot{\mu} = 0 \tag{3.1}$$

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)

$$W_{\mathrm{loc}}^c = \{(F(\xi, \mu), \xi, \mu) : (\xi, \mu) \in U(0) \subset \mathbb{R}^4\}$$

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

$$F(0, 0) = 0, \qquad F'(0, 0) = 0. \tag{3.2}$$

The reduced system within W^c_{loc} is

$$\dot{\xi} = \begin{pmatrix} \tilde{f}_{n-1} \\ \tilde{f}_n \end{pmatrix}(F(\xi, \mu), \xi, \mu) =: r(\xi, \mu). \tag{3.3}$$

By (2.9), (2.12), (2.13) and (3.2) the right hand side has a Taylor expansion

$$r(\xi, \mu) = \begin{pmatrix} \tilde{f}_{n-1} \\ \tilde{f}_n \end{pmatrix}(0, \xi, \mu) + O_3 = \begin{pmatrix} \xi_2 \\ \mu_1 \end{pmatrix} + \tfrac{1}{2} Q \begin{pmatrix} \xi \\ \mu \end{pmatrix}^2 + O_3$$

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

$$Q\begin{pmatrix} \xi \\ \mu \end{pmatrix}^2 = \begin{pmatrix} \zeta^{0T} \\ \Psi^{0T} \end{pmatrix} f^{n0} \left[\begin{pmatrix} v^0 & w^0 & D \\ 0 & 0 & B \end{pmatrix} \begin{pmatrix} \xi \\ \mu \end{pmatrix} \right]^2. \tag{3.4}$$

If we consider μ_1, μ_2 as the third and fourth component of $\begin{pmatrix} \xi \\ \mu \end{pmatrix}$ then the indices in (2.18) are consistent with (3.4). In addition to the coefficients from (2.18) we also need

$$Q_{2,44} = \Psi^{0T} f^{n0} \delta_2^2. \tag{3.5}$$

All the other quadratic and higher-order coefficients will not be used later on, so we work with the system

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 + \tfrac{1}{2} Q_{1,11} \xi_1^2 + Q_{1,14} \xi_1 \mu_2 + \cdots \\ \dot{\xi}_2 &= \mu_1 + \tfrac{1}{2} Q_{2,11} \xi_1^2 + Q_{2,12} \xi_1 \xi_2 + Q_{2,14} \xi_1 \mu_2 + Q_{2,24} \xi_2 \mu_2 + \tfrac{1}{2} Q_{2,44} \mu_2^2 + \cdots . \end{aligned} \tag{3.6}$$

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, ϵ, τ) as follows

$$\mu_1 = (\alpha_0 + \alpha_1 \tau^2)\epsilon^4, \qquad \mu_2 = \tau \epsilon^2 \tag{3.7a}$$

$$\xi_1(t) = \alpha_2 \epsilon^2 (x(\alpha_3 \epsilon t) + \alpha_4 \tau), \qquad \xi_2(t) = \alpha_5 \epsilon^3 y(\alpha_3 \epsilon t). \tag{3.7b}$$

Here the constants α_i are yet to be determined. The terms which involve α_1, α_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

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = g(x, y, \epsilon, \tau) = \begin{pmatrix} y \\ x^2 - 4 \end{pmatrix} + \epsilon \begin{pmatrix} \tfrac{1}{2} a_1 x^2 + a_2 \tau x + a_3 \tau^2 \\ b_1 xy + b_2 \tau y \end{pmatrix} + O(\epsilon^2) \tag{3.8}$$

by setting

$$\alpha_0 = -(1/2Q_{2,11}), \qquad \alpha_1 = \alpha_0(Q_{2,11}Q_{2,44} + Q_{2,14}^2), \qquad \alpha_2 = -\alpha_0,$$
$$\alpha_3 = \tfrac{1}{2}, \qquad \alpha_4 = -2Q_{2,14}, \qquad \alpha_5 = -\tfrac{1}{2}\alpha_0.$$

The coefficients in the quadratic terms of (3.8)—except for a_3, which is not

needed later on—are found to be

$$a_1 = \bar{q}Q_{1,11}, \qquad a_2 = 2\bar{q}(Q_{1,14}Q_{2,11} - Q_{1,11}Q_{2,14}), \qquad \bar{q} = 1/Q_{2,11} \tag{3.9a}$$

$$b_1 = Q_{2,12}\bar{Q}, \qquad b_2 = 2\bar{q}(Q_{2,24}Q_{2,11} - Q_{2,14}Q_{2,12}). \tag{3.9b}$$

We notice that it is this step which requires the quadratic turning point condition

$$Q_{2,11} = \Psi^{0T} f^0_{uu} v^0 v^0 \neq 0 \tag{3.10}$$

which we were able to avoid in the previous section. Moreover, the transversality condition (T2) now translates into

$$a_2 + b_2 \neq 0. \tag{3.11}$$

For $\epsilon = 0$ the system (3.8) is Hamiltonian and has the well-known homoclinic orbit (cf. Guckenheimer & Holmes, 1983).

$$(\bar{x}(t), \bar{y}(t)) = 2(1 - 3\,\mathrm{sech}^2(t), 6\,\mathrm{sech}^2(t)\tanh(t)), \qquad t \in \mathbb{R}. \tag{3.12}$$

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 view-point which will be used in Section 5 to establish convergence of our numerical method.

We introduce the Banach spaces

$$X_0 = \left\{ z \in C(\mathbb{R}, \mathbb{R}^2) : \lim_{t \to -\infty} z(t) \text{ and } \lim_{t \to \infty} z(t) \text{ exist} \right\}$$

with norm

$$\|z\|_0 = \sup\{\|z(t)\| : t \in \mathbb{R}\}, \qquad \|\cdot\| \text{ some norm in } \mathbb{R}^m$$

and

$$X_1 = \{z \in C^1(\mathbb{R}, \mathbb{R}^2) : z, \dot{z} \in X_0\}, \qquad \|z\|_1 = \|z\|_0 + \|\dot{z}\|_0.$$

With $w = (z, \epsilon) = (x, y, \epsilon) \in X_1 \times \mathbb{R}$ let us write (3.8) as

$$F(w, \tau) = 0 \tag{4.1}$$

where $F : X_1 \times \mathbb{R} \times \mathbb{R} \to X_0 \times \mathbb{R}$ is defined by

$$F(w, \tau) = \begin{pmatrix} \dot{z} - g(z, \epsilon, \tau) \\ y(0) \end{pmatrix}. \tag{4.2}$$

The condition $y(0) = 0$ is used to fix the phase of the homoclinic orbits.

A careful inspection of the transformation (3.7*b*) shows that F is only defined for arguments $(w, \tau) = (z, \varepsilon, \tau)$ satisfying $|\epsilon| \leq \varepsilon_0$, $|\tau| \leq C\epsilon^{-2}$, $\|z\|_0 \leq C\epsilon^{-2}$ for some $\epsilon_0 > 0$. This will be sufficient for our purposes and we will not always meniton it explicitly.

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

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

Hence we have a trivial branch $(\bar{w}, \tau)$ of homoclinic orbits and we look for values of τ 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

$$\bar{w}, \tau_0 = \frac{10}{7}\frac{a_1 + b_1}{a_2 + b_2}. \tag{4.3}$$

The emanating C^1-branch can be parametrized by ϵ

$$\big(w(\epsilon), \tau(\epsilon)\big) = \big(z(\epsilon), \epsilon, \tau(\epsilon)\big) \in X_1 \times \mathbb{R}^2. \tag{4.4}$$

It has tangent

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

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

$$\begin{aligned} \dot{x} - y &= \tfrac{1}{2}a_1\bar{x}^2 + a_2\tau_0\bar{x} + a_3\tau_0^2 \\ \dot{y} - 2\bar{x}x &= b_1\bar{x}\bar{y} + b_2\tau_0\bar{y} \end{aligned} \qquad \text{and} \quad y(0) = 0. \tag{4.5}$$

REMARK By our construction we only know that both limits

$$\lim_{t\to\infty} z(t, \epsilon) = z_+ \qquad \text{and} \qquad \lim_{t\to-\infty} z(t, \epsilon) = z_- \qquad \text{exist.}$$

These points must be stationary points of (3.8) at $\tau = \tau(\epsilon)$ and lie close to the stationary point $(2, 0)$ of the unperturbed system ($\epsilon = 0$). 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 $z(\epsilon)$ are in fact homoclinic.

Proof. Consider the linearization of F with respect to $w \in X_1 \times \mathbb{R}$ about the trivial solution $(\bar{w}, \tau)$.

$$F_w(\bar{w}, \tau)w = \begin{pmatrix} Lz - g_\epsilon(\bar{z}, 0, \tau)\epsilon \\ y(0) \end{pmatrix}, \qquad w = (z, \epsilon) = (x, y, \epsilon) \in X_1 \times \mathbb{R} \tag{4.6}$$

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

$$Lz = \begin{pmatrix} \dot{x} - y \\ \dot{y} - 2\bar{x}x \end{pmatrix}, \qquad z = (x, y). \tag{4.7}$$

We now need a few facts from the Fredholm theory of linear differential operators $Lz = \dot{z} - A(t)z$ which have the property that $\lim_{t\to\infty} A(t) = \lim_{t\to-\infty} A(t)$

exists and has no eigenvalues on the imaginary axis (see Beyn, 1990*a*, Lemma 2.2, and also Palmer, 1984, Lemma 4.2; Hale, 1983).

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

For the special case (4.7) we have

$$N(L) = \text{span}\,\{\hat{z}\}, \qquad N(L^*) = \text{span}\,\{(-\hat{y}, \hat{x})\}. \tag{4.8}$$

Notice that $\hat{z} \in N(L)$ 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 $t \to \infty$ in contradiction to Liouvilles theorem. The representation of $N(L^*)$ then follows easily from (ii).

Using (i) and the bordering lemma (Beyn, 1990*a*, Lemma 2.3) we find that $F_w(\bar{w}, \tau) : X_1 \times \mathbb{R} \to X_0 \times \mathbb{R}$ also has Fredholm index 0. Since $\hat{y}(0) \neq 0$ the only way that $F_w(\bar{w}, \tau)$ can have a non-trivial null space is in case

$$g_\epsilon(\bar{z}, 0, \tau) \in R(L).$$

By (iii), (4.8) and (3.12) this is equivalent to a vanishing Melnikov integral

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} (-\hat{y}, \hat{x}) g_\epsilon(\hat{z}, 0, \tau)\, dt \\ &= \int_{-\infty}^{\infty} -\ddot{\bar{x}}(\tfrac{1}{2}a_1\bar{x}^2 + a_2\tau\bar{x} + a_3\tau^2) + \dot{\bar{x}}b_1\bar{x}\dot{\bar{x}} + b_2\tau\dot{\bar{x}}\, dt \\ &= \int_{-\infty}^{\infty} (a_1 + b_1)\bar{x}\dot{\bar{x}}^2 + \tau(a_2 + b_2)\dot{\bar{x}}^2\, dt = 192\big(-\tfrac{2}{7}(a_1 + b_1) + (\tau/5)(a_2 + b_2)\big). \end{aligned}$$

This is satisfied at $\tau = \tau_0$ and we obtain

$$N\big(F_w(\bar{w}, \tau_0)\big) = \text{span}\,\{(z_0, 1)\}, \qquad R\big(F_w(\bar{w}, \tau_0)\big) = R(L) \times \mathbb{R} \tag{4.9}$$

where z_0 is the unique solution of (4.5).

The final condition for bifurcation from $(\bar{w}, \tau_0)$ is (cf. Crandall & Rabinowitz, 1971)

$$F_{w\tau}(\bar{w}, \tau_0)\begin{pmatrix} z_0 \\ 1 \end{pmatrix} \notin R\big(F_w(\bar{w}, \tau_0)\big).$$

Using (4.9) and once more (iii) and (4.8) this turns out to be equivalent to

$$0 \neq \int_{-\infty}^{\infty} (-\hat{y}, \hat{x}) g_{\epsilon\tau}(\bar{z}, 0, \tau_0)\, dt = \int_{-\infty}^{\infty} -\ddot{\bar{x}}(a_2\bar{x} + 2a_3\tau_0) + \dot{\bar{x}}b_2\hat{y}\, dt = \tfrac{192}{5}(a_2 + b_2)$$

which is true by assumption.

Finally, (4.9) shows that we may parametrize the non-trivial branch such that it has tangent $\big(z_0, 1, \tau'(0)\big)$ 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

$$\big(z(t,\epsilon),\epsilon,\tau(\epsilon)\big) = \big(x(t,\epsilon), y(t,\epsilon),\epsilon,\tau(\epsilon)\big), \qquad t\in\mathbb{R},\ |\epsilon|<\epsilon_0$$

satisfies

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

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

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

$$D_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad D_\epsilon = \begin{pmatrix} \alpha_2\epsilon^2 & 0 \\ 0 & \alpha_5\epsilon^3 \end{pmatrix} \qquad \text{for } \epsilon\neq 0.$$

From (3.7) we find the relation

$$\Phi(s,\eta_0,\epsilon,\tau) = D_\epsilon^{-1}\Xi\big[(s/\alpha_3\epsilon), D_\epsilon\big(\eta_0+\alpha_4\tau(1,0)\big), (\alpha_0+\alpha_1\tau^2)\epsilon^4, \tau\epsilon^2\big]$$

and hence

$$\Phi(s, D_0\eta_0, -\epsilon, \tau) = D_0\Phi(-s,\eta_0,\epsilon,\tau), \qquad \epsilon\neq 0.$$

Therefore, if $\big(x(t,\epsilon), y(t,\epsilon),\epsilon,\tau(\epsilon)\big)$ is a homoclinic solution of (4.1) then so is $\big(x(-t,\epsilon), -y(-t,\epsilon), -\epsilon, \tau(\epsilon)\big)$. 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. □

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

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

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

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

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

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 $\epsilon>0$ 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. Once

a homoclinic orbit of (2.1) has been found, we can use standard algorithms (see Doedel & Friedman, 1989; Beyn, 1990*a*; 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 ϵ (Theorem 4.1) and rewrite (4.1) as

$$H(v, \epsilon) = 0 \tag{5.1}$$

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

$$H(v, \epsilon) = \big(\dot{z} - g(z, \epsilon, \tau), y(0)\big) \in X_0 \times \mathbb{R}. \tag{5.2}$$

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

$$v_{n+1} = T(v_n, \epsilon), \qquad n = 0, 1, 2, \ldots \tag{5.3}$$

where either

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

is the chord operator or

$$T(v, \epsilon) = v - H_v(v, \epsilon)^{-1} H(v, \epsilon)$$

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_v^0 is Fredholm of index 0 with $N(H_v^0) = \text{span}\,\{\varphi\}$ for

some $\varphi \neq 0$. Take splittings $V = N(H_v^0) \oplus V_1$, $W = R(H_v^0) \oplus W_1$ and let $P : V \to N(H_v^0)$, $Q : W \to W_1$ be the corresponding projectors.

(iii) $QH_{v\epsilon}^0 \varphi \neq 0$.

(iv) $QH_\epsilon^0 = 0$.

(v) $QH_{vv}^0 \varphi v = 0$ for all $v \in V$.

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

$$H(v, \epsilon) = 0, \qquad 0 < \epsilon \leqslant \epsilon_0 \tag{5.4}$$

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

$$K_\epsilon = \{v \in V : \|P(v - v_0)\| \leqslant \epsilon^\alpha, \ \|(I - P)(v - v_0)\| \leqslant c_0 \epsilon\}. \tag{5.5}$$

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

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 $H_\epsilon^0 \in R(H_v^0)$ and

$$N(H'^0) = \text{span}\,\{(\varphi, 0), (\zeta, 1)\}$$

where $H_v^0 \zeta = -H_\epsilon^0$. The quadratic form induced by H'^0 on $N(H'^0)$ is given by

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

where $\alpha = QH_{vv}^0 \varphi^2 = 0$ by (v), $\beta = QH_{v\epsilon}^0 \varphi \neq 0$ by (iii) and $\gamma = QH'^0(\zeta, 1)^2$. Therefore, the tangents to the bifurcating branches are

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

In general, γ and ζ 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 $QH_{vv}^0 \varphi^2 = 0$.

2. Gluing the balls K_ϵ together gives us the domain of convergence

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

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 τ and ϵ interchanged.

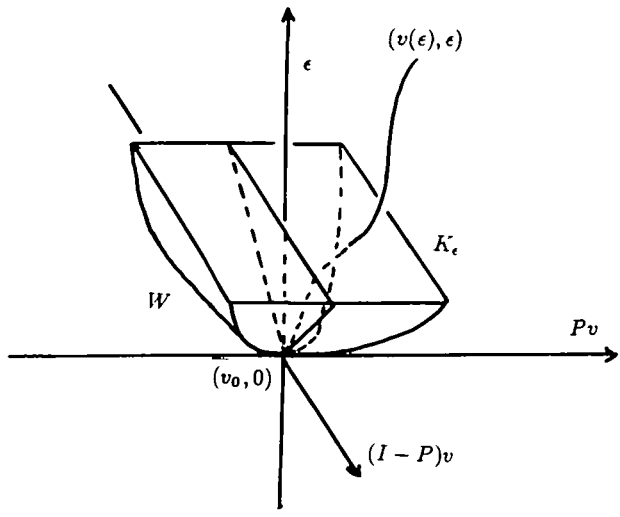

FIG. 2. Convergence wedge for Newton's method near a bifurcation point (Theorem 5.1).

From the proof of Theorem 4.1 we obtain

$$H(v_0, 0) = \big(\dot{\bar{z}} - g(\bar{z}, 0, \tau_0), \bar{y}(0)\big) = 0$$

$$H_v^0(z, \tau) = \big(Lz, y(0)\big) \qquad \text{with } L \text{ from (4.7)},$$

$$N(H_v^0) = \text{span}\,\{\varphi\} \qquad \text{with } \varphi = (0, 0, 1) \in X_1 \times \mathbb{R},$$

$$R(H_v^0) = \{(z, s) \in X_0 \times \mathbb{R} : \Psi(z, s) = 0\}, \qquad \Psi(z, s) = \int_{-\infty}^{\infty} (-\hat{y}, \hat{x}) z \, dt,$$

$$H_{v\epsilon}^0 \varphi \notin R(H_v^0) \qquad \text{since}$$

$$\Psi(H_{v\epsilon}^0 \varphi) = \int_{-\infty}^{\infty} \ddot{\bar{x}}(a_2 + 2a_3\tau_0) - \hat{x} b_2 \hat{x} \, dt = -(a_2 + b_2)\int_{-\infty}^{\infty} \hat{x}^2 \, dt \neq 0,$$

$$\Psi\big(H(v_0, \epsilon)\big) = O(\epsilon^2) - \epsilon \int_{-\infty}^{\infty} (-\hat{y}, \hat{x}) g_\epsilon(\bar{z}, 0, \tau_0) \, dt = O(\epsilon^2),$$

$$H_{vv}^0 \varphi v = \big(-g_{\tau z}(\bar{z}, 0, \tau_0) z - g_{\tau\tau}(\bar{z}, 0, \tau_0)\tau, 0\big) = 0 \qquad \text{for all } v = (z, \tau) \in X_1 \in \mathbb{R}.$$

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

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = g(x, y, \epsilon, \tau), \qquad y(0) = 0 \tag{5.6}$$

converges to the non-trivial homoclinic orbits $\big(x(\epsilon), y(\epsilon), \tau(\epsilon)\big)$ of Theorem 4.1 when started at $(\bar{x}, \bar{y}, \tau_0)$.

Of course, Newton's method in the Banach space $X_1 \times \mathbb{R}$ 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 (u^0, λ^0) (see Appendix A). This approximation depends on a parameter ϵ as follows

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

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

$$\xi_1(t, \epsilon) = (\epsilon^2/Q_{2,11})[1 - 3\,\mathrm{sech}^2((\epsilon/2)t) - Q_{2,14}\tau_0] \tag{6.1c}$$

$$\xi_2(t, \epsilon) = (3\epsilon^3/Q_{2,11})\,\mathrm{sech}^2((\epsilon/2)t)\tanh((\epsilon/2)t). \tag{6.1d}$$

We then fix one parameter, say $\lambda_1 = \tilde{\lambda}_1$, and solve a well-posed boundary value problem for $(u, \lambda_2) \in X_1 \times \mathbb{R}$ (cf. Doedel & Friedman, 1989; Beyn, 1990*a* 1990*b*; and Section 5).

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

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

where $\sigma : X_1 \to \mathbb{R}$ 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))

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

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

$$\sigma(u) = \int_{-\infty}^{\infty} \dot{\tilde{u}}(t)^T(u(t) - \tilde{u}(t))\,\mathrm{d}t = 0. \tag{6.4}$$

In the following examples we have concentrated on the local behaviour of the starting algorithm, i.e. the initial homoclinic orbit (u, λ_2) was computed and it was then continued by slightly varying λ_1.

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

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

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

$$\sigma_J(u) = 0 \tag{6.5b}$$

$$B_-(u(T_-), \lambda_2) = 0, \qquad B_+(u(T_+), \lambda_2) = 0. \tag{6.5c}$$

Here, σ_J denotes a finite phase condition, e.g. the point condition (6.3) or the truncated integral condition

$$\sigma_J(u) = \int_{T_-}^{T_+} \dot{\bar{u}}(t)^T \big(u(t) - \bar{u}(t)\big)\, dt. \tag{6.6}$$

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

$$B_-(v, \lambda_2) = L_s(\lambda_2)\big(v - v(\lambda_2)\big) \in \mathbb{R}^{n_s},$$
$$B_+(v, \lambda_2) = L_u(\lambda_2)\big(v - v(\lambda_2)\big) \in \mathbb{R}^{n_u}$$

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

$$\dot{u} = f(u, \bar{\lambda}_1, \lambda_2) \tag{6.7}$$

close to $\bar{u}(-\infty) = \bar{u}(+\infty)$. Moreover, the rows of $L_s(\lambda_2) \in \mathbb{R}^{n_s, n}$ (resp. $L_u(\lambda_2) \in \mathbb{R}^{n_u, n}$) form a basis of the stable (resp. unstable) subspace of $f_u^T\big(v(\lambda_2), \bar{\lambda}_1, \lambda_2\big)$. 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_- were determined from

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

where δ_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, 1990*a*). The solver D02RAF 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 $\epsilon = 0{\cdot}1$, $\delta_0 = 10^{-4}$ and $\bar{\lambda}_1 = 0{\cdot}25 \times 10^{-4}$ the b.v.p. (6.5) could be solved easily. Then λ_1 was increased and a branch of homoclinic orbits was calculated up to $\lambda_1 = 0{\cdot}35 \times 10^{-2}$. Figure 3 shows the evolving homoclinics in phase space. During the continuation the adaptation

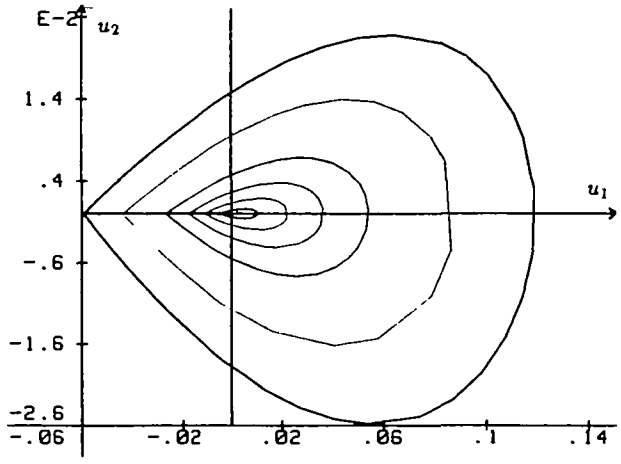

FIG. 3. Branch of homoclinic orbits in phase space for Example 1 ($\lambda_1 = 0{\cdot}25 \times 10^{-4}, \ldots, 0{\cdot}35 \times 10^{-2}$).

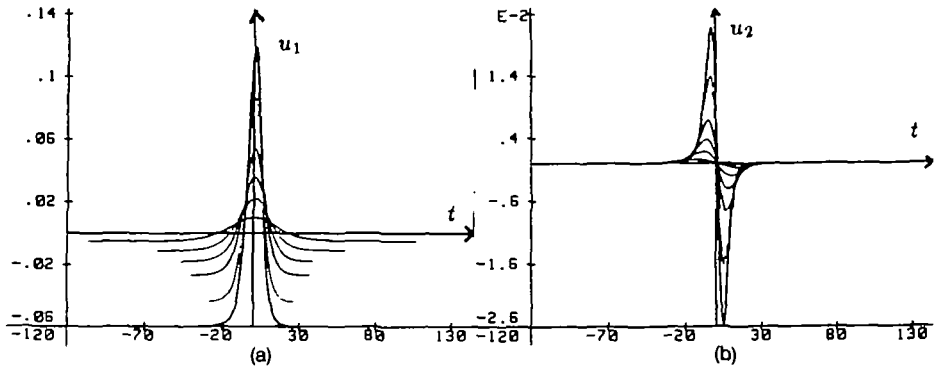

FIG. 4. (a), (b): Time diagrams of u_1, u_2 for the homoclinic orbits from Fig. 3.

strategy for $[T_-, T_+]$, as developed in Beyn (1990*b*), 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 λ_2, T_-, T_+ with increasing λ_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 β-cells. It is a modification by Chay and Keizer (1983) of the Hodgkin–Huxley model, further simplified by Rinzel (1985). The parameters are $\lambda_1 = \mathrm{Ca}\,[\mu\mathrm{M}]$ (intracellular calcium concentration) and $\lambda_2 = T[°\mathrm{C}]$ (the temperature).

The dimensionless equations are

$$C_M \dot{V} = 2\bar{g}_{\mathrm{Ca.HH}} m_\infty^3(V) h_\infty(V)(V_{\mathrm{Ca}} - V) + \left(\bar{g}_{\mathrm{K.HH}} n^4 + \bar{g}_{\mathrm{K.Ca}} \frac{\mathrm{Ca}}{1+\mathrm{Ca}}\right)(V_K - V) + \bar{g}_L(V_L - V)$$

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

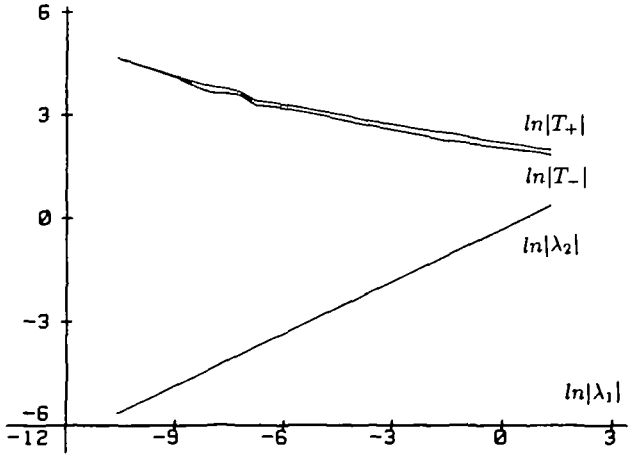

FIG. 5. Dependence of $\ln|T_-|$, $\ln(T_+)$, $\ln|\lambda_2|$ on the continuation parameter λ_1 for Example 1.

where

$$C_M = 1, \quad \bar{g}_{Ca,HH} = 6{\cdot}5, \quad \bar{g}_{K,HH} = 12, \quad \bar{g}_{K,Ca} = 0{\cdot}09, \quad \bar{g}_L = 0{\cdot}04,$$
$$V_{Ca} = 100, \quad V_K = -75, \quad V_L = -40, \quad \Phi = 3^{(T-6{\cdot}3)/10},$$
$$m_\infty(V) = \frac{\alpha_m(V)}{\alpha_m(V) + \beta_m(V)}, \qquad h_\infty(V) = \frac{\alpha_h(V)}{\alpha_h(V) + \beta_h(V)}.$$
$$\alpha_m(V) = \bar{e}\left(\frac{-25 - V}{10}\right), \qquad \alpha_n(V) = \bar{e}\left(\frac{-20 - V}{10}\right), \qquad \bar{e}(x) = \frac{x}{e^x - 1},$$
$$\alpha_h(V) = 0{\cdot}07 \exp\left(\frac{-50 - V}{20}\right), \qquad \beta_h(V) = \left(1 + \exp\left(\frac{-20 - V}{10}\right)\right)^{-1}$$
$$\beta_m(V) = 4 \exp\left(\frac{-50 - V}{18}\right), \qquad \beta_n(V) = 0{\cdot}125 \exp\left(\frac{-30 - V}{80}\right).$$

We found a TB-point at $(\lambda_1, \lambda_2) = (1{\cdot}88, 14{\cdot}94)$, $(V, n) = (-34{\cdot}21, 0{\cdot}26)$. We did some experiments on the quality of the predicted homoclinic (6.1) with varying values of ϵ (see also Rodriguez Luis *et al*, 1990). Figure 6 compares the predicted values $\tilde{\lambda}_2 = \tilde{\lambda}_2(\epsilon)$ with the computed ones $\lambda_2(\epsilon)$. Note that $\lambda_2(\epsilon)$ is obtained by solving (6.5) with $\lambda_1 = \tilde{\lambda}_1(\epsilon)$ fixed. The predictor was found to be reliable up to $\epsilon = 0{\cdot}3$. 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))

$$\dot{x} = 2k_1 z^2 - 2k_{-1} x^2 - k_3 xy, \qquad z = 1 - x - y - s$$
$$\dot{y} = k_2 z - k_{-2} y - k_3 xy$$
$$\dot{s} = k_4\left(z - \frac{k_{-4}}{k_4} s\right).$$

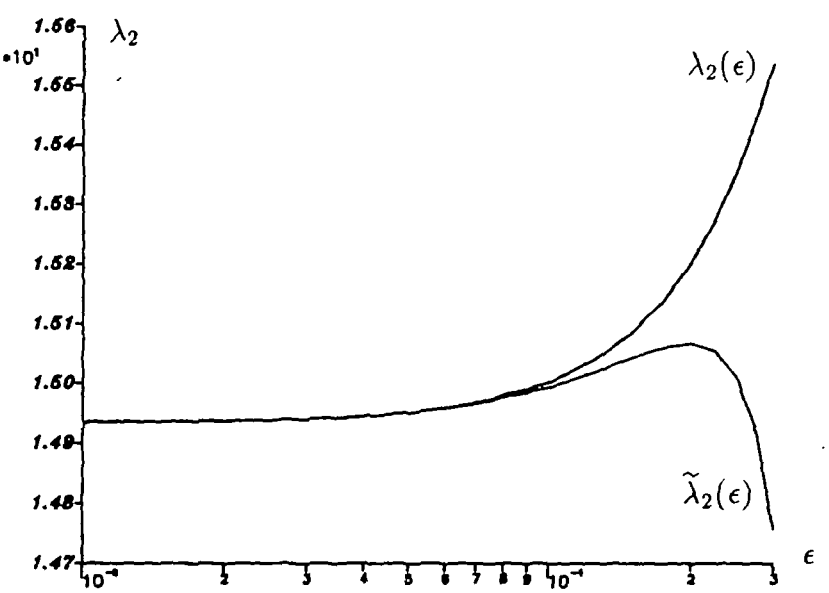

FIG. 6. Comparison of parameter values $\tilde{\lambda}_2(\epsilon)$, predicted by the starting procedure, with computed values $\lambda_2(\epsilon)$ (Example 2).

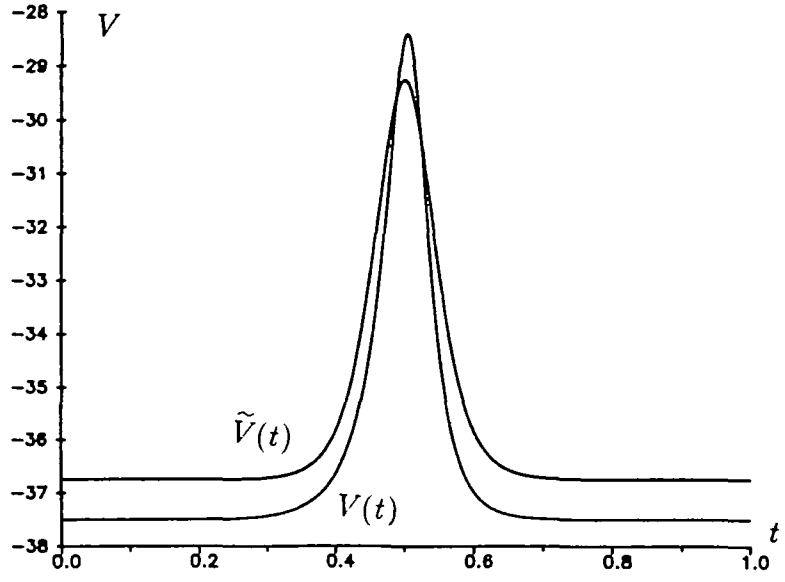

FIG. 7. Comparison of predicted ($\tilde{V}(t)$) and computed ($V(t)$) voltage function for $\epsilon = 0{\cdot}3$ (Example 2).

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)

$$k_1 = 2{\cdot}5, \qquad k_{-1} = 1, \qquad k_3 = 10, \qquad k_{-2} = 0{\cdot}1, \qquad k_4 = 0{\cdot}0675$$

and a TB-point was calculated at

$$(x, y, s) = (0{\cdot}1159,\ 0{\cdot}31547,\ 0{\cdot}28844)$$
$$\lambda_1 = k_2 = 1{\cdot}4176, \qquad \lambda_2 = k_{-4}/k_4 = 0{\cdot}9714.$$

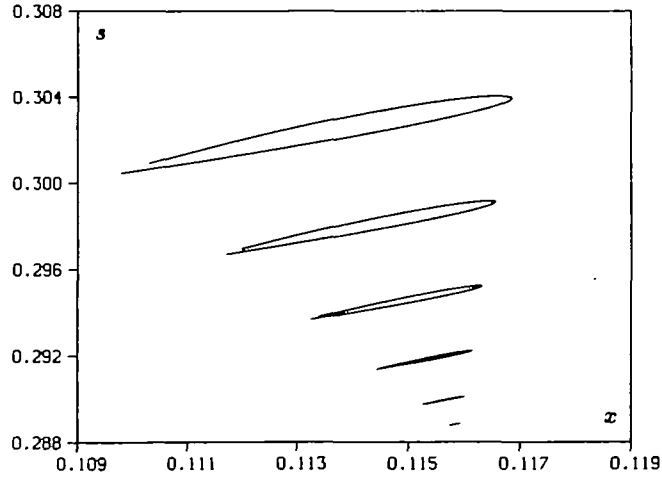

FIG. 8. Homoclinic orbits in (x, s)-phase space for Example 3 ($\epsilon = 0{\cdot}005, \ldots, 0{\cdot}03$).

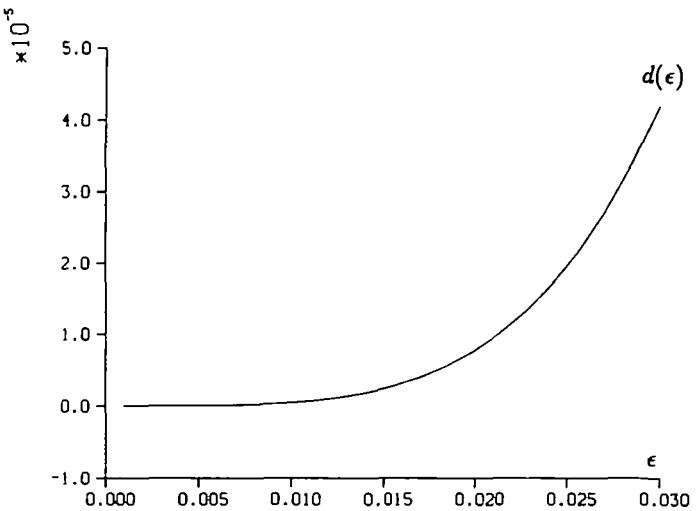

FIG. 9. Distance of homoclinic orbits from the tangent plane to the centre manifold for Example 3.

The emanating homoclinic orbits for a few values of ϵ are shown in Fig. 8.

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

$$\mathrm{d}(\epsilon) = \sup_{t \in [T_-, T_+]} \left| \gamma^T \left(\bar{u}(t, \epsilon) - u^0 - D \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \right) \right|.$$

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, λ^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, λ_0). However, we try to suppress this index whenever it does not affect the readability. Compute $v, w, \Psi, \zeta \in \mathbb{R}^n$ from the linear systems:

$$A^0 \begin{pmatrix} v \\ g \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad A^0 \begin{pmatrix} w \\ h \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix}, \qquad A^{0T} \begin{pmatrix} \Psi \\ g \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad A^{0T} \begin{pmatrix} \zeta \\ h \end{pmatrix} = \begin{pmatrix} \Psi \\ 0 \end{pmatrix}.$$

Normalize

$$\alpha \leftarrow \Psi^T w; \qquad \zeta \leftarrow \frac{1}{\alpha} \left(\zeta - \frac{1}{\alpha} (\zeta^T w) \Psi \right), \qquad \Psi \leftarrow \frac{1}{\alpha} \Psi.$$

A.2 *Linear normal form*

$$\beta_i \leftarrow \Psi^T f^0_{\lambda_i}, \qquad i = 1, 2,$$

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

$$C \leftarrow (w \;\; 0) - f^0_\lambda B \in \mathbb{R}^{n,2}.$$

Solve the two linear systems

$$A^0 \begin{pmatrix} D \\ d^T \end{pmatrix} = \begin{pmatrix} C \\ 0 \end{pmatrix}.$$

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

$$v_1 \leftarrow (1/s)\big(f_u(u^0 + sv, \lambda^0) - f^0_u\big)v;$$

$$v_4 \leftarrow (1/s)\big(f_z(u^0 + sv, \lambda^0) - f^0_z\big)\delta_2;$$

$$Q_{1,11} \leftarrow \zeta^T v_1; \qquad Q_{2,11} \leftarrow \Psi^T v_1; \qquad Q_{1,14} \leftarrow \zeta^T v_4; \qquad Q_{2,14} \leftarrow \Psi^T v_4;$$

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

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

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

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

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

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

$$\tau_0 \leftarrow \frac{5}{7\Delta}(Q_{1,11} + Q_{2,12}); \qquad \sigma_0 \leftarrow \frac{1}{2Q_{2,11}}\big((Q^2_{1,14} - Q_{2,11}Q_{2,44})\tau_0^2 - 1\big);$$

choose an $\epsilon > 0$ and let

$$\mu_1 \leftarrow \sigma_0 \epsilon^4; \qquad \mu_2 \leftarrow \tau_0 \epsilon^2$$

$$\xi_1(t) \leftarrow \frac{\epsilon^2}{Q_{2,11}}\left(1 - 3\,\mathrm{sech}^2\left(\frac{\epsilon}{2}t\right) - Q_{2,14}\tau_0\right)$$

$$\xi_2(t) \leftarrow \frac{3\epsilon^3}{Q_{2,11}}\,\mathrm{sech}^2\left(\frac{\epsilon}{2}t\right)\tanh\left(\frac{\epsilon}{2}t\right).$$

A.5 *Approximate homoclinic for the original system* (2.1)

$$\bar{u}(t) = u^0 + \xi_1(t)v + \xi_2(t)w + D\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$$

$$\bar{\lambda} = \lambda^0 + B\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}.$$

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

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

and show that $T(\cdot, \epsilon)$ is a contraction on K_ϵ 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.

$$|v| \geqslant 0, \qquad |v| = 0 \Leftrightarrow v = 0, \qquad |v_1 + v_2| \leqslant |v_1| + |v_2|$$

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

$$|v| = \begin{pmatrix}\|Pv\|\\ \|(I-P)v\|\end{pmatrix}, \qquad v \in V;$$

$$|w| = \begin{pmatrix}\|Qw\|\\ \|(I-Q)w\|\end{pmatrix}, \qquad w \in W. \tag{B1}$$

We may then write

$$K_\epsilon = \{v \in V : |v - v_0| \leqslant r_\epsilon\}, \qquad r_\epsilon = \begin{pmatrix}\epsilon^\alpha\\ c_0\epsilon\end{pmatrix}.$$

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

$$B = \{v \in V : |v - v_0| \leqslant r\}, \qquad \text{where } r \in \mathbb{R}^m, \qquad r > 0,$$

an operator $T : B \to V$ and a non-negative matrix $L \in \mathbb{R}^{m,m}$ with the following properties:

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

$$|T(v_0) - v_0| < (I - L)r. \tag{B3}$$

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 $\bar{v}$. Moreover,

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

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

$$\|v\|_r = \max\left\{\frac{|v|_i}{r_i} : i = 1,\ldots,m\right\}.$$

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)

$$|H_v(v_0, \epsilon)^{-1} v| \leq C^* \begin{bmatrix} \epsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix} |v| \qquad \forall v \in V. \tag{B5}$$

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 ϵ.

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

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

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

$$c_0 = 2C \qquad \text{and} \qquad c_0 \epsilon \leq \epsilon^\alpha \qquad \text{for } 0 < \epsilon \leq \epsilon_1. \tag{B6}$$

Then the following holds

$$|T(v_0, \epsilon) - v_0| \leq \tfrac{1}{2} r_\epsilon \qquad \text{for } 0 < \epsilon \leq \epsilon_1. \tag{B7}$$

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

$$\begin{aligned} |T(v, \epsilon) - T(y, \epsilon)| &= |H_v(v_0, \epsilon)^{-1}[H_v(v_0, \epsilon)(v - y) - (H(v, \epsilon) - H(y, \epsilon))]| \\ &\leq C^* \begin{bmatrix} \epsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix} \left| \int_0^1 [H_v(v_0, \epsilon) - H_v(y + t(v - y), \epsilon)]\, dt (v - y) \right| \\ &= C^* \begin{bmatrix} \epsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix} |z|, \end{aligned}$$

where

$$z = \int_0^1 \int_0^1 H_{vv}[(1-s)v_0 + s(y + t(v-y)), \epsilon](y + t(v-y) - v_0)(v-y)\,ds\,dt.$$

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

$$\begin{aligned}
\|Qz\| &\leq \Big\| Q \int_0^1 \int_0^1 [H_{vv}((1-s)v_0 + s(y + t(v-y)), \epsilon) - H^0_{vv}] \\
&\quad \times (y + t(v-y) - v_0)(v-y)\,ds\,dt \Big\| \\
&\quad + \Big\| Q \int_0^1 \int_0^1 H^0_{vv}\,ds(y + t(v-y) - v_0)\,dt(v-y) \Big\| \\
&\leq C(\|v - v_0\| + \|y - v_0\| + \epsilon)(\|v - v_0\| + \|y - v_0\|)\,\|v - y\| \\
&\quad + \|QH^0_{vv}(\tfrac{1}{2}(v+y) - v_0)(v-y)\| \\
&\leq C\epsilon^{2\alpha}\,\|v - y\| + \|QH^0_{vv}[(I-P)(\tfrac{1}{2}(v+y) - v_0)][(I-P)(v-y)]\| \\
&\leq C\{\epsilon^{2\alpha}(\|P(v-y)\| + \|(I-P)(v-y)\|) + \epsilon\,\|(I-P)(v-y)\|\} \\
&\leq C\{\epsilon^{2\alpha}\,\|P(v-y)\| + \epsilon\,\|(I-P)(v-y)\|\}.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\|(I-Q)z\| &\leq C(\|v - v_0\| + \|y - v_0\|)\,\|v - y\| \\
&\leq C\epsilon^{\alpha}(\|P(v-y)\| + \|(I-P)(v-y)\|).
\end{aligned}$$

Summarizing these estimates we obtain

$$|T(v, \epsilon) - T(y, \epsilon)| \leq C\begin{bmatrix} \epsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \epsilon^{2\alpha} & \epsilon \\ \epsilon^{\alpha} & \epsilon^{\alpha} \end{bmatrix}|v - y| \leq L_\epsilon\,|v - y|$$

where

$$L_\epsilon = C\begin{bmatrix} \epsilon^{2\alpha-1} & 1 \\ \epsilon^{\alpha} & \epsilon^{\alpha} \end{bmatrix}.$$

With this Lipschitz bound we calculate

$$(I - L_\epsilon)r_\epsilon = \begin{pmatrix} \epsilon^{\alpha} - C\epsilon^{2\alpha-1} - Cc_0\epsilon \\ c_0\epsilon - C\epsilon^{2\alpha} - Cc_0\epsilon^{\alpha+1} \end{pmatrix}.$$

Thus, for ϵ sufficiently small, (B7) yields

$$(I - L_\epsilon)r_\epsilon \geq \tfrac{2}{3}r_\epsilon > |T(v_0, \epsilon) - v_0|.$$

An application of Theorem B1 finishes the proof.

REMARK From (B4) and (B5) we obtain

$$\begin{aligned}
|v - y| &\leq C\begin{bmatrix} 1 & 1 \\ \epsilon^{\alpha} & 1 \end{bmatrix}|H_v(v_0, \epsilon)^{-1}(H(v, \epsilon) - H(y, \epsilon))| \\
&\leq C\begin{bmatrix} \epsilon^{-1} & 1 \\ \epsilon^{\alpha-1} & 1 \end{bmatrix}|H(v, \epsilon) - h(y, \epsilon)| \qquad \text{for } v, y \in K_\epsilon. \qquad \text{(B8)}
\end{aligned}$$

This is the non-linear generalization of (B5).

Proof of Lemma B2. Let us approximate $H_v(v_0, \epsilon)$ by the linear block-diagonal operator $A(\epsilon) := \epsilon Q H_{v\epsilon}^0 P + (I - Q) H_v^0 (I - P)$. Expanding $H_v(v_0, \epsilon)$ with respect to ϵ it is easily seen that

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

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

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

Both inequalities together imply for $v \in V$

$$|H_v(v_0, \epsilon)v| \geq M(\epsilon)\,|v|, \qquad M(\epsilon) = \begin{bmatrix} C_1\epsilon - C_2\epsilon^2 & -C_2\epsilon \\ -C_2\epsilon & C_1 - C_2\epsilon \end{bmatrix}. \tag{B9}$$

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

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

Therefore, (B9) yields

$$|v| \leq C \begin{bmatrix} \epsilon^{-1} & 1 \\ 1 & 1 \end{bmatrix} |H_v(v_0, \epsilon)v| \qquad \forall v \in V. \tag{B10}$$

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

Note on Newton's method

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

$$T(v, \epsilon) = v - H_v(v, \epsilon)^{-1} H(v, \epsilon), \qquad v \in K_\epsilon.$$

First, we need an analogue of Lemma B2

$$|H_v(y, \epsilon)^{-1} v| \leq C \begin{bmatrix} \epsilon^{-1} & 1 \\ \epsilon^{\alpha-1} & 1 \end{bmatrix} |v| \qquad \text{for } y \in K_\epsilon, \qquad v \in V.$$

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 $v(\epsilon) \in K_\epsilon$ and show

$$|T(v, \epsilon) - T(v(\epsilon), \epsilon)| \leq Q_\epsilon\,|v - v(\epsilon)|^2, \qquad v \in K_\epsilon, \tag{B11}$$

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

$$Q_\epsilon \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^2 = C \begin{pmatrix} \epsilon^{\alpha-1} x_1^2 + \epsilon^{-1} x_2^2 \\ x_1^2 + \epsilon^{\alpha-1} x_2^2 \end{pmatrix}.$$

Similar to the chord method the crucial estimate uses assumption (v).

$$\begin{aligned}\|Q[H(v(\epsilon),\epsilon)-(H(v,\epsilon)+H_v(v,\epsilon)(v(\epsilon)-v))]\| \\ \leq \left\|Q\int_0^1 (1-t)[H_{vv}(v+t(v(\epsilon)-v))-H^0_{vv}](v(\epsilon)-v)^2\,dt\right\| \\ +\tfrac{1}{2}\|H^0_{vv}[(I-P)(v(\epsilon)-v)]^2\| \\ \leq C(\epsilon^\alpha\|P(v(\epsilon)-v)\|^2+\|(I-P)(v(\epsilon)-v)\|^2).\end{aligned}$$

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

$$|T(v,\epsilon)-T(v(\epsilon),\epsilon)|\leq C\begin{bmatrix}\epsilon^{2\alpha-1} & 1\\ \epsilon^\alpha & \epsilon^\alpha\end{bmatrix}|v-v(\epsilon)|, \qquad v\in K_\epsilon$$

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_ϵ on ϵ we have to take an ϵ-dependent norm for this purpose. For example,

$$\|v\|_\epsilon=[\epsilon^\rho\|Pv\|^2+(\epsilon^\rho\|(I-P)v\|)^2]^{\frac{1}{2}}.$$

with $\rho=\min(2(\alpha-1),-\frac{2}{3})$ 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

Arnold, V. I. 1988 Geometrical methods in the theory of ordinary differential equations. *Grundlehren der Mathematischen Einzelwissenschaften*, Berlin: Springer.

Bai, F., Spence, A., & Stuart, A. M. 1992 The numerical computation of heteroclinic connections in systems of gradient partial differential equations. Preprint, University of Bath.

Beyn, W.-J. 1990*a* The numerical computation of connecting orbits in dynamical systems. *IMA J. Numer. Anal.* **9**, 379–405.

Beyn, W.-J. 1990*b* Global bifurcations and their numerical computation. *Continuation and Bifurcations: Numerical Techniques and Applications (NATO ASI, Series C)* vol. 313 (D. Roose, B. de Dier and A. Spence, eds.). Kluwer, 169–181.

Beyn, W.-J. 1991 Numerical methods for dynamical systems. *Advances in Numerical Analysis, vol. I, Nonlinear Partial Differential Equations and Dynamical Systems* (W. Light, ed.). Clarendon Press, 175–236.

Bogdanov, R. I. 1975 Versal deformations of a singular point on the plane in the case of zero eigenvalues. *Funct. Anal. Appl.* **9**, 144–145.

Bohl, E. 1978 P-boundedness of inverses of nonlinear operators. *ZAMM* **58**, 277–287.

Bykov, V. I., Yablonskii, G. S., & Kim, V. F. 1978 On one simple model of kinetic self-oscillations in catalytic reaction of CO oxidation. *Dokl. Sov. Math.*, 637–639.

Carr, J. 1981 *Applications of Centre Manifold Theory*. New York: Springer.

CHAN, T. E. 1984 Deflation techniques and block-elimination algorithms for solving bordered singular systems. *SIAM J. Sci. Stat. Comput.* **5,** 121–134.

CHAY, R. R., & KEIZER, J. 1983 Minimal model for membrane oscillations in the pancreatic β-cell. *Biophys. J.* **42,** 181–190.

CRANDALL, M. G., & RABINOWITZ, P. H. 1971 Bifurcation from simple eigenvalues. *J. Funct. Anal.* **8,** 321–340.

DECKER, D. W., KELLER, H. B., & KELLEY, C. T. 1983 Convergence rates for Newton's method at singular points. *SIAM J. Numer. Anal.* **20,** 296–314.

DOEDEL, E. J., & FRIEDMAN, M. J. 1989 Numerical computation of heteroclinic orbits. *J. Comput. Appl. Math.* **26,** 159–170.

DOEDEL, E. J., & KERNÉVEZ, J. P. 1986 AUTO: Software for continuation and bifurcation problems in ordinary differential equations. *Appl. Math. Technical Report,* CALTECH.

FRIEDMAN, M. J., & DOEDEL, E. J. 1992 Numerical computation and continuation of invariant manifolds connecting fixed points. *SIAM J. Numer. Anal.* **28,** 789–808.

GASPARD, P. 1992 Local birth of homoclinic chaos. Preprint, Université Libre de Bruxelles.

GOVAERTS, W. 1991 Stable solvers and block elimination for bordered systems. *SIAM J. Matrix Anal. Appl.* **12,** 469–483.

GOVAERTS, W., & PRYCE, J. D. 1991 Mixed block elimination for linear systems with wider borders. Preprint.

GRIEWANK, A., & REDDIEN, G. W. 1989 Computation of cusp singularities for operator equations and their discretizations. *J. Comput. Appl. Math.* **26,** 133–153.

GUCKENHEIMER, J. 1984 Multiple bifurcation problems in codimension two. *SIAM J. Math. Anal.* **15,** 1–49.

GUCKENHEIMER, J., & HOLMES, PH. 1983 Nonlinear oscillations, dynamical systems, and bifurcations of vector fields. *Appl. Math. Sci.* vol 42. New York: Springer.

HALE, J. 1983 Introduction to dynamic bifurcation. *Bifurcation Theory and Applications* (*Springer Lecture Notes in Mathematics, vol. 1057*) (L. Salvadori, ed.). Berlin: Springer.

JEPSON, A. D., & DECKER, D. W. 1986 Convergence cones near bifurcation. *SIAM J. Numer. Anal.* **23,** 959–975.

KELLER, H. B. 1977 Constructive methods for bifurcation and nonlinear eigenvalue problems. *Computing Methods in Applied Sciences and Engineering* (*Springer Lecture Notes, vol. 704*) (R. Glowinski and J. L. Lions, eds.) Berlin: Springer, 241–251.

KELLER, H. B., & JEPSON, A. D. 1984 Steady state and periodic solution paths: their bifurcation and computations. *Numerical Methods for Bifurcation Problems* ISNM 70 (T. Küpper, H. D. Mittelmann, and H. Weber, eds.). Birkhäuser, 219–246.

KIRK, V. 1991 Breaking of symmetry in the saddle-node Hopf bifurcation. *Phys. Lett.* A **154,** 243–248.

KHIBNIK, A. I. 1990 LINLBF: a program for continuation and bifurcation analysis of equilibria up to codimenson three. *Continuation and Bifurcations: Numerical Techniques and Applicatins* (*NATO ASI, Series C*) vol. 313. (D. Roose, B de Dier and A. Spence, eds.). Kluwer, 283–296.

KHIBNIK, A. I., KUZNETSOV, Y. A., LEVITIN, V. L., & NIKOLAEV, E. V., 1993 Continuation techniques and interactive software for bifurcation analysis of ODEs and iterated maps. *Physica* D, special issue: *Proceedings of the NATO Advanced Research Workshop on Homoclinic Chaos,* (*Brussels,* 1991). To appear.

KUZNETSOV, Y. A. 1990 Computation of invariant manifold bifurcations. *Continuation and Bifurcations: Numerical Techniques and Applications* (*NATO ASI, Series C*) vol. 313 (D. Roose, B. de Dier and A. Spence, eds). Kluwer, 183–195.

LANGFORD, W. F. 1979 Periodic and steading-state mode interaction lead to tori. *SIAM J. Appl. Math.* **37,** 22–48.

MELNIKOV, V. K. 1963 On the stability of the center for time periodic perturbations. *Trans. Moscow Math. Soc.* **12,** 1–57.

MOORE, G. 1987 Some remarks on the deflated block elimination method. *Bifurcation: Analysis, Algorithms, Applications* ISNM 79 (T. Küpper, R. Seydel and H. Troger, eds.). Birkhäuser, 222–234.

MOROZOV, E. I. 1982 Bifurcation of a vector field near a singular point in the case of two pairs of imaginary eigenvalues. *Comptes Rendues de l'Acad. Bulgaires des Sci.* **35,** 149–152.

PALMER, K. J. 1984 Exponential dichotomies and transversal homoclinic points. *J. Differ. Equations* **55,** 225–256.

PÖNISCH, G. 1991 An indirect method for computing origins for Hopf bifurcation in two-parameter problems. *Computing* **46,** 307–320.

RINZEL, J. 1985 Bursting oscillations in an excitable membrane model. *Ordinary and Partial Differential Equations, (Proceedings Dundee) (Springer Lecture Notes in Mathematics, vol. 1151)* (B. D. Sleeman and R. J. Jarvis, eds.). Berlin: Springer, 304–316.

RODRIGUEZ LUIS, A. J., FREIRE, E., & PONCE, E. 1990 A method for homoclinic and heteroclinic continuation in two and three dimensions. *Continuation and Bifurcations: Numerical Techniques and Applications (NATO ASI, Series C)* vol. 313 (D. Roose, B. de Dier and A. Spence, eds.). Kluwer, 197–210.

ROOSE, D. 1987 Numerical computation of origins for Hopf bifurcation in a two-parameter problem. *Bifurcation: Analysis, Algorithms, Applications* ISNM 79. (T. Küpper, R. Seydel and H. Troger, eds.) Birkhäuser, 268–273.

ROOSE, D., DE DIER, B., & SPENCE, A. (eds.) 1990 *Continuation and Bifurcations: Numerical Techniques and Applications (NATO ASI, Series C)* vol. 313. Kluwer.

SPENCE, A., CLIFFE, K. A., & JEPSON, A. D. 1989 A note on the calculation of paths of Hopf bifurcations. *J. Comput. Appl. Math.* **26,** 125–131.

STIEFENHOFER, M. 1991 Private communication.

TAKENS, F. 1974 Singularities of vector fields. *Publ. Math. IHES* **43,** 47–100.

WIGGINS, S. 1990 *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* Berlin: Springer.