Schwerpunktprogramm der
Deutschen Forschungsgemeinschaft
Ergodentheorie, Analysis und effiziente Simulation
dynamischer Systeme

Numerical Taylor expansions of invariant manifolds in large dynamical systems

W.-J. Beyn and W. Kleß

Preprint 6/96

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

version from 6 Aug. 1997
submitted to
"Numerische Mathematik"

Numerical Taylor expansions of invariant manifolds in large dynamical systems

Wolf–Jürgen Beyn*
email: beyn@mathematik.uni-bielefeld.de

Winfried Kleß*
email: kless@mathematik.uni-bielefeld.de

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

Abstract

In this paper we develop a numerical method for computing higher order local approximations of invariant manifolds, such as stable, unstable or center manifolds near steady states of a dynamical system. The underlying system is assumed to be large in the sense that a large sparse Jacobian at the equilibrium occurs, for which only a linear (black box) solver and a low dimensional invariant subspace is available, but for which methods like the QR–Algorithm are considered to be too expensive. Our method is based on an analysis of the multilinear Sylvester equations for the higher derivatives which can be solved under certain nonresonance conditions. These conditions are weaker than the standard gap conditions on the spectrum which guarantee the existence of the invariant manifold. The final algorithm requires the solution of several large linear systems with a bordered Jacobian. To these systems we apply a block elimination method recently developed by Govaerts and Pryce (1991, 1993).

Keywords: Invariant manifolds, numerical methods, higher order approximations, large dynamical systems.

Mathematics Subject Classification (1991): 34C35, 58F14, 65F50, 65H17.

1 Introduction

Our primary goal is the local approximation of invariant manifolds near a steady state of a smooth dynamical system

$$\dot{u} = F(u), \quad F : \mathbb{R}^m \longrightarrow \mathbb{R}^m. \tag{1.1}$$

We assume that the spectrum of the linearization $DF(u_0)$ at the equilibrium u_0 can be decomposed as

$$\sigma\big(DF(u_0)\big) = \sigma_l \cup \sigma_h, \quad \sigma_l \cap \sigma_h = \emptyset. \tag{1.2}$$

*supported by DFG Schwerpunktprogramm “Ergodentheorie, Analysis und effiziente Simulation dynamischer Systeme”.

Here σ_l contains a few eigenvalues (usually close to or on the imaginary axis) and the corresponding invariant subspace $E_l \subset \mathbb{R}^m$ is known, whereas σ_h contains the remaining (usually large) part of the spectrum with an invariant subspace $E_h \subset \mathbb{R}^m$ which is generally unknown. The indices 'l' and 'h' will be used to indicate low and high dimensional objects.

Under some well–known gap conditions on the spectrum, such as the following (cf. [38])

$$\mathrm{Re}\ \sigma_h < \alpha < \beta < \mathrm{Re}\ \sigma_l \quad \text{and} \quad \alpha < r \cdot \beta \tag{GC$_r$}$$

for some $r \in \mathbb{N}$, there exists a locally invariant C^r–manifold which may be written as a graph over E_l

$$W = \{(x, g(x)) : x \in V_l \subset E_l\}, \quad V_l \text{ a suitable neighborhood in } E_l, \tag{1.3}$$

for some C^r–function $g : V_l \to E_h$. Depending on the position of the spectral set σ_l, this manifold is called strongly unstable ($0 < \mathrm{Re}\ \sigma_l$), center–unstable ($0 \le \mathrm{Re}\ \sigma_l$), center ($0 = \mathrm{Re}\ \sigma_l$) or pseudo unstable ($\beta < \mathrm{Re}\ \sigma_l$, $\beta < 0$), and further types occur under time reversal of system (1.1) (cf. [3], [38]).

Assuming that the remaining spectrum is stable, the dynamics of (1.1) close to u_0 can be reduced to the low dimensional system

$$\dot{x} = P_l F(x, g(x)), \tag{1.4}$$

where P_l is the projection onto E_l along E_h. Reduction processes of this type are very well studied analytically and the applications to ordinary and partial differential equations are abundant (for example [7], [16], [21], [26], [29], [41]).

In this paper we develop a numerical method for computing the higher derivatives of the function g under the sole assumption that σ_l, E_l are known, that a linear solver for the translates $DF(u_0) - \lambda I$ is provided and that the same information on the transpose $DF(u_0)^T$ is available. Such a situation typically occurs for spatial discretizations of time–dependent partial differential equations. In some of these applications even the requirement on the transpose might be too restrictive.

In section 2 we summarize some well–known facts from invariant manifold theory and we formally set up the equations for the higher derivatives $D^n g$, $n = 1, \dots, r$ at the equilibrium. These are linear equations for the multilinear forms and, since for the first derivative one obtains the well–known matrix equation of Sylvester (see [40]), we call them the multilinear Sylvester equations. We show that a suitable reformulation of the multilinear Sylvester equation does not only determine $D^n g$ but at the same time the n–th derivative of the reduced vector field in (1.4).

In section 3 we show that the multilinear Sylvester equations can be solved via a generalization of Rosenblum's formula (cf. [36]) under the nonresonance conditions

$$\left\{ \sum_{\nu=1}^{n} \lambda_\nu : \lambda_\nu \in \sigma_l \right\} \cap \sigma_h = \emptyset \tag{NRC$_n$}$$

for $n = 1, \dots, r$. These conditions are weaker than the gap condition (GC)$_r$. In the linearization theory the nonresonance conditions are imposed on any spectral decomposition $\sigma_l \cup \sigma_h$ and this leads to the well–known Sternberg conditions (see [32], [37], [17], [23]).

While the complex integrals in the generalized Rosenblum formula have the advantage of being generalizable to an infinite dimensional setting, the formula itself seems to be of little practical use.

As an alternative we set up a solution method which generalizes the well–known Bartels–Stewart algorithm for the linear Sylvester equation, (cf. [11]). However, this algorithm still does not meet the above requirements because it works with a representation of the Jacobian $DF(u_0)$ on the large space E_h.

Therefore, in section 4 we develop a new version of this algorithm in which all computations are based on solving linear systems with bordered Jacobians of the form

$$\Gamma = \begin{pmatrix} DF(u_0) - \lambda I & \varphi \\ \psi^T & 0 \end{pmatrix}. \tag{1.5}$$

Here the number of columns in φ and ψ equals the dimension of the low dimensional space E_l and λ is a shift parameter of the form $\lambda = \sum_{\nu=1}^{n} \lambda_\nu,\ \ \lambda_\nu \in \sigma_l$.

For these linear systems we use the so–called mixed block elimination method developed by Govaerts and Pryce ([12], [14]). The remarkable feature of their algorithm is that it works with a black box solver for the large matrix $DF(u_0) - \lambda I$. The only requirements are that the complete matrix Γ in (1.5) is well–conditioned and that the solver for the large matrix $DF(u_0) - \lambda I$ is normwise backward stable in the sense of [18]; the matrix $DF(u_0) - \lambda I$ itself however may be almost singular.

Some generalizations of our algorithm to discrete dynamical systems are indicated in sections 3 and 4, and an application to the Brusselator system is provided in section 5.

Let us finally remark that the overall algorithm is not a completely new approach but is inherent in many concrete analytical computations (see for example the projector method in [26]). But we think it is useful to show that there is a general theory for the multilinear Sylvester equation and that this leads to a general solution algorithm which works under the linear algebra restrictions described above.

2 Prerequisites from invariant manifold theory

2.1 Locally invariant manifolds in the case of general spectral decomposition

Consider the autonomous first order differential equation

$$\frac{du}{dt} \stackrel{\text{def}}{=} \dot{u} = F(u), \tag{2.1}$$

where F is in $C^r(\Omega, \mathbb{R}^m)$, Ω is an open subset of $\mathbb{R}^m$ and r is a natural number. Throughout, we assume that $u_0 \in \Omega$ is a steady state of (2.1) and we use A to denote the Jacobian $DF(u_0)$ at u_0. Locally invariant C^r–manifolds near u_0 are known to exist under certain gap conditions on the spectrum $\sigma(A)$. For the following theorem, adopted from Shub (1987) [38] (see also [20]), we assume that real numbers $\alpha < \beta < 0 < \gamma < \delta$ are given, such that the sets

$$\sigma_s := \{\lambda \in \sigma(A) \,:\, Re\,\lambda < \alpha\},$$

$$\sigma_c := \{\lambda \in \sigma(A) \,:\, \beta < Re\,\lambda < \gamma\},$$

$$\sigma_u := \{\lambda \in \sigma(A) \,:\, \delta < Re\,\lambda\}$$

Fig. 1 form a disjoint decomposition of the spectrum $\sigma(A)$ (Figure 1).

Let E_s, E_c and E_u denote the corresponding invariant subspaces of $\mathbb{R}^m$ with respect to this decomposition, i.e. for $* \in \{s, c, u\}$ the space E_* is the maximal subspace of $\mathbb{R}^m$ which satisfies

$$A(E_*) \subset E_* \quad \text{and} \quad \sigma(A|_{E_*}) = \sigma_*.$$

Theorem 2.1: *Under the above assumptions there exist zero neighborhoods $V_s \subset E_s$, $V_c \subset E_c$, $V_u \subset E_u$ and Lipschitz functions $g_{ss} : V_s \to V_c \oplus V_u$, $g_{sc} : V_s \oplus V_c \to V_u$, $g_c : V_c \to V_s \oplus V_u$, $g_{cu} : V_c \oplus V_u \to V_s$ as well as $g_{uu} : V_u \to V_s \oplus V_c$, such that each of these five functions maps 0 onto 0 and has a vanishing derivative at the origin. The shifted graphs of these functions*

$$W^V_{ss}(u_0) := \{u_0 + v_{ss} + g_{ss}(v_{ss}) \,:\, v_{ss} \in V_s\},$$

$$W^V_{sc}(u_0) := \{u_0 + v_{sc} + g_{sc}(v_{sc}) \,:\, v_{sc} \in V_s \oplus V_c\},$$

$$W^V_{c}(u_0) := \{u_0 + v_{c} + g_{c}(v_{c}) \,:\, v_{c} \in V_c\},$$

$$W^V_{cu}(u_0) := \{u_0 + v_{cu} + g_{cu}(v_{cu}) \,:\, v_{cu} \in V_c \oplus V_u\} \quad \textit{and}$$

$$W^V_{uu}(u_0) := \{u_0 + v_{uu} + g_{uu}(v_{uu}) \,:\, v_{uu} \in V_u\}$$

are locally invariant manifolds for the flow of (2.1).

Moreover, on suitable neighborhoods one can choose g_{ss} and g_{uu} as C^r–functions. For the C^r smoothness of the three remaining functions the following **spectral gap conditions** *are sufficient: g_{cu} is in C^r if*

$$\alpha < r \cdot \beta \tag{2.2}$$

and g_{sc} is in C^r under the condition

$$r \cdot \gamma < \delta. \tag{2.3}$$

If both gap conditions (2.2) and (2.3) are satisfied, then also g_c may be chosen as a C^r–function.

We recall that the local invariance of any of the manifolds W in Theorem 2.1 holds with respect to the neighborhood $V := u_0 + V_s \oplus V_c \oplus V_u$, which means that the flow ϕ of the system (2.1) satisfies

$$v \in M,\ \phi(t, v) \in V \text{ for } t_0 < t < t_1,\ t_0 < 0 < t_1 \quad \Longrightarrow \quad \phi(t, v) \in M \text{ for } t_0 < t < t_1.$$

The existence and smoothness of the strong stable and strong unstable manifold have been known for a long time (for example Irwin (1972) [22] and 1980 [23] in the discrete case; see also [6]). For the smoothness of the generalized center stable, center and center unstable manifold, see [20] and [38]. A short overview of various methods of proof can be found in [3], and a reference to Hölder continuity is [19].

2.2 Taylor expansion

In the next sections we assume that the spectrum of the Jacobian A splits into two parts, σ_l and $\sigma_h = \sigma(A) \setminus \sigma_l$, with

$$\sigma_l = \overline{\sigma_l} \quad \text{and} \quad \sigma_h = \overline{\sigma_h}, \tag{2.4}$$

where the bar denotes the conjugate complex. Let $\mathbb{R}^m = E_l \oplus E_h$ be the associated decomposition into invariant subspaces of A and let P_l, P_h be the corresponding projectors. Moreover, we set

$$F_l(u) := P_l F(u) \quad \text{and} \quad F_h(u) := P_h F(u). \tag{2.5}$$

We assume the existence of a locally invariant C^r manifold

$$W = \{(x, g(x)) \,:\, x \in V_l\}, \tag{2.6}$$

where $g \in C^r(V_l, E_h)$ is defined on an open set $V_l \subset E_l$ containing $P_l u_0 =: u_{0,l}$ and satisfies $g(u_{0,l}) = P_h u_0 =: u_{0,h}$.

It is well–known that W is locally invariant if and only if $g(x)$ satisfies the differential equation

$$g'(x) F_l(x, g(x)) = F_h(x, g(x)) = P_h F(x, g(x)), \quad x \in V_l. \tag{2.7}$$

This means that the vector field restricted to W is tangent to the manifold. On this manifold the dynamics is then determined by the **reduced system**

$$\dot{x} = F_l(x, g(x)) = P_l F(x, g(x)). \tag{2.8}$$

We are interested in the local approximation of the manifold W and the reduced vector field, hence we expand F and g in their Taylor-series at the steady state. In the following it is convenient to introduce the function

$$\widetilde{g}(x) := (x, g(x)), \quad x \in V_l,$$

and to abbreviate evaluations at u_0, $u_{0,l}$ by an upper index '^{0}', for example

$$D^\nu F^0 = D^\nu F(u_0), \quad D^\nu g^0 = D^\nu g(u_{0,l}), \quad D^\nu \widetilde{g}^0 = D^\nu \widetilde{g}(u_{0,l}).$$

In order to express higher derivatives of equation (2.7) we employ the higher order Leibniz and chain rule (cf. [2], [32], [1], notice the inconsistencies in all of them). To avoid factorials and to reduce the number of indices we write down simplified versions of the formulae. If B is some bilinear form and f_1, f_2 are smooth maps, then the Leibniz rule for the n–th derivative of $B(f, g)(x)$ reads as follows

$$D^n B(f_1, f_2)(x)(h_1, \ldots, h_n) = \sum_{\pi \in \mathcal{P}_n} B\big(D^{|\pi_1|} f_1(x) h_{\pi_1}, D^{|\pi_2|} f_2(x) h_{\pi_2}\big), \tag{2.9}$$

where the sum is taken over all partitions $\pi = (\pi_1, \pi_2) \in \mathcal{P}_n$ of $\{1, \ldots, n\}$ into two disjoint subsets π_1 and π_2 (with $\pi_1 = \emptyset$ or $\pi_2 = \emptyset$ included). For any finite subset $\omega = \{\omega_1, \ldots, \omega_n\}$ of $\mathbb{N}$, let $n = |\omega| \in \mathbb{N}_0$ denote its cardinality, and ordering the elements according to $\omega_1 < \cdots < \omega_n$ we use the notation

$$h_\omega = (h_{\omega_1}, \ldots, h_{\omega_n}), \quad h_\emptyset = 1.$$

The precise definition of $\mathcal{P}_n$ is

$$\mathcal{P}_n = \big\{\pi = (\pi_1, \pi_2) \,:\, \pi_1, \pi_2 \subset \{1, \dots, n\},\ \pi_1 \cup \pi_2 = \{1, \dots, n\},\ \pi_1 \cap \pi_2 = \emptyset\big\}.$$

For the chain rule of order $n \geq 1$ it is convenient to consider nontrivial, ordered partitions $\pi = (\pi_1, \dots, \pi_k)$ of a given subset $\omega \subset \mathbb{N}$, $n = |\omega|$ as follows

$$\emptyset \neq \pi_i \subset \omega, \quad \bigcup_{i=1}^{k} \pi_i = \omega, \quad \pi_i \cap \pi_j = \emptyset \quad \text{and} \quad \min \pi_i < \min \pi_j \quad \text{for all} \quad i < j.$$

Let $\mathcal{Q}_k(\omega)$ denote the set of all such partitions $\pi = (\pi_1, \dots, \pi_k)$ and let $|\pi| = k$ denote their length. Moreover, we set $\mathcal{Q}(\omega) = \bigcup_{k=1}^{n} \mathcal{Q}_k(\omega)$, $\mathcal{Q}(\emptyset) = \{\emptyset\}$ for formal reasons, and we abbreviate $\mathcal{Q}(n) = \mathcal{Q}(\{1, \dots, n\})$. With these settings we may write the chain rule of order n for given maps f_1, f_2 and a finite subset $\omega \subset \mathbb{N}$, $|\omega| = n$ in the following form

$$D^n(f_1 \circ f_2)(x) h_\omega \;=\; \sum_{\substack{\pi \in \mathcal{Q}(\omega) \\ (k := |\pi|)}} D^k f_1(f_2(x)) \Big(D^{|\pi_1|} f_2(x) h_{\pi_1}, \dots, D^{|\pi_k|} f_2(x) h_{\pi_k} \Big). \tag{2.10}$$

The proof is easily accomplished by an induction argument on $n = |\omega|$.

Applying the Leibniz rule (2.9) and the chain rule (2.10) the n–th derivative of equation (2.7) reads

$$\begin{aligned}
&\sum_{\substack{\pi \in \mathcal{P}_n \\ (\pi = (\pi_1, \pi_2))}} \sum_{\substack{\tau \in \mathcal{Q}(\pi_2) \\ (k := |\tau|)}} D^{|\pi_1|+1} g^0 \big[h_{\pi_1}, D^k F_l^0 \big(D^{|\tau_1|} \widetilde{g}^0 h_{\tau_1}, \dots, D^{|\tau_k|} \widetilde{g}^0 h_{\tau_k}\big)\big] \\
&\quad = \sum_{\pi \in \mathcal{P}_n} D^{|\pi_1|+1} g^0 \big[h_{\pi_1}, D^{|\pi_2|} (F_l \circ \widetilde{g})^0 h_{\pi_2}\big] \;=\; D^n \big((g') \cdot (F_l \circ \widetilde{g})\big)^0 (h_1, \dots, h_n) \qquad (2.11) \\
&\quad = D^n (F_h \circ \widetilde{g})^0 (h_1, \dots, h_n) \;=\; \sum_{\substack{\pi \in \mathcal{Q}(n) \\ (k := |\pi|)}} D^k F_h^0 \big(D^{|\pi_1|} \widetilde{g}^0 h_{\pi_1}, \dots, D^{|\pi_k|} \widetilde{g}^0 h_{\pi_k}\big)
\end{aligned}$$

for $(h_1, \dots, h_n) \in E_l^n$.

Let $\mathcal{L}^n(E_1, \dots, E_n; E_0)$ denote the Banach space of n–multilinear continuous maps from the Banach spaces $E_1, \dots, E_n$ into the Banach space E_0. If $E_\nu = E_1$, $2 \leq \nu \leq n$, this space is denoted by $\mathcal{L}^n(E_1, E_0)$ and we set $\mathcal{L}(E_1, E_0) := \mathcal{L}^1(E_1; E_0)$.

Because of the invariance of the subspaces E_l and E_h we get

$$D_h F_l(u_0) = 0 \quad \text{and} \quad D_l F_h(u_0) = 0,$$

where D_l, D_h denotes differentiation with respect to E_l and E_h, respectively. By setting

$$A_l := D_l F_l(u_{0,l}, u_{0,h}) \in \mathcal{L}(E_l, E_l),$$

$$A_h := D_h F_h(u_{0,l}, u_{0,h}) \in \mathcal{L}(E_h, E_h),$$

we may write

$$A = A_l \oplus A_h,$$

where

$$\sigma(A_l) = \sigma_l \quad \text{and} \quad \sigma(A_h) = \sigma_h.$$

Equation (2.11) looks almost intractable for general n. But the first cases are easily written down. For $n = 1$ we have the linear Sylvester equation

$$Dg^0 \cdot A_l - A_h \cdot Dg^0 = 0. \tag{2.12}$$

Since $\sigma(A_l) \cap \sigma(A_h) = \sigma_l \cap \sigma_h = \emptyset$, this has the unique solution

$$Dg^0 = 0 \tag{2.13}$$

(cf. [40]). Using this the next equations are

$$\begin{aligned} & D^2g^0(A_lh_1, h_2) + D^2g^0(h_1, A_lh_2) - A_h \cdot D^2g^0(h_1, h_2) \\ &\quad = \quad D_l^2F_h^0(h_1, h_2) \qquad \text{for all} \quad (h_1, h_2) \in E_l \times E_l, \end{aligned}$$

$$\begin{aligned} & D^3g^0(A_lh_1, h_2, h_3) + D^3g^0(h_1, A_lh_2, h_3) + D^3g^0(h_1, h_2, A_lh_3) - A_hD^3g^0(h_1, h_2, h_3) \\ = \;& D_l^3F_h^0(h_1, h_2, h_3) + D_hD_lF_h^0\left[h_1, D^2g^0(h_2, h_3)\right] + D_hD_lF_h^0\left[h_2, D^2g^0(h_1, h_3)\right] \\ & + D_hD_lF_h^0\left[h_3, D^2g^0(h_1, h_2)\right] - D^2g^0\left[h_1, D_l^2F_l^0(h_2, h_3)\right] - D^2g^0\left[h_2, D_l^2F_l^0(h_1, h_3)\right] \\ & - D^2g^0\left[h_3, D_l^2F_l^0(h_1, h_2)\right] \qquad \text{for all} \quad (h_1, h_2, h_3) \in E_l \times E_l \times E_l. \end{aligned}$$

In the general case by application of (2.13) and $F(u_0) = 0$ we delete in (2.11) the summands corresponding to $\pi_1 = \emptyset$ and $\pi_2 = \emptyset$ on the left–hand side. Moreover, we extract the n summands corresponding to $|\pi_1| = n-1$, $|\pi_2| = 1$ from the left–hand side and the summand with $|\pi| = 1$, i. e. $\pi = (\pi_1) = (\{1, \ldots, n\})$, from the right–hand side. This yields the following result.

Theorem 2.2: *For $n = 1, \ldots, r-1$ the n–th derivative $X := D^ng^0 \in \mathcal{L}^n(E_l, E_h)$ at $u_{0,l}$ satisfies the multilinear Sylvester equation*

$$\begin{aligned} \big(L(X)\big)(h_1, \ldots, h_n) \;&:=\; \sum_{\nu=1}^{n} X(h_1, .., h_{\nu-1}, A_lh_\nu, h_{\nu+1}, .., h_n) - A_hX(h_1, \ldots, h_n) \\ &=\; C_{h,n}(h_1, \ldots, h_n) \qquad \textit{for all} \quad (h_1, \ldots, h_n) \in E_l^n, \end{aligned} \tag{2.14}$$

where $C_{h,n} \in \mathcal{L}^n(E_l, E_h)$ is determined by the derivatives $D^\nu F_l^0$, $\nu = 1, \ldots, n-1$, and $D^\nu F_h^0$, $\nu = 1, \ldots, n$, as well as by the previous derivatives $D^2g^0, \ldots, D^{n-1}g^0$.

In the following corollary we reformulate the multilinear Sylvester equation (2.14) in a form useful for later purposes (see section 4.2) and write down an explicit formula for the right–hand side.

Corollary 2.3: (i) *Setting $Y := Y_n := \left\{\begin{matrix} 0 & \text{for } n = 1 \\ D^n(F_l \circ \widetilde{g})^0 & \text{for } n \geq 2 \end{matrix}\right\} \in \mathcal{L}^n(E_l, E_l)$ the multilinear forms X and Y satisfy*

$$L(X) + Y = C_n, \tag{2.15}$$

where $C_n \in \mathcal{L}^n(E_l, \mathbb{R}^m)$ *is determined by the derivatives* $D^\nu F_l^0$, $\nu = 1, \dots, n-1$, *and* $D^\nu F^0$, $\nu = 1, \dots, n$, *as well as by the previous derivatives* $D^2 g^0, \dots, D^{n-1} g^0$. *The explicit form is*

$$\begin{aligned} C_n(h_1, \dots, h_n) &= \sum_{\substack{\pi \in \mathcal{Q}^*(n) \\ (k := |\pi|)}} D^k F^0 \big(D^{|\pi_1|} \widetilde{g}^0 h_{\pi_1}, \dots, D^{|\pi_k|} \widetilde{g}^0 h_{\pi_k} \big) \\ &\quad - \sum_{\pi \in \mathcal{P}_n^*} \sum_{\substack{\tau \in \mathcal{Q}(\pi_2) \\ (k := |\tau|)}} D^{|\pi_1|+1} g^0 \big[h_{\pi_1}, D^k F_l^0 \big(D^{|\tau_1|} \widetilde{g}^0 h_{\tau_1}, \dots, D^{|\tau_k|} \widetilde{g}^0 h_{\tau_k} \big) \big], \end{aligned} \tag{2.16}$$

where we use the abbreviations

$$\mathcal{P}_n^* := \big\{ \pi = (\pi_1, \pi_2) \in \mathcal{P}_n \,:\, \pi_1, \pi_2 \neq \emptyset,\ |\pi_2| \neq 1 \big\} \quad \textit{and} \quad \mathcal{Q}^*(n) := \big\{ \pi \in \mathcal{Q}(n) \,:\, |\pi| > 1 \big\}$$

The derivatives of $\widetilde{g}$ *are given by*

$$D\widetilde{g}^0 = \begin{pmatrix} I \\ Dg^0 \end{pmatrix} = \begin{pmatrix} I \\ 0 \end{pmatrix} \quad \textit{and} \quad D^\nu \widetilde{g}^0 = \begin{pmatrix} 0 \\ D^\nu g^0 \end{pmatrix} \quad \textit{for} \quad \nu \geq 2,$$

and none of the summands in (2.16) contains $D^n \widetilde{g}^0$ *or* $D^n g^0$.

(ii) *The right–hand side* $C_{h,n}$ *in equation (2.14) fulfils*

$$C_{h,n} = P_h C_n.$$

$C_{h,n}$ *has the same form as* C_n *in (2.16), but with* $D^k F^0$ *replaced by* $D^k F_h^0$.

Proof of the corollary: In case $n = 1$ we have $L(X) = 0$, see (2.12). For $n \geq 2$ using the n-th derivative of equation (2.7) we find

$$D^n \big((g') \cdot (F_l \circ \widetilde{g}) \big)^0 = D^n (F \circ \widetilde{g})^0 - Y,$$

and hence the multilinear form C_n equals

$$D^n (F \circ \widetilde{g})^0 - D^n \big((g') \cdot (F_l \circ \widetilde{g}) \big)^0 + L(X).$$

Now we use $F(u_0) = 0$, (2.13) and the chain rule (2.10) on the first term $D^n(F \circ \widetilde{g})^0$ and the Leibniz rule (2.9) on the second term (see (2.11)). Extracting the summands as above our conclusion follows. □

If equation (2.14) has a unique solution $X \in \mathcal{L}^n(E_l, E_h)$ for each $C_{h,n} \in \mathcal{L}^n(E_l, E_h)$, then obviously equation (2.15) has a unique solution $(X, Y) \in \mathcal{L}^n(E_l, E_h) \times \mathcal{L}^n(E_l, E_l)$ for every right–hand side $C_n \in \mathcal{L}^n(E_l, \mathbb{R}^m)$ and vice versa. In case of unique solvability, both Theorem 2.2 and Corollary 2.3 (i) define an iterative process for computing the derivatives

$$D^1 g^0, \dots, D^{n-1} g^0, D^n g^0,$$

from which the formal Taylor expansion of the function g is obtained. Therefore, in the following sections we analyze the solvability of (2.14) as well as methods for its numerical solution.

There are two ways of computing the derivatives of the reduced vector field $F_l \circ \widetilde{g}$ at the steady state $u_{0,l}$.

First, if instead of (2.14) we solve the enlarged system (2.15) (this will be done in section 4) the derivative $Y = D^n(F_l \circ \widetilde{g})^0$ is obtained as a by–product of the computation of the n–th derivative $D^n g^0$.

Second, if the derivative $D^n g^0$ has been computed we can use it to evaluate the next derivative $D^{n+1}(F_l \circ \widetilde{g})^0$ as follows. From equation (2.10) we have

$$D^{n+1}(F_l \circ \widetilde{g})(x) = G(x) + D_h F_l(x, g(x)) \cdot D^{n+1} g(x), \quad x \in V_l, \ n \in \{1, \dots, r-1\},$$

with a function $G(x) \in \mathcal{L}^{n+1}(E_l, E_h)$ that can be computed from $D^1 g(x), \dots, D^n g(x)$ and $D^1 F_l(x, g(x)), \dots, D^{n+1} F_l(x, g(x))$. Since $D_h F_l^0 = 0$, we can evaluate the $(n+1)$–th derivative at the steady state through

$$D^{n+1}(F_l \circ \widetilde{g})^0 = G^0. \tag{2.17}$$

The first terms in a Taylor expansion of the reduced system (2.8) are given by

$$\begin{aligned} \dot{\xi} = A_l \xi + \frac{1}{2} D_l^2 F_l^0 \xi^2 + \frac{1}{6}\Big(D_l^3 F_l^0 \xi^3 + 3 D_h D_l F_l^0 [\xi, D^2 g^0 \xi^2]\Big) + \frac{1}{24}\Big(D_l^4 F_l^0 \xi^4 + \\ 6 D_h D_l^2 F_l^0 [\xi^2, D^2 g^0 \xi^2] + 3 D_h^2 F_l^0 [D^2 g^0 \xi^2]^2 + 4 D_h D_l F_l^0 [\xi, D^3 g^0 \xi^3]\Big) + \dots \end{aligned} \tag{2.18}$$

with $\xi = x - u_{0,l}$.

For a fair comparison of the second method with the first we would have to go up to order $n+1$ in the first method. This seems to be more costly than the second approach and, moreover, it requires the conditions of solvability to hold up to order $n+1$ rather than n.

In summary, we propose to evaluate the Taylor expansion of order r for the reduced vector field by solving (2.15) successively for $n = 2, \dots, r-1$. This automatically yields $D^n(F_l \circ \widetilde{g})^0$, $n = 2, \dots, r-1$ and $D^r(F_l \circ \widetilde{g})^0$ is then computed from (2.17) with $n = r-1$.

2.3 The mapping case

We briefly report the analogs of the previous results for a discrete dynamical system

$$u^{n+1} = F(u^n), \quad n \in \mathbb{Z}, \tag{2.19}$$

with a regular fixed point u_0. Now the spectrum of $A := DF(u_0)$ is decomposed into the annuli $0 < |\lambda| < \alpha$ and $\beta < |\lambda| < \gamma$ and $\delta < |\lambda|$, where $\beta < 1 < \gamma$. The analog of Theorem 2.1 then asserts the existence of locally invariant manifolds tangent to the corresponding invariant subspaces, and the gap conditions for C^r–smoothness read:

$$\alpha < \beta^r \quad \text{or} \quad \gamma^r < \delta, \tag{2.20}$$

respectively (see Shub (1987) [38]).

The local invariance of the manifold $W = \{(x, g(x)) : x \in V_l\}$ can now be characterized by the functional equation

$$g\big(F_l(x, g(x))\big) = F_h(x, g(x)) \tag{2.21}$$

for all $x \in V_l$, such that $F_l(x, g(x)) \in V_l$, and the reduced system is

$$x^{n+1} = F_l(x^n, g(x^n)).$$

As above, for the expansion of g and the reduced map the higher order derivatives of equation (2.21) are computed and evaluated at the fixed point. One obtains the following analog of Theorem 2.2.

Theorem 2.4: *For $n \in \{1, \dots, r\}$, the multilinear form $X := D^n g^0 \in \mathcal{L}^n(E_l, E_h)$ satisfies the equation*

$$\begin{aligned} X(A_l h_1, \dots, A_l h_n) - A_h \cdot X(h_1, \dots, h_n) \\ = C_{h,n}(h_1, \dots, h_n) \qquad \text{for all} \quad (h_1, \dots, h_n) \in E_l^n, \end{aligned} \tag{2.22}$$

where $C_{h,n} \in \mathcal{L}^n(E_l, E_h)$ is determined by derivatives of F_l and F_h of order less than or equal to n evaluated at u_0 and by $D^1 g^0, \dots, D^{n-1} g^0$.

In the case $n = 1$, one finds $C_{h,1} = 0$, therefore, $Dg^0 = 0$ is the unique solution of the linear Sylvester equation

$$Dg^0 \cdot A_l - A_h \cdot Dg^0 = 0.$$

Using this result the next equations for $D^2 g^0$ and $D^3 g^0$ are given by

$$D^2 g^0(A_l h_1, A_l h_2) - A_h \cdot D^2 g^0(h_1, h_2) = D_l^2 F_h^0(h_1, h_2)$$

and

$$\begin{aligned} & D^3 g^0(A_l h_1, A_l h_2, A_l h_3) - A_h \cdot D^3 g^0(h_1, h_2, h_3) \\ = \ & D_l^3 F_h^0(h_1, h_2, h_3) + D_h D_l F_h^0 \left[h_1, D^2 g^0(h_2, h_3)\right] + D_h D_l F_h^0 \left[h_2, D^2 g^0(h_1, h_3)\right] \\ & + D_h D_l F_h^0 \left[h_3, D^2 g^0(h_1, h_2)\right] - D^2 g^0 \left[A_l h_1, D_l^2 F_l^0(h_2, h_3)\right] \\ & - D^2 g^0 \left[A_l h_2, D_l^2 F_l^0(h_1, h_3)\right] - D^2 g^0 \left[A_l h_3, D_l^2 F_l^0(h_1, h_2)\right]. \end{aligned}$$

3 The multilinear Sylvester equation

In this section E_0 and E_ν ($\nu = 1, \dots, n$) denote subspaces of $\mathbb{C}^m$ or $\mathbb{R}^m$, respectively. For given maps $A_\nu \in \mathcal{L}(E_\nu, E_\nu)$, $B \in \mathcal{L}(E_0, E_0)$ and $C \in \mathcal{L}^n(E_1, \dots, E_n; E_0)$ we seek a solution $X \in \mathcal{L}^n(E_1, \dots, E_n; E_0)$ of the multilinear Sylvester equation

$$\begin{aligned} \sum_{\nu=1}^{n} X(h_1, \dots, h_{\nu-1}, A_\nu h_\nu, h_{\nu+1}, \dots, h_n) - BX(h_1, \dots, h_n) \\ = C(h_1, \dots, h_n) \qquad \text{for all} \quad (h_1, \dots, h_n) \in E_1 \times \dots \times E_n. \end{aligned} \tag{3.1}$$

Notice that this is a linear system for $X \in \mathcal{L}^n(E_1, \ldots, E_n; E_0)$. We start with a theoretical result.

3.1 A generalization of Rosenblum's formula

In case $n = 1$ equation (3.1) turns into the linear Sylvester equation

$$X \cdot A - B \cdot X = C. \tag{3.2}$$

Under the spectral condition

$$\sigma(A) \cap \sigma(B) = \emptyset$$

this equation has a unique solution $X \in \mathcal{L}(E_1, E_0)$ given by Rosenblum's formula

$$X = \frac{1}{2\pi i} \int_{\mathcal{G}} (\zeta I - B)^{-1} C (\zeta I - A)^{-1} \, d\zeta, \tag{3.3}$$

where $\mathcal{G}$ is a simple closed curve enclosing $\sigma(A)$ and excluding $\sigma(B)$ (see Rosenblum [36] and [40]). This result can be generalized as follows.

Theorem 3.1: *The following statements (i) and (ii) are equivalent:*
(i) For every right–hand side $C \in \mathcal{L}^n(E_1, \ldots, E_n; E_0)$, the equation (3.1) has a unique solution $X \in \mathcal{L}^n(E_1, \ldots, E_n; E_0)$.
(ii) The linear maps A_ν $(\nu = 1, \ldots, n)$ and B fulfil the nonresonance condition of order n:

$$\left\{ \sum_{\nu=1}^{n} \lambda_\nu : \lambda_\nu \in \sigma(A_\nu) \right\} \cap \sigma(B) = \emptyset. \tag{3.4}$$

Under the nonresonance condition (3.4) the unique solution $X \in \mathcal{L}^n(E_1, \ldots, E_n; E_0)$ has the representation

$$\begin{aligned} X(h_1, \ldots, h_n) = \frac{1}{(2\pi i)^n} \int_{\mathcal{G}_n} \cdots \int_{\mathcal{G}_1} \left(\sum_{\nu=1}^{n} \zeta_\nu I - B \right)^{-1} \\ \cdot\, C\Big((\zeta_1 I - A_1)^{-1} h_1, \ldots, (\zeta_n I - A_n)^{-1} h_n \Big) \, d\zeta_1 \cdots d\zeta_n, \end{aligned} \tag{3.5}$$

where for $\nu = 1, \ldots, n$, $\mathcal{G}_\nu$ is a simple closed curve enclosing $\sigma(A_\nu)$ and excluding

$$\sigma\Big(B - \sum_{\substack{\mu=1 \\ \mu \neq \nu}}^{n} \zeta_\mu I \Big)$$

for all $\zeta_\mu \in \mathcal{G}_\mu$ $(\mu = 1, \ldots, n,\ \mu \neq \nu)$.

Remark: The statements of the theorem are true in the real and in the complex case. Moreover, in the special case $E_1 = E_\nu$ and $A_1 = A_\nu$ for $\nu = 2, \ldots, n$ it is easy to see that the symmetry of C implies the symmetry of X under the assumption (3.4). In this case it is also obvious that the gap condition $(\mathrm{GC})_r$ implies the nonresonance (3.4) for any $n = 1, \ldots, r$. Finally, for a different integral representation of the solution X under the assumption of such a spectral gap we refer to Iooss and Adelmeyer [21]; the precise relationship to our approach is worked out in [25].

Proof of the theorem: Let us first assume that (i) holds and that (ii) is false. Then, there exist $\lambda_\nu \in \sigma(A_\nu)$, $\nu = 1, \ldots, n$, such that

$$\beta := \sum_{\nu=1}^{n} \lambda_\nu \in \sigma(B). \tag{3.6}$$

Consider first the complex case $E_\nu \subset \mathbb{C}^m$. Then we have eigenvectors $x_\nu \in E_\nu$ and $y \in E_0$ satisfying $x_\nu^H A_\nu = \lambda_\nu x_\nu^H$ and $By = \beta y$. Define $X \in \mathcal{L}^n(E_1, \ldots, E_n; E_0) \setminus \{0\}$ by

$$X(h_1, \ldots, h_n) := y \cdot \prod_{\nu=1}^{n} \left(x_\nu^H \cdot h_\nu\right) \qquad \text{for} \quad h_\nu \in E_\nu \quad (\nu = 1, \ldots, n)$$

and find that X is a nontrivial solution of the homogeneous Sylvester equation:

$$\sum_{\nu=1}^{n} X(h_1, \ldots, h_{\nu-1}, A_\nu h_\nu, h_{\nu+1}, \ldots, h_n) = y \cdot \sum_{\nu=1}^{n} \Bigg(\lambda_\nu x_\nu^H h_\nu \cdot \prod_{\substack{\mu=1 \\ \mu \neq \nu}}^{n} \left(x_\mu^H \cdot h_\mu\right) \Bigg)$$

$$= \beta y \prod_{\mu=1}^{n} \left(x_\mu^H \cdot h_\mu\right) = BX(h_1, \ldots, h_n).$$

In the real case one proceeds in the standard way. First, consider the complexifications $E_{\nu,\mathbb{C}}$, $A_{\nu,\mathbb{C}}$, $B_\mathbb{C}$, $C_\mathbb{C}$ of E_ν, A_ν, B, C and construct as above the nontrivial solution $X_\mathbb{C} \in \mathcal{L}^n(E_{1,\mathbb{C}}, \ldots, E_{n,\mathbb{C}}; E_{0,\mathbb{C}})$ of the complexified equation. Then, both $\mathrm{Re}(X_\mathbb{C})$ and $\mathrm{Im}(X_\mathbb{C})$ restricted to the real spaces E_ν are in $\mathcal{L}^n(E_1, \ldots, E_n; E_0)$, they solve the real equation and they cannot vanish simultaneously.

It remains to show that under condition (ii) formula (3.5) yields a solution for any C so that uniqueness follows from linear algebra. First, notice that by the nonresonance condition we may choose suitable curves $\mathcal{G}_\nu \subset \mathbb{C}$ which satisfy the requirements of the theorem. For example take small circles enclosing the eigenvalues of A_ν. Then, inserting (3.5) into the multilinear Sylvester equation and splitting the terms $A_\nu h_\nu$ into $(A_\nu - \zeta_\nu I)h_\nu + \zeta_\nu h_\nu$ implies

$$\sum_{\nu=1}^{n} X(h_1, \ldots, h_{\nu-1}, A_\nu h_\nu, h_{\nu+1}, \ldots, h_n) - BX(h_1, \ldots, h_n) \;=$$

$$\frac{1}{(2\pi i)^n} \int_{\mathcal{G}_n} \cdots \int_{\mathcal{G}_1} \sum_{\nu=1}^{n} \left(\sum_{\mu=1}^{n} \zeta_\mu I - B \right)^{-1} C\Big((\zeta_1 I - A_1)^{-1} h_1, .., -h_\nu, .., (\zeta_n I - A_n)^{-1} h_n \Big)$$

$$+ C\Big((\zeta_1 I - A_1)^{-1} h_1, \ldots, (\zeta_n I - A_n)^{-1} h_n \Big) \; d\zeta_1 \cdots d\zeta_n.$$

The first integral terms on the right–hand side vanish as can be seen by evaluating the integral with respect to $\mathcal{G}_\nu$. The second term yields $C(h_1, \ldots, h_n)$ due to Cauchy's integral formula.

In the real case, the spectra are symmetric with respect to the real axis, and, by taking symmetric curves $\mathcal{G}_\nu$, our formula defines a real multilinear form X. $\square$

We briefly mention the analog of Theorem 3.1 for the mapping case. The multilinear Sylvester equation is now of the form (cf. (2.22))

$$X(A_1h_1,\ldots,A_nh_n) - B\cdot X(h_1,\ldots,h_n) = C(h_1,\ldots,h_n) \qquad \text{for} \quad h_\nu \in E_\nu \tag{3.7}$$

and the nonresonance condition reads

$$\left\{\prod_{\nu=1}^{n} \lambda_\nu : \lambda_\nu \in \sigma(A_\nu)\right\} \cap \sigma(B) = \emptyset. \tag{3.8}$$

Theorem 3.2: *Equation (3.7) has a unique solution X for any right–hand side C if and only if the nonresonance condition (3.8) holds. In that case X has the integral representation*

$$\begin{aligned} X(h_1,\ldots,h_n) = \frac{1}{(2\pi i)^n}\int_{\mathcal{G}_n}\cdots\int_{\mathcal{G}_1}&\left(\prod_{\nu=1}^{n}\zeta_\nu I - B\right)^{-1} \\ &\cdot C\Big((\zeta_1 I - A_1)^{-1}h_1,\ldots,(\zeta_n I - A_n)^{-1}h_n\Big)\, d\zeta_1\cdots d\zeta_n. \end{aligned} \tag{3.9}$$

Here each curve $\mathcal{G}_\nu$ does not contain 0, has $\sigma(A_\nu)$ in its interior and

$$\sigma\left(\left(\prod_{\substack{\mu=1\\ \mu\neq\nu}}^{n}\zeta_\mu\right)^{-1}\cdot B\right)$$

in its exterior for all $\zeta_\mu \in \mathcal{G}_\mu$ $(\mu \neq \nu)$.

3.2 The tensor approach

In this subsection we reconsider the multilinear Sylvester equation from the point of view of tensor analysis, and we indicate some generalization of our previous result. We use the fact that the pair $\big(\mathcal{L}^n(E_1,\ldots,E_n;E_0),\otimes\big)$ is a tensor product of the vector spaces $E_n^*,\ldots,E_1^*$, E_0 (see e.g. [15], [28]), if the multilinear function $\otimes : E_n^*\times\cdots\times E_1^*\times E_0 \to \mathcal{L}^n(E_1,\ldots,E_n;E_0)$, $(u_n,\ldots,u_0)\mapsto u_n\otimes\cdots\otimes u_0$, is defined by

$$(u_n\otimes\cdots\otimes u_0)(h_1,\ldots,h_n) := u_0\cdot\prod_{\nu=1}^{n} u_\nu h_\nu \qquad \text{for} \quad h_\nu\in E_\nu \quad (\nu = 1,\ldots,n).$$

Here E_ν^* denotes the dual space of the finite dimensional $\mathbb{K}$–vector space E_ν and $L_\nu^* \in \mathcal{L}(E_\nu^*,E_\nu^*)$ denotes the dual of a linear mapping $L_\nu \in \mathcal{L}(E_\nu,E_\nu)$ defined by $(L_\nu^* u)(h) = u(L_\nu h)$, $h\in E_\nu$, $u\in E_\nu^*$. The tensor product $L_n^*\otimes\cdots L_1^*\otimes L_0$ of the linear transformations

$L_\nu^* \in \mathcal{L}(E_\nu^*, E_\nu^*)$, $\nu = 1, \dots, n$, and $L_0 \in (E_0, E_0)$ is given by

$$\begin{aligned}
&\big((L_n^* \otimes \cdots L_1^* \otimes L_0)(u_n \otimes \cdots \otimes u_0)\big)(h_1, \dots, h_n) \\
&\quad = (L_n^* u_n \otimes \cdots \otimes L_1^* u_1 \otimes L_0 u_0)(h_1, \dots, h_n) = L_0 u_0 \cdot \prod_{\nu=1}^{n} (L_\nu^* u_\nu) h_\nu \qquad (3.10) \\
&\quad = L_0 u_0 \cdot \prod_{\nu=1}^{n} u_\nu (L_\nu h_\nu) = L_0 (u_n \otimes \cdots \otimes u_0)(L_1 h_1, \dots, L_n h_n)
\end{aligned}$$

for any $(u_n \otimes \cdots \otimes u_0) \in \mathcal{L}^n(E_1, \dots, E_n; E_0)$ and $(h_1, \dots, h_n) \in E_1 \times \cdots \times E_n$.

Let $L_0, \dots, L_n$ be maps as above and let p be a polynomial in the variables $x_0, \dots, x_n$ with coefficients $c_{k_0, \dots, k_n} \in \mathbb{K}$

$$p(x_0, \dots, x_n) = \sum_{k_0, \dots, k_n = 0}^{l} c_{k_0, \dots, k_n} \cdot \prod_{\nu=0}^{n} x_\nu^{k_\nu}. \qquad (3.11)$$

With these we associate linear operators in $\mathcal{L}\big(\mathcal{L}^n(E_1, .., E_n; E_0), \mathcal{L}^n(E_1, .., E_n; E_0)\big)$ through

$$p(L_0, L_1^* \dots, L_n^*) = \sum_{k_0, \dots, k_n = 0}^{l} c_{k_0, \dots, k_n} \cdot {L_n^*}^{k_n} \otimes \cdots \otimes {L_1^*}^{k_1} \otimes L_0^{k_0}$$

and

$$\big(T_{L_0, \dots, L_n}(X)\big)(h_1, \dots, h_n) := \sum_{k_0, \dots, k_n = 0}^{l} c_{k_0, \dots, k_n} L_0^{k_0} X\big(L_1^{k_1} h_1, \dots, L_n^{k_n} h_n\big), \qquad h_\nu \in E_\nu .$$

Using equation (3.10) and $(L_\nu^*)^{k_\nu} = (L_\nu^{k_\nu})^*$ implies

$$p(L_0, L_1^*, \dots, L_n^*) = T_{L_0, \dots, L_n} .$$

With this approach we can express the equations (3.1), (3.7) with the aid of tensor products as

$$p(B, A_1^*, \dots, A_n^*) \cdot X = C, \qquad (3.12)$$

where in case of (3.1) we have

$$p(x_0, \dots, x_n) = -x_0 + \sum_{\nu=1}^{n} x_\nu , \qquad (3.13)$$

and in case (3.7)

$$p(x_0, \dots, x_n) = -x_0 + \prod_{\nu=1}^{n} x_\nu . \qquad (3.14)$$

This reformulation of the problem is helpful in two different respects.

First, we may generalize the Theorem of Stéphanos ([27], [39]) as follows (the proof is by extending the proof of Theorem 2.1 in [28, chapter 2], cf. [25] for details).

Theorem 3.3: *Let $U_0, \ldots, U_n$ be finite dimensional $\mathbb{K}$–vector spaces where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $m_\nu := \dim U_\nu$. Let p be the polynomial from equation (3.11). If the maps $L_\nu \in \mathcal{L}(U_\nu, U_\nu)$, $\nu = 0, \ldots, n$, have eigenvalues $\mu_{\nu,1}, \ldots, \mu_{\nu,m_\nu}$ (repeated according to their algebraic multiplicity) then the $m_0 \cdots m_n$ numbers*

$$p(\mu_{0,r_0}, \ldots, \mu_{n,r_n}), \qquad r_\nu = 1, \ldots, m_\nu \quad (\nu = 0, \ldots, n),$$

are the eigenvalues of the linear map $p(L_0, \ldots, L_n)$ repeated according to their algebraic multiplicity.

Theorems 3.1 and 3.2 are then simple consequences of Theorem 3.3 when applied to (3.12), (3.13), (3.14).

The second aspect is the relationship between tensor products of linear maps and Kronecker products of matrices (see [28]):

For $\nu = 1, \ldots, n$ let $\{e_{\nu,1}, \ldots, e_{\nu,m_\nu}\}$ be an ordered basis of E_ν, and let $\{e^*_{\nu,1}, \ldots, e^*_{\nu,m_\nu}\}$ be the corresponding dual basis. If $M_\nu \in \mathbb{K}^{m_\nu, m_\nu}$ denotes the matrix representation of $L_\nu \in \mathcal{L}(E_\nu, E_\nu)$ then M_ν^T is the matrix representation of L_ν^*. Further, if one chooses the following basis of the space $\mathcal{L}^n(E_1, \ldots, E_n; E_0) = E_n^* \otimes \cdots \otimes E_1^* \otimes E_0$

$$\left\{ e^*_{n,r_n} \otimes \cdots \otimes e^*_{1,r_1} \otimes e_{0,r_0} \; : \; r_\nu \in \{1, \ldots, m_\nu\}, \; \nu = 0, \ldots, n \right\} \tag{3.15}$$

ordered lexicographically with priority to the first components of $(r_n, \ldots, r_0)$, then the matrix representation of the tensor product $L_n^* \otimes \cdots L_1^* \otimes L_0$ is given by the Kronecker product

$$M_n^T \otimes \cdots M_1^T \otimes M_0 \in \mathbb{K}^{\widetilde{m}, \widetilde{m}},$$

where $\widetilde{m} := \prod_{\nu=0}^n m_\nu$.

Similarly, for the polynomial p defined in (3.11) we use the Kronecker product and define

$$p(M_0, \ldots, M_n) = \sum_{k_0, \ldots, k_n = 0}^{l} c_{k_0, \ldots, k_n} \cdot M_n^{k_n} \otimes \cdots \otimes M_0^{k_0}.$$

Let $\mathrm{vec} : \mathcal{L}^n(E_1, \ldots, E_n; E_0) \to \mathbb{K}^{\widetilde{m}}$ be the isomorphism which assigns to each multilinear form X its coordinate vector with respect to the basis (3.15). A recursive definition of this map is given in [25]. It generalizes the well–known vec–function of a matrix as used for example in [33], [27]. With these assumptions and settings, and by identifying the linear maps A_ν and B with their matrix representations, equation (3.12) may be written as

$$p(B, A_1^T, \ldots, A_n^T) \cdot \mathrm{vec}(X) = \mathrm{vec}(C). \tag{3.16}$$

For equation (3.13) we use

$$p(B, A_1^T, \ldots, A_n^T) = \sum_{\nu=1}^{n} \left(I_{m_n \cdots m_{\nu+1}} \otimes A_\nu^T \otimes I_{m_{\nu-1} \cdots m_0} \right) - I_{m_n \cdots m_1} \otimes B,$$

and for (3.14)

$$p(B, A_1^T, \ldots, A_n^T) = A_n^T \otimes \cdots \otimes A_1^T \otimes I_{m_0} - I_{m_n \cdots m_1} \otimes B.$$

Since the matrix $p(B, A_1^T, \ldots, A_n^T)$ is of dimension $m_0 \cdots m_n$, this approach seems feasible for numerical computations only in case the number n and the dimensions m_ν are very small.

3.3 A multilinear Bartels–Stewart algorithm

The integral representation of the solution X in Theorem 3.1 is certainly of limited practical value. Its evaluation by a quadrature rule would require solving a multitude of linear systems. Furthermore, the tensor approach seems to be practicable for very small dimensions only. We therefore give an alternative proof of existence and uniqueness under the nonresonance condition (3.4). Our method of proof can be formulated as an algorithm. It is in fact a generalization of the Bartels–Stewart algorithm for solving the linear Sylvester equation (see [4], [11]).

From now on we work with bases in finite dimensional spaces, hence, we identify E_ν with $\mathbb{K}^{m_\nu}$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, for some m_ν and regard A_ν, B as matrices in $\mathbb{K}^{m_\nu, m_\nu}$ and $\mathbb{K}^{m_0, m_0}$, respectively.

In the first step we construct unitary matrices U_ν to form the complex Schur decompositions

$$U_\nu^H A_\nu U_\nu = A'_\nu = \begin{pmatrix} A'_{\nu,11} & \cdots & A'_{\nu,1m_\nu} \\ & \ddots & \vdots \\ \bigcirc & & A'_{\nu,m_\nu m_\nu} \end{pmatrix}. \tag{3.17}$$

Notice that in theory this form always exists and in practice is numerically stable ([11]). Furthermore, note that in the later application it requires the solution of a low dimensional eigenvalue problem ($m_\nu \ll m_0$).

Then, we perform a linear change of bases in the domain of X and C. Thus equation (3.1) is equivalent to

$$\begin{aligned} &\sum_{\nu=1}^{n} X(U_1 h_1, \ldots, U_{\nu-1} h_{\nu-1}, U_\nu A'_\nu h_\nu, U_{\nu+1} h_{\nu+1}, \ldots, U_n h_n) - BX(U_1 h_1, \ldots, U_n h_n) \\ &\quad = C(U_1 h_1, \ldots, U_n h_n) \qquad \text{for all} \quad h_\nu \in \mathbb{C}^{m_\nu} \;\; (\nu = 1, \ldots, n), \end{aligned}$$

or to

$$\sum_{\nu=1}^{n} X'(h_1, \ldots, h_{\nu-1}, A'_\nu h_\nu, h_{\nu+1}, \ldots, h_n) - BX'(h_1, \ldots, h_n) = C'(h_1, \ldots, h_n), \tag{3.18}$$

where we define

$$X'(h_1, \ldots, h_n) := X(U_1 h_1, \ldots, U_n h_n) \quad \text{and} \quad C'(h_1, \ldots, h_n) := C(U_1 h_1, \ldots, U_n h_n). \tag{3.19}$$

We consider multi–indices $(\alpha_1, \ldots, \alpha_n) = \alpha \in \mathbb{N}^n$, $\alpha_\nu \le m_\nu$, in a given strict order $\alpha^1 \prec \ldots \prec \alpha^M$, $M := m_1 \cdots m_n$, such that the following holds for every index α and every $\nu \in \{1, \ldots, n\}$:

$$(\alpha_1, \ldots, \alpha_{\nu-1}, k, \alpha_{\nu+1}, \ldots, \alpha_n) \prec \alpha \qquad \text{for all} \quad k \in \{1, \ldots, \alpha_\nu - 1\}. \tag{3.20}$$

For example, take a lexicographic order with arbitrarily chosen priority of the components. Let e_i denote the i–th coordinate vector. From equation (3.18) we successively compute

$$X'_{\alpha^j} := X'\big(e_{\alpha_1^j}, \ldots, e_{\alpha_n^j}\big), \quad j = 1, \ldots, M,$$

as follows:

$$
\begin{aligned}
C'_{\alpha^j} &= \sum_{\nu=1}^{n} X'(e_{\alpha_1^j},\ldots,A'_\nu e_{\alpha_\nu^j},\ldots,e_{\alpha_n^j}) - BX'_{\alpha^j} \\
&= \Big[\Big(\sum_{\nu=1}^{n} A'_{\nu,\alpha_\nu^j\alpha_\nu^j}\Big)I - B\Big]X'_{\alpha^j} + \sum_{\nu=1}^{n}\sum_{k=1}^{\alpha_\nu^j-1} A'_{\nu,k\alpha_\nu^j} X'_{\beta,\beta=(\alpha_1^j,\ldots,\alpha_{\nu-1}^j,k,\alpha_{\nu+1}^j,\ldots,\alpha_n^j)}\,.
\end{aligned}
\tag{3.21}
$$

The matrix $\Big[\Big(\sum_{\nu=1}^{n} A'_{\nu,\alpha_\nu^j\alpha_\nu^j}\Big)I - B\Big]$ is regular due to the nonresonance condition (3.4) and because of assumption (3.20) the terms X'_β in the second sum on the right–hand side satisfy $\beta \prec \alpha^j$ and hence are known at stage j. Therefore, X'_{α^j} can be computed and is uniquely determined by the above equation. In the complex case we set

$$
X(h_1,\ldots,h_n) := X'(U_1^H h_1,\ldots,U_n^H h_n) \qquad \text{for all} \quad h_\nu \in \mathbb{C}^{m_\nu} \ \ (\nu=1,\ldots,n)
$$

and find the solution of the original multilinear Sylvester equation (3.1). In the real case we work with the complexifications of A_ν. Since (3.1) has a unique solution, the above X provided by the algorithm is then a real multilinear form. As an alternative we can set up an algorithm based on the real Schur decomposition, see section 4.4. Equation (3.21) then becomes a low dimensional Sylvester equation for which the tensor approach is appropriate.

Further remarks in [4], [10] and [11] about the overall algorithm for the linear Sylvester equation apply to the multilinear case in an analogous fashion. For example, it may be appropriate to pretransform the matrix B before the linear systems (3.21) are solved.

As in section 3.1 we sketch here the analog of the above algorithm for the mapping case. Starting with (3.7) we form the Schur decompositions (3.17), define X' and C' as in (3.19) and determine the coefficients X'_{α^j} from the equations

$$
\begin{aligned}
C'_{\alpha^j} &= X'(A'_1 e_{\alpha_1^j},\ldots,A'_n e_{\alpha_n^j}) - BX'_{\alpha^j} \\
&= \Big[\Big(\prod_{\nu=1}^{n} A'_{\nu,\alpha_\nu^j\alpha_\nu^j}\Big)I - B\Big]X'_{\alpha^j} + \sum_{\substack{\beta\prec\alpha^j \\ \beta_\nu\le\alpha_\nu^j \\ (\nu=1,\ldots,n)}} \Big(\prod_{\nu=1}^{n} A'_{\nu,\beta_\nu\alpha_\nu^j}\Big)X'_\beta\,.
\end{aligned}
\tag{3.22}
$$

Each of these equations has a unique solution and the back transformation works as above.

4 Numerical Approximation

4.1 General approach

According to Theorem 3.1 we can determine the derivatives $D^\nu g^0 = D^\nu g(u_{0,l}) \in \mathcal{L}^\nu(E_l,E_h)$, $\nu=1,\ldots,n$, recursively from (2.14) if the nonresonance conditions $(\mathrm{NRC})_\nu$, $\nu=1,\ldots,n$, are satisfied.

For the practical implementation of the multilinear Bartels–Stewart algorithm from 3.3 we must choose bases $\varphi \in \mathbb{R}^{m,m_l}$ of E_l and $\Phi \in \mathbb{R}^{m,m_h}$ of E_h. We define

$$S := \big(\varphi, \Phi\big), \quad S^{-1} =: \begin{pmatrix} \psi^T \\ \Psi^T \end{pmatrix}, \quad \psi \in \mathbb{R}^{m,m_l}, \quad \Psi \in \mathbb{R}^{m,m_h}$$

and obtain the block diagonal form

$$S^{-1}AS = \begin{pmatrix} \psi^T A\varphi & \psi^T A\Phi \\ \Psi^T A\varphi & \Psi^T A\Phi \end{pmatrix} =: \begin{pmatrix} \widehat{A}_l & 0 \\ 0 & \widehat{A}_h \end{pmatrix}, \tag{4.1}$$

where $\widehat{A}_l$, $\widehat{A}_h$ are the basis representations of $A_l|_{E_l}$, $A_h|_{E_h}$. The projectors P_l, P_h are given by

$$P_l = \varphi\psi^T \quad \text{and} \quad P_h = \Phi\Psi^T = I - P_l,$$

and the function $\widehat{g}$ representing the invariant manifold in the new coordinate system assumes the form

$$\widehat{g}(x) = \Psi^T g(\varphi x), \quad x \in U_0 \subset \mathbb{R}^{m_l},$$

where U_0 is a suitable neighborhood of $x_0 := \psi^T u_0$.

With these transformations and the setting $D^n\widehat{g}^0 = D^n\widehat{g}(x_0)$ the multilinear Sylvester equation (2.14) may now be written as

$$\begin{aligned} \sum_{\nu=1}^{n} D^n\widehat{g}^0(h_1, \ldots, h_{\nu-1}, \widehat{A}_l h_\nu, h_{\nu+1}, \ldots, h_n) - \widehat{A}_h D^n\widehat{g}^0(h_1, \ldots, h_n) \\ = \widehat{C}_{h,n}(h_1, \ldots, h_n) \qquad \text{for all} \quad h_1, \ldots, h_n \in \mathbb{R}^{m_l}, \end{aligned} \tag{4.2}$$

with $\widehat{C}_{h,n}$ given by

$$\begin{aligned} \widehat{C}_{h,n}(h_1, \ldots, h_n) &= \sum_{\substack{\pi \in \mathcal{Q}^*(n) \\ (k:=|\pi|)}} \Psi^T D^k F^0\big(G_{|\pi_1|} h_{\pi_1}, \ldots, G_{|\pi_k|} h_{\pi_k}\big) \\ &- \sum_{\pi \in \mathcal{P}_n^*} \sum_{\substack{\tau \in \mathcal{Q}(\pi_2) \\ (k:=|\tau|)}} D^{|\pi_1|+1}\widehat{g}^0\big[h_{\pi_1}, \psi^T D^k F^0\big(G_{|\tau_1|} h_{\tau_1}, \ldots, G_{|\tau_k|} h_{\tau_k}\big)\big]. \end{aligned}$$

Here we have used the functions $G_\nu \in \mathcal{L}^\nu(\mathbb{R}^{m_l}, \mathbb{R}^m)$ defined by

$$G_1 = \varphi, \quad G_\nu = \Phi D^\nu \widehat{g}^0 \quad \text{for} \quad \nu \geq 2. \tag{4.3}$$

We can now apply the multilinear Bartels–Stewart algorithm to (4.2), but notice that this still involves the matrices Φ, Ψ, $\widehat{A}_h$ and hence is only suitable for small systems.

4.2 The modified algorithm for large systems

In large systems it is reasonable to assume that the matrices $\widehat{A}_l$, φ, ψ satisfying

$$A\varphi = \varphi\widehat{A}_l, \quad \psi^T A = \widehat{A}_l\psi^T, \quad \psi^T\varphi = I$$

are available, but that any explicit use of $\widehat{A}_h$, Φ, Ψ should be avoided. It is then useful to introduce the partially transformed mapping

$$X := D^n(g\circ\varphi)(x_0) = \Phi D^n\widehat{g}(x_0) \in \mathcal{L}^n(\mathbb{R}^{m_l},\mathbb{R}^m) \tag{4.4}$$

which equals G_n from (4.3) in case $n\geq 2$ and is the unique solution of the blown up system

$$\begin{aligned} \sum_{\nu=1}^{n} X(h_1,\ldots,h_{\nu-1},\widehat{A}_l h_\nu,h_{\nu+1},\ldots,h_n) - AX(h_1,\ldots,h_n) &= \Phi\widehat{C}_{h,n}(h_1,\ldots,h_n) \\ \psi^T X(h_1,\ldots,h_n) &= 0. \end{aligned} \tag{4.5}$$

Notice that the second equation enforces the range of X to be in E_h. Moreover, in the expression $\Phi\widehat{C}_{h,n}$ the terms involving $\Phi\Psi^T$ can be replaced by $I-\varphi\psi^T$ and the quantities $\Phi D^\nu\widehat{g}^0 = G_\nu,\ \ 2\leq\nu\leq n-1$, are known at this stage.

Equation (4.5) is an overdetermined system which by construction has a unique solution. In practice, however, rounding errors or inaccuracies in the matrices A_l', φ, ψ may destroy this property. Therefore, we propose to increase not only the number of equations but also the number unknowns and to solve an equivalent square system

$$\begin{aligned} \sum_{\nu=1}^{n} X(h_1,..,h_{\nu-1},\widehat{A}_l h_\nu,h_{\nu+1},..,h_n) + (\varphi Y - AX)(h_1,\ldots,h_n) &= \Phi\widehat{C}_{h,n}(h_1,\ldots,h_n) \\ \psi^T X(h_1,\ldots,h_n) &= 0. \end{aligned} \tag{4.6}$$

It is easy to see that this system is regular and that the unique solution is given by the above X and $Y=0$. Therefore it is possible to use $\|Y\|$ for an a–posteriori test of the accuracy.

As an alternative we can simply replace the leading terms $(I-\varphi\psi^T)D^kF^0$ in $\Phi\widehat{C}_{h,n}$ on the right–hand side of (4.6) by D^kF^0, which corresponds to the change from $C_{h,n}$ to C_n in Corollary 2.3. This does not affect X but changes the multilinear form Y to

$$Y(h_1,\ldots,h_n) := \sum_{\substack{\pi\in\mathcal{Q}^*(n)\\ (k:=|\pi|)}} \psi^T D^k F^0\big(G_{|\pi_1|}h_{\pi_1},\ldots,G_{|\pi_k|}h_{\pi_k}\big). \tag{4.7}$$

In case $n=1$ we have $Y=0$. For $n\geq 2$ using the chain rule (2.10) we find

$$Y = D^n\big(\psi^T F\circ(\varphi+g\circ\varphi)\big)(x_0). \tag{4.8}$$

Hence the multilinear form Y equals the term Y_n of Corollary 2.3 in our current coordinate system and the pair $(X,Y)\in\mathcal{L}^n(\mathbb{R}^{m_l},\mathbb{R}^m)\times\mathcal{L}^n(\mathbb{R}^{m_l},\mathbb{R}^{m_l})$ as in (4.4) and (4.7) is the unique solution of the system

$$\begin{aligned} \sum_{\nu=1}^{n} X(h_1,..,h_{\nu-1},\widehat{A}_l h_\nu,h_{\nu+1},..,h_n) + (\varphi Y - AX)(h_1,\ldots,h_n) &= \widehat{C}_n(h_1,\ldots,h_n) \\ \psi^T X(h_1,\ldots,h_n) &= 0, \end{aligned} \tag{4.9}$$

with $\widehat{C}_n$ given by

$$\begin{aligned}\widehat{C}_n(h_1,\ldots,h_n) &= \sum_{\substack{\pi\in\mathcal{Q}^*(n)\\(k:=|\pi|)}} D^kF^0\big(G_{|\pi_1|}h_{\pi_1},\ldots,G_{|\pi_k|}h_{\pi_k}\big)\\ &- \sum_{\pi\in\mathcal{P}_n^*}\sum_{\substack{\tau\in\mathcal{Q}(\pi_2)\\(k:=|\tau|)}} G_{|\pi_1|+1}\big[h_{\pi_1},\psi^TD^kF^0\big(G_{|\tau_1|}h_{\tau_1},\ldots,G_{|\tau_k|}h_{\tau_k}\big)\big].\end{aligned}$$

As in section 3.3 we proceed with the Schur decomposition of the small matrix $\widehat{A}_l$:

$$U^H\widehat{A}_lU = A'_l, \quad A'_l \quad \text{upper triangular.} \tag{4.10}$$

Then, introduce the transformed mappings

$$\begin{aligned}X'(h_1,\ldots,h_n) &:= X(Uh_1,\ldots,Uh_n),\\ Y'(h_1,\ldots,h_n) &:= Y(Uh_1,\ldots,Uh_n),\\ C'_n(h_1,\ldots,h_n) &:= \widehat{C}_n(Uh_1,\ldots,Uh_n).\end{aligned} \tag{4.11}$$

With these settings (4.9) is equivalent to

$$\begin{aligned}\sum_{\nu=1}^n X'(h_1,..,h_{\nu-1},A'_lh_\nu,h_{\nu+1},..,h_n)+(\varphi Y'-AX')(h_1,\ldots,h_n) &= C'_n(h_1,\ldots,h_n)\\ \psi^TX'(h_1,\ldots,h_n) &= 0.\end{aligned} \tag{4.12}$$

Ordering the multi–indices $(\alpha_1^j,\ldots,\alpha_n^j) = \alpha^j \in \{1,\ldots,m_l\}^n$, $j = 1,\ldots,M := (m_l)^n$, according to (3.20) we can successively compute

$$X'_{\alpha^j} := X'\big(e_{\alpha_1^j},\ldots,e_{\alpha_n^j}\big), \quad j=1,\ldots,M, \quad e_i = i\text{–th unit vector in } \mathbb{C}^{m_l},$$

from the linear systems

$$\begin{pmatrix}\Big[\Big(\sum\limits_{\nu=1}^n(A'_l)_{\alpha_\nu^j\alpha_\nu^j}\Big)I-A\Big] & \varphi\\ \psi^T & 0\end{pmatrix}\cdot\begin{pmatrix}X'_{\alpha^j}\\ Y'_{\alpha^j}\end{pmatrix} = \begin{pmatrix}R_{\alpha^j}\\ 0\end{pmatrix}. \tag{4.13}$$

with

$$R_{\alpha^j} = C'_{n,\alpha^j} - \sum_{\nu=1}^n\sum_{k=1}^{\alpha_\nu^j-1}(A'_l)_{k\alpha_\nu^j}X'_{\beta,\beta=(\alpha_1^j,\ldots,\alpha_{\nu-1}^j,k,\alpha_{\nu+1}^j,\ldots,\alpha_n^j)}. \tag{4.14}$$

The C'_n–terms in (4.14) can be computed with the help of $G_1=\varphi$ and the previous derivatives $G_\nu = \Phi D^\nu\widehat{g}^0$, $\nu=2,\ldots,n-1$, via

$$\begin{aligned}C'_n(h_1,..,h_n) &= \sum_{\substack{\pi\in\mathcal{Q}^*(n)\\(k:=|\pi|)}} D^kF^0\big(G_{|\pi_1|}Uh_{\pi_1},\ldots,G_{|\pi_k|}Uh_{\pi_k}\big)\\ &- \sum_{\pi\in\mathcal{P}_n^*}\sum_{\substack{\tau\in\mathcal{Q}(\pi_2)\\(k:=|\tau|)}} G_{|\pi_1|+1}\big[Uh_{\pi_1},\psi^TD^kF^0\big(G_{|\tau_1|}Uh_{\tau_1},..,G_{|\tau_k|}Uh_{\tau_k}\big)\big],\end{aligned} \tag{4.15}$$

where we use the notation $Uh_\omega = (Uh_{\omega_1},\ldots,Uh_{\omega_l})$, $\omega=(\omega_1,\ldots,\omega_l)$.

Finally, we transform back to obtain the n–th derivative in the original coordinates

$$X(h_1,\ldots,h_n) := X'(U^H h_1,\ldots,U^H h_n) \qquad \text{for} \quad h_1,\ldots,h_n \in \mathbb{R}^{m_l}. \tag{4.16}$$

From the unique solvability of equation (4.5) it follows that this X is indeed a real multilinear form. Our complete algorithm is now defined by the formulae (4.10) and (4.13) together with (4.14), (4.15) and (4.16).

Several remarks concerning the practical implementation follow:

First, as in (4.16) we obtain by

$$Y(h_1,\ldots,h_n) := Y'(U^H h_1,\ldots,U^H h_n) \qquad \text{for} \quad h_1,\ldots,h_n \in \mathbb{R}^{m_l}$$

a real multilinear form. According to equation (4.8) for $n \geq 2$ (the only case of interest) this multilinear form is the n-th derivative of the right–hand side of the reduced system in our current coordinate system. Thus our algorithm computes the unique solution of equation (2.15) and, as remarked at the end of section 2.2, we obtain as a by–product of the computation of $D^n(g \circ \varphi)$ the n-th derivative of the reduced system.

Second, due to the symmetry of the multilinear forms $C'_{h,n}$ and X' the number of coefficients X'_{α^j} to be computed reduces considerably: If the algorithm has reached some stage $j \in \{1,\ldots,M\}$ then the computation of X'_{α^j} may be skipped if α^j is a permutation of some $\beta \prec \alpha^j$. The number of systems (4.13) to be solved then reduces from

$$(m_l)^n \quad \text{to} \quad \binom{n+m_l-1}{m_l-1}.$$

We do not organize the index arrangement here but we show the savings for the bilinear case in section 4.4.

Finally, through the triangular Schur decomposition we have switched to a complex calculation even in case of a real multilinear Sylvester equation. One possibility which avoids complex numbers will be indicated for the bilinear case in section 4.4.

4.3 Solution of bordered systems

Linear systems with bordered matrices and large but sparse upper left block are quite typical in numerical continuation and bifurcation calculations (see [5], [8], [9], [13], [24], [30], [35]).

The upper left block is of the form $\lambda I - A$, $\lambda \in \mathbb{R}$ or $\lambda \in \mathbb{C}$, and it may be singular or almost singular for example if A has a zero eigenvalue in σ_l.

Various approaches for solving such systems have been developed and we propose to use the mixed block elimination method of Govaerts and Pryce (see [12], [14]).

In [12] linear systems of the type

$$\begin{array}{c} \\ m \\ m_l \end{array} \!\!\begin{array}{c} \begin{array}{cc} m & m_l \end{array} \\ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \end{array} \cdot \begin{array}{c} 1 \\ \begin{pmatrix} x \\ y \end{pmatrix} \end{array} = \begin{array}{c} 1 \\ \begin{pmatrix} f \\ g \end{pmatrix} \end{array}\!\! \begin{array}{c} \\ m \\ m_l \end{array} \quad \text{or} \quad \Gamma z = h \quad \text{for short} \tag{4.17}$$

are studied for the case $m_l = 1$. The assumption is that Γ is well–conditioned and that black box solvers S for A and S^* for A^T are given which are normwise backward stable (see [18]). This means there is a modest constant C_S, such that, for any RHS $r \in \mathbb{R}^m$, $S(r)$ is an approximation of the solution of $As = r$ satisfying

$$(A+E)S(r) = r + e$$

for some perturbation matrix E and vector e such that

$$\|E\| \leq \operatorname{eps} C_S \|A\| \quad \text{and} \quad \|e\| \leq \operatorname{eps} C_S \|r\|,$$

eps being the floating point roundoff unit ($\sim 2.2 * 10^{-16}$ for our example in section 5).

The method mixes the Block Crout and Doolittle factorization with one iterative refinement and it needs one call of S^* and two calls of S. It is shown in [12] that this yields a stable procedure BEM for solving (4.17) provided the condition number of A is at most of the order $\mathcal{O}\big(\frac{1}{\text{eps}}\big)$. This algorithm BEM is generalized in a recursive way to systems with wider borders and then is called BEMW in [14]. Typically, $1 \leq m_l \ll m$ holds and the method is stable (with a constant growing only linearly in m_l) if the leading principal sub–matrices of Γ of orders $m+1, \dots, m+m_l$ are not too ill–conditioned. Altogether, BEMW requires m_l solves with A^T and $m_l + 1$ solves with A and an additional number of flops of the order

$$4mm_l(2m_l+1) + \frac{8}{3}m_l^3$$

if real matrices are considered.

4.4 The bilinear case

We now exploit the details of the case $n = 2$ and show how the real Schur decomposition can be used to compute the bilinear form $X = D^2(g \circ \varphi)^0 = \Phi D^2 \widehat{g}^0 \in \mathcal{L}^2(\mathbb{R}^{m_l}, \mathbb{R}^m)$. Of course, this can be generalized to arbitrary n.

For $n = 2$ equation (4.9) reads

$$\begin{aligned} X(\widehat{A}_l h_1, h_2) + X(h_1, \widehat{A}_l h_2) - AX(h_1, h_2) + \varphi Y(h_1, h_2) &= \widehat{C}_2(h_1, h_2) \\ &= D^2 F^0\big(\varphi h_1, \varphi h_2\big) \\ \psi^T X(h_1, h_2) &= 0. \end{aligned} \tag{4.18}$$

According to (4.7), (4.8) the solution fulfils $Y(h_1, h_2) = \psi^T D^2 F^0\big(\varphi h_1, \varphi h_2\big)$.

For the small matrix $\widehat{A}_l$ we compute the real Schur decomposition

$$U^T \widehat{A}_l U = A'_l = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ & \ddots & \vdots \\ \bigcirc & & a_{pp} \end{pmatrix}. \tag{4.19}$$

Here $U \in \mathbb{R}^{m_l, m_l}$ is orthogonal and A'_l is upper quasi–triangular, where the blocks a_{ii} either denote a 1–by–1 matrix or a 2–by–2 matrix corresponding to a pair of complex conjugate eigenvalues. Defining X', Y' and C_2' as in (4.11) and (4.15) we may rewrite (4.18) as

$$\begin{aligned} X'(A_l'h_1,h_2)+X'(h_1,A_l'h_2)-AX'(h_1,h_2)+\varphi Y'(h_1,h_2) &= C_2'(h_1,h_2) \\ &= D^2F^0\big(\varphi Uh_1,\varphi Uh_2\big) \\ \psi^T X'(h_1,h_2) &= 0. \end{aligned} \tag{4.20}$$

Let the mapping $\tau:\{1,\ldots,p\}\to\{1,2\}$ define the sizes of the blocks a_{ik} via $a_{ik}\in\mathbb{R}^{\tau(i),\tau(k)}$. Then we partition the bilinear forms X' and C_2' into small bilinear forms by using the same pattern as in (4.19).

For example, in the case $\tau(1)=2$, $\tau(2)=1$ and $\tau(3)=2$ the bilinear form X' may be pictured as follows

$$X'=\big(X_{ik}'\big)=$$

$$\left[\begin{array}{cc|c|cc|cc}
X_{11}' & X_{12}' & X_{13}' & X_{14}' & X_{15}' & \cdots & X_{1m_l}' \\
X_{21}' & X_{22}' & X_{23}' & X_{24}' & X_{25}' & \cdots & X_{2m_l}' \\ \hline
\vdots & \vdots & X_{33}' & X_{34}' & X_{35}' & \cdots & X_{3m_l}' \\ \hline
\vdots & \vdots & \vdots & X_{44}' & X_{45}' & \cdots & X_{4m_l}' \\
\vdots & \vdots & \vdots & X_{54}' & X_{55}' & \cdots & X_{5m_l}' \\ \hline
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdots & X_{m_lm_l}'
\end{array}\right]
=
\left[\begin{array}{c|c|c|cc}
x_{11} & x_{12} & x_{13} & \cdots & x_{1p} \\ \hline
\vdots & x_{22} & x_{23} & \cdots & x_{2p} \\ \hline
\vdots & \vdots & x_{33} & \cdots & x_{3p} \\ \hline
\vdots & \vdots & \vdots & \ddots & \vdots \\
\cdot & \cdot & \cdot & \cdots & x_{pp}
\end{array}\right],$$

where $X_{ik}':=X'\big(e_i,e_k\big)$ and e_i is the i–th unit vector in $\mathbb{R}^{m_l}$. Elements in the left lower part which can be obtained by symmetry are replaced by dots.

The blocks $x_{\alpha^j}\in\mathcal{L}^2(\mathbb{R}^{\tau(\alpha_1^j)},\mathbb{R}^{\tau(\alpha_2^j)};\mathbb{R}^m)$, $\alpha^j=(\alpha_1^j,\alpha_2^j)\in\{1,\ldots,p\}^2$, are defined through

$$x_{\alpha^j}(\eta_1,\eta_2):=X'\big(E_{\alpha_1^j}\cdot\eta_1,E_{\alpha_2^j}\cdot\eta_2\big),\quad j=1,..,M:=p^2,\quad \eta_1\in\mathbb{R}^{\tau(\alpha_1^j)},\ \eta_2\in\mathbb{R}^{\tau(\alpha_2^j)},$$

where the i–th column of $E_k\in\mathbb{R}^{m_l,\tau(k)}$ is the $\big[\sum_{\mu=1}^{k-1}\tau(\mu)+i\big]$–th unit vector in $\mathbb{R}^{m_l}$ ($i\in\{1,\tau(k)\}$). Similarly, we construct y_{α^j} by partitioning Y' and c_{α^j} by partitioning C_2':

$$y_{\alpha^j}(\eta_1,\eta_2)=Y'\big(E_{\alpha_1^j}\cdot\eta_1,\,E_{\alpha_2^j}\cdot\eta_2\big)=\psi^TD^2F^0\big(\varphi UE_{\alpha_1^j}\cdot\eta_1,\,\varphi UE_{\alpha_2^j}\cdot\eta_2\big),$$

$$c_{\alpha^j}(\eta_1,\eta_2):=C_2'\big(E_{\alpha_1^j}\cdot\eta_1,E_{\alpha_2^j}\cdot\eta_2\big)=D^2F^0\big(\varphi UE_{\alpha_1^j}\cdot\eta_1,\varphi UE_{\alpha_2^j}\cdot\eta_2\big).$$

For $k_1,k_2\in\{1,\ldots,p\}$ we have $A_l'E_{k_1}=\sum_{i=1}^{k_1}E_ia_{ik_1}$ and hence

$$X'(A_l'E_{k_1}\eta_1,E_{k_2}\eta_2)=\sum_{i=1}^{k_1}X'(E_ia_{ik_1}\eta_1,E_{k_2}\eta_2)=\sum_{i=1}^{k_1}x_{(i,k_2)}(a_{ik_1}\eta_1,\eta_2)$$

as well as

$$X'(E_{k_1}\eta_1, A_l' E_{k_2}\eta_2) = \sum_{i=1}^{k_2} X'\left(E_{k_1}\eta_1, E_i a_{ik_2}\eta_2\right) = \sum_{i=1}^{k_2} x_{(k_1,i)}(\eta_1, a_{ik_2}\eta_2).$$

Summarizing, we have transformed (4.20) into a system of equations, $j = 1,\ldots,M$, given by

$$\begin{aligned} x_{\alpha^j}(a_{\alpha_1^j\alpha_1^j}\eta_1,\eta_2) + x_{\alpha^j}(\eta_1, a_{\alpha_2^j\alpha_2^j}\eta_2) - Ax_{\alpha^j}(\eta_1,\eta_2) + \varphi y_{\alpha^j}(\eta_1,\eta_2) & \\ = c_{\alpha^j}(\eta_1,\eta_2) - \sum_{i=1}^{\alpha_1^j-1} x_{(i,\alpha_2^j)}(a_{i\alpha_1^j}\eta_1,\eta_2) - \sum_{i=1}^{\alpha_2^j-1} x_{(\alpha_1^j,i)}(\eta_1, a_{i\alpha_2^j}\eta_2) & \qquad (4.21) \\ \psi^T x_{\alpha^j}(\eta_1,\eta_2) = 0. & \end{aligned}$$

If we now order the multi–indices α^j, $j = 1,\ldots,M$, according to (3.20), then as in the complex case in section 4.2 the terms x_β on the right–hand side of (4.21) satisfy $\beta \prec \alpha^j$ and the c_{α^j} can be computed from known quantities.

By the symmetry of X' we have $x_{(i,k)}(\eta_1,\eta_2) = x_{(k,i)}(\eta_2,\eta_1)$, and the symmetry of Y', C_2' yields the same property for the y_{α^j}, c_{α^j}. Therefore, we can omit the calculation of those pairs $(x_{\alpha^j}, y_{\alpha^j})$ for which α^j is a permutation of some $\beta \prec \alpha^j$. Moreover, the symmetry of the forms $x_{(i,i)}$, $y_{(i,i)}$ and $c_{(i,i)}$ on the block diagonal will be used in case 4 below.

We notice that the equation (4.21) has the same structure as the original bilinear Sylvester equation (4.18), but now η_i is in $\mathbb{R}$ or $\mathbb{R}^2$ and hence the total size of (4.21) can be estimated by $m \cdot \tau(\alpha_1^j) \cdot \tau(\alpha_2^j) \le 4m$.

For the calculation of x_{α^j} we transform (4.21) into an equivalent square system with a bordered matrix which will then be solved by BEMW.

In the following we consider the cases $(\tau(\alpha_1^j), \tau(\alpha_2^j)) = (1,1)$, $(1,2)$, $(2,1)$, $(2,2)$, where, in addition, we have to distinguish between $\alpha_1^j = \alpha_2^j$ and $\alpha_1^j \ne \alpha_2^j$ in the case $\tau(\alpha_1^j) = \tau(\alpha_2^j) = 2$. The cases are easily recognized in our example above. We abbreviate the right–hand side of (4.21) as r_{α^j} and put $e_1 := \binom{1}{0}$, $e_2 := \binom{0}{1}$.

Case 1: $\tau(\alpha_1^j) = \tau(\alpha_2^j) = 1$. Equation (4.21) is equivalent to the bordered system

$$\begin{pmatrix} \left[(a_{\alpha_1^j\alpha_1^j} + a_{\alpha_2^j\alpha_2^j})I - A\right] & \varphi \\ \psi^T & 0 \end{pmatrix} \cdot \begin{pmatrix} x_{\alpha^j} \\ y_{\alpha^j} \end{pmatrix} = \begin{pmatrix} r_{\alpha^j} \\ 0 \end{pmatrix}. \qquad (4.22)$$

The solution fulfils $y_{\alpha^j} = \psi^T D^2 F^0\left(\varphi U E_{\alpha_1^j}, \varphi U E_{\alpha_2^j}\right)$.

Case 2: $\tau(\alpha_1^j) = 2$, $\tau(\alpha_2^j) = 1$. Now we abbreviate $x_i := x_{\alpha^j}(e_i,1)$, $y_i := y_{\alpha^j}(e_i,1)$, $i = 1,2$, and extract $u = \left(\begin{smallmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{smallmatrix}\right) := a_{\alpha_1^j\alpha_1^j}$, $v := a_{\alpha_2^j\alpha_2^j}$ from (4.19). In (4.21) we successively insert $(e_1,1)$ and $(e_2,1)$ and find the following systems

$$\begin{pmatrix} \left[(u_{11}+v)I - A\right] & u_{21}I & \varphi \\ \psi^T & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ y_1 \end{pmatrix} = \begin{pmatrix} r_{\alpha^j}(e_1,1) \\ 0 \end{pmatrix} \qquad (4.23)$$

and

$$\begin{pmatrix} u_{12}I & \Big[(u_{22}+v)I-A\Big] & \varphi \\ 0 & \psi^T & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} r_{\alpha^j}(e_2,1) \\ 0 \end{pmatrix}. \tag{4.24}$$

This may be written as a regular square system of order $2m+2m_l$:

$$\left(\begin{array}{cc|cc} \Big[(u_{11}+v)I-A\Big] & u_{21}I & \varphi & 0 \\ u_{12}I & \Big[(u_{22}+v)I-A\Big] & 0 & \varphi \\ \hline \psi^T & 0 & 0 & 0 \\ 0 & \psi^T & 0 & 0 \end{array}\right) \cdot \begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} r_{\alpha^j}(e_1,1) \\ r_{\alpha^j}(e_2,1) \\ 0 \\ 0 \end{pmatrix} \tag{4.25}$$

Case 3: $\tau(\alpha_1^j)=1$, $\tau(\alpha_2^j)=2$. Analogous to case 2.

Case 4: $\tau(\alpha_1^j)=\tau(\alpha_2^j)=2$ and $\alpha_1^j=\alpha_2^j$. Define $x_{ik}:=x_{\alpha^j}(e_i,e_k)$, $y_{ik}:=y_{\alpha^j}(e_i,e_k)$, $i,k=1,2$, and $u=\left(\begin{smallmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{smallmatrix}\right):=a_{\alpha_1^j\alpha_1^j}=a_{\alpha_2^j\alpha_2^j}$. From the symmetry of X', Y' and $\alpha_1^j=\alpha_2^j$ we find $x_{12}=x_{21}$, $y_{12}=y_{21}$ and $r_{\alpha^j}(e_1,e_2)=r_{\alpha^j}(e_2,e_1)$. In (4.21) we successively insert $(\eta_1,\eta_2)=(e_i,e_k)$ with $(i,k)=(1,1)$, $(1,2)$, $(2,2)$ and end up with a bordered system of dimension $3(m+m_l)$ given by

$$\left[\begin{array}{ccc|ccc} \Big[2u_{11}I-A\Big] & 2u_{21}I & 0 & \varphi & 0 & 0 \\ u_{12}I & \Big[(u_{11}+u_{22})I-A\Big] & u_{21}I & 0 & \varphi & 0 \\ 0 & 2u_{12}I & \Big[2u_{22}I-A\Big] & 0 & 0 & \varphi \\ \hline \psi^T & 0 & 0 & 0 & 0 & 0 \\ 0 & \psi^T & 0 & 0 & 0 & 0 \\ 0 & 0 & \psi^T & 0 & 0 & 0 \end{array}\right] \cdot \begin{bmatrix} x_{11} \\ x_{12} \\ x_{22} \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} r_{\alpha^j}(e_1,e_1) \\ r_{\alpha^j}(e_1,e_2) \\ r_{\alpha^j}(e_2,e_2) \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.26}$$

Case 5: $\tau(\alpha_1^j)=\tau(\alpha_2^j)=2$ and $\alpha_1^j\neq\alpha_2^j$. We proceed as in case 4 but now the matrices $u:=a_{\alpha_1^j\alpha_1^j}$ and $v:=a_{\alpha_2^j\alpha_2^j}$ are different. In contrast to case 4 we no longer have $x_{12}=x_{21}$ and hence we find a regular bordered system of dimension $4(m+m_l)$ with the left upper block of the form

$$\begin{bmatrix} \Big[(u_{11}+v_{11})I-A\Big] & u_{21}I & v_{21}I & 0 \\ u_{12}I & \Big[(u_{22}+v_{11})I-A\Big] & 0 & v_{21}I \\ v_{12} & 0 & \Big[(u_{11}+v_{22})I-A\Big] & u_{21}I \\ 0 & v_{12}I & u_{12}I & \Big[(u_{22}+v_{22})I-A\Big] \end{bmatrix} \tag{4.27}$$

In all five cases the upper left block of the square bordered matrices can be singular or almost singular for example if A has a zero eigenvalue in σ_l.

Another observation is that, if complex eigenvalues occur, but a real computation is required then it is of course not sufficient to have a linear solver for real translates $A - \lambda I$, $\lambda \in \mathbb{R}$. Rather, certain bordered block matrices as in (4.25)–(4.27) appear, where A is on the diagonal and multiples of the identity form the off-diagonal elements. For a direct sparse linear solver (as we will use it in the example below) this structure is still acceptable. In this case one can also use the solver on (4.25)–(4.27) directly without recourse to BEMW.

After the computation of the coefficients x_{α^j}, $j = 1, \ldots, M$, the bilinear form X' can be constructed from these via

$$X'(e_{i_1}, e_{i_2}) = x_{(k_1,k_2)}(\epsilon_{1,j_1}, \epsilon_{2,j_2}) \quad \text{for } i_1, i_2 = 1, .., m_l, \quad e_i \text{ the } i\text{–th unit vector in } \mathbb{R}^{m_l},$$

where, for $\nu = 1, 2$, the integers $k_\nu \in \{1, \ldots, p\}$ as well as $j_\nu \in \{1, \tau(k_\nu)\}$ are uniquely determined by

$$i_\nu = \sum_{\mu=1}^{k_\nu - 1} \tau(\mu) + j_\nu$$

and ϵ_{ν,j_ν} is the j_ν–th unit vector in $\mathbb{R}^{\tau(k_\nu)}$. Now, analogous to (4.16), we transform back to obtain the n–th derivative in the original coordinates

$$X(h_1, \ldots, h_n) := X'(U^T h_1, \ldots, U^T h_n) \qquad \text{for} \quad h_1, \ldots, h_n \in \mathbb{R}^{m_l}. \tag{4.28}$$

Clearly, we can avoid this final transformation (4.28), if φ and ψ are chosen in such a way that $\widehat{A}_l$ in (4.1) already has the real Schur form (4.19).

5 Numerical Results

As a typical example we consider a spatial discretization of the well–known Brusselator ([34], [7], [26], [31])

$$\begin{aligned} u_t &= D_1 \Delta u + \lambda \left(A - (B+1)u + u^2 v\right) \qquad & t &\geq 0 \\ v_t &= D_2 \Delta v + \lambda \left(Bu - u^2 v\right) & x &\in X \end{aligned} \tag{5.1}$$

first in one then in two space dimensions ($X = [0,1]$ and $X = [0,1]^2$), in both cases subject to Dirichlet boundary conditions

$$u = A, \quad v = \frac{B}{A} \quad \text{at} \quad x \in \partial X. \tag{5.2}$$

A, B, D_1, D_2, λ are positive parameters and $\sqrt{\lambda}$ is the original length scale in the underlying trimolecular reactions. With the boundary conditions (5.2), equation (5.1) admits for all $\lambda > 0$ a single uniform steady–state solution:

$$u_0 = A, \quad v_0 = \frac{B}{A}. \tag{5.3}$$

In the one dimensional example at parameter values $A = 2$, $B = 6$, $D_1 = 2D_2$, $D_2 = 10^{-3}$ we discretize the second derivatives u_{xx}, v_{xx} by the standard three point difference method with $\Delta x = \frac{1}{200}$. After eliminating the boundary conditions, the resulting line system is of dimension 398. The eigenvalues of the Jacobian can still be computed explicitly and at

$$\lambda_0 = 240 * (1 - \cos(\pi/200)) \approx 0.0296082$$

a Hopf bifurcation occurs.

For the values $\lambda = 0.01$, $\lambda = \lambda_0$ and $\lambda = 0.08$ we compute the quadratic approximation of the invariant manifold which belongs to the leading complex conjugate pair of eigenvalues. It is a weakly stable manifold for $\lambda = 0.01$, a center manifold at $\lambda = \lambda_0$ and the unstable manifold for $\lambda_0 < \lambda < \overline{\lambda} = 0.11842$. Notice that this is covered by case 4 in section 4.4. The computations were done by combining Govaerts' solver [14] with a sparse linear solver from the package MATLAB, which applies a nonsymmetric minimum degree ordering algorithm. Fig. 2

In Figures 2–4 we show in the same scale the linear and quadratic approximation for these three parameter values, while Figure 5 contains all quadratic approximations in one picture. In each figure we plot the first component of both Fig. 3 Fig. 4 Fig. 5

$$\varphi s \quad \text{and} \quad \varphi s + X(s,s), \quad s \in [-1,1]^2,$$

where $\varphi = (\varphi_1, \varphi_2)$ spans the linear invariant subspace (with φ_1, φ_2 normalized with respect to the L_2–norm) and X is the computed bilinear form.

Table 1 shows the growth of the number of flops as given by MATLAB with increasing dimension for the case $\lambda = \lambda_0$. In each column we show the factor by which the number of flops increases from the previous column. For these computations we use the "tight" settings of the minimum degree ordering parameters as defined by the MATLAB function `spparms`, which may lead to orderings with less fill-in.

Table 1: One dimensional case: Growth of the number of flops

dim(line sys.) $= m$	48	98	198	398	798	1598	3198
dim(system (4.26))	150	300	600	1200	2400	4800	9600
factor	1	2.055	2.008	2.013	2.006	1.999	1.999

The left upper $3m \times 3m$ block of the bordered matrix in (4.26) has the bandwidth $2m + 1$ which grows linearly in m. From Table 1 it can be seen that in the current example the quadratic approximations can still be computed with an amount of flops which grows only linearly in m.

In the two–dimensional example we compute in an analogous way the quadratic approximations of the invariant manifolds belonging to the leading complex conjugate pair of eigenvalues using the same parameter values as above and discretizing the Laplace operator by the standard five point difference method. It turned out here, however, that the MATLAB solver for sparse matrices, which is used as the black box solver in Govaerts' algorithm BEMW, is efficient when combined with a reverse Cuthill-McKee ordering to reduce bandwidth. Table 2 shows the growth of the amount of flops and is arranged like Table 1.

Table 2: Two dimensional case: Growth of the number of flops

dim(line sys.) $= m$	72	162	338	648	1250	2450
dim(system (4.26))	222	492	1020	1950	3756	7356
factor	1	3.729	4.012	2.938	3.423	3.603

Since the Jacobian has an upper and lower bandwidth of $\sqrt{2m}$ in this example, the factors 3 up to 4 lie within the expected range. These factors certainly reduce further if some general iterative solver or a multigrid method is combined with BEMW.

References

[1] Abraham, R., Marsden, J. E., Ratiu, T. (1988): Manifolds, Tensor Analysis, and Applications. 2. ed., Appl. Math. Sci. **75**, Springer, Berlin, New York.

[2] Abraham, R., Robbin, J. (1967): Transversal Mappings and Flows. W. A. Benjamin, New York, Amsterdam.

[3] Aulbach, B., Garay, B. M. (1993): Linearizing the expanding part of noninvertible mappings. Z. Angew. Math. Phys. **44**, 469–494.

[4] Bartels, R. H., Stewart, G. W. (1972): Algorithm 432: Solution of the matrix equation $AX + XB = C$. Commun. ACM **15**, 820–826.

[5] Beyn, W.–J. (1991): Numerical methods for dynamical systems. In: W. Light, ed., Nonlinear Partial Differential Equations and Dynamical Systems. Advances in Numerical Analysis **1**, pp. 175–236. Clarendon, Oxford.

[6] Brunovský, P., Fiedler, B. (1986): Numbers of zeros on invariant manifolds in reaction–diffusion equations. Nonlinear Anal., Theory Methods Appl. **10**, 179–193.

[7] Carr, J. (1981): Applications of Centre Manifold Theory. Appl. Math. Sci. **35**, Springer, Berlin, New York.

[8] Chan, T. F. (1984a): Deflation techniques and block–elimination algorithms for solving bordered singular systems. SIAM J. Sci. Stat. Comput. **5**, 121–134.

[9] Chan, T. F. (1984b): Techniques for large sparse systems arising from continuation methods. In: T. Küpper, H. Mittelmann, H. Weber, eds., Numerical Methods for Bifurcation Problems. ISNM **70**, pp. 116-128. Birkhäuser, Basel, Boston.

[10] Golub, G. H., Nash, S., Van Loan, C. (1979): A Hessenberg–Schur method for the problem $AX + XB = C$. IEEE Trans. Autom. Control **24**, 909–913.

[11] Golub, G. H., Van Loan, C. F. (1989): Matrix Computations. 2. ed., The Johns Hopkins University Press, Baltimore, Maryland.

[12] Govaerts, W. (1991): Stable solvers and block elimination for bordered systems. SIAM J. Matrix Anal. Appl. **12**, 469–483.

[13] Govaerts, W. (1995): Bordered matrices and singularities of large nonlinear systems. Int. J. Bifurcation Chaos Appl. Sci. Eng. **5**, 243–250.

[14] Govaerts, W., Pryce, J. D. (1993): Mixed block elimination for linear systems with wider borders. IMA J. Numer. Anal. **13**, 161–180.

[15] Greub, W. (1978): Multilinear Algebra. 2. ed., Springer, Berlin, New York.

[16] Guckenheimer, J., Holmes, P. (1990): Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. 3. print., rev. and corr., Appl. Math. Sci. **42**, Springer, Berlin, New York.

[17] Hartman, P. (1982): Ordinary Differential Equations. Repr. of the 2. ed., Birkhäuser, Boston, Basel.

[18] Higham, N. J. (1996): Accuracy and Stability of Numerical Algorithms. SIAM, Philadelphia.

[19] Hilger, S. (1992): Smoothness of invariant manifolds. J. Funct. Anal. **106**, 95–129.

[20] Hirsch, M. W., Pugh, C. C., Shub, M. (1977): Invariant Manifolds. Lect. Notes Math. **583**, Springer, Berlin, New York.

[21] Iooss, G., Adelmeyer, M. (1992): Topics in Bifurcation Theory and Applications. Advanced Series in Nonlinear Dynamics **3**, World Scientific, Singapore, London.

[22] Irwin, M. C. (1972): On the smoothness of the composition map. Quart. J. Math., Oxford II. Ser. **23**, 113–133.

[23] Irwin, M. C. (1980): Smooth Dynamical Systems. Academic Press, London, New York.

[24] Keller, H. B. (1983): The bordering algorithm and path following near singular points of higher nullity. SIAM J. Sci. Stat. Comput. **4**, 573–582.

[25] Kleß, W. (1997): Numerische Approximation lokal invarianter Mannigfaltigkeiten in großen dynamischen Systemen. PhD thesis, University of Bielefeld.

[26] Kuznetsov, Y. A. (1995): Elements of Applied Bifurcation Theory. Appl. Math. Sci. **112**, Springer, Berlin, New York.

[27] Lancaster, P., Tismenetsky, M. (1985): The Theory of Matrices. 2. ed., Academic Press, Orlando, London, New York.

[28] Marcus, M. (1973): Finite Dimensional Multilinear Algebra, Part I. Pure and Appl. Math. **23**, Marcel Dekker, New York.

[29] Mielke, A. (1991): Hamiltonian and Lagrangian Flows on Center Manifolds with Applications to Elliptic Variational Problems. Lect. Notes Math. **1489**, Springer, Berlin, New York.

[30] Moore, G. (1987): Some remarks on the deflated block elimination method. In: T. Küpper, R. Seydel, H. Troger, eds., Bifurcation: Analysis, Algorithms, Applications. ISNM **79**, pp. 222–234. Birkhäuser, Basel, Boston.

[31] Murray, J. D. (1993): Mathematical Biology. 2., corr. ed., Biomathematics **19**, Springer, Berlin, New York.

[32] Nelson, E. (1969), Topics in Dynamics I: Flows. Mathematical Notes, Princeton University Press, Princeton, New Jersey.

[33] Neudecker, H. (1969): Some theorems on matrix differentiation with special reference to Kronecker matrix products. J. Am. Stat. Assoc. **64**, 953–963.

[34] Nicolis, G., Prigogine, I. (1977): Self–Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations. John Wiley & Sons, New York.

[35] Rheinboldt, W. C. (1986): Numerical Analysis of Parametrized Nonlinear Equations. Univ. Arkansas Lect. Notes Math. Sci. **7**, John Wiley & Sons, New York.

[36] Rosenblum, M. (1956): On the operator equation $BX - XA = Q$. Duke Math. J. **23**, 263–269.

[37] Sell, G. R. (1985): Smooth linearization near a fixed point. Am. J. Math. **107**, 1035–1091.

[38] Shub, M. (1987): Global Stability of Dynamical Systems. Translated by J. Christy, Springer, Berlin, New York. Original edition: Stabilité globale des systèmes dynamiques. Astérisque **56**, Société Mathématique de France (1978).

[39] Stéphanos, C. (1900): Sur une extension du calcul des substitutions linéaires. J. Math. Pures Appl., V. Sér. **6**, 73–128.

[40] Stewart, G. W., Sun, Ji–guang (1990): Matrix Perturbation Theory. Academic Press, San Diego, Boston, New York.

[41] Wiggins, S. (1990): Introduction to Applied Nonlinear Dynamical Systems and Chaos. Texts Appl. Math. **2**, Springer, Berlin, New York.

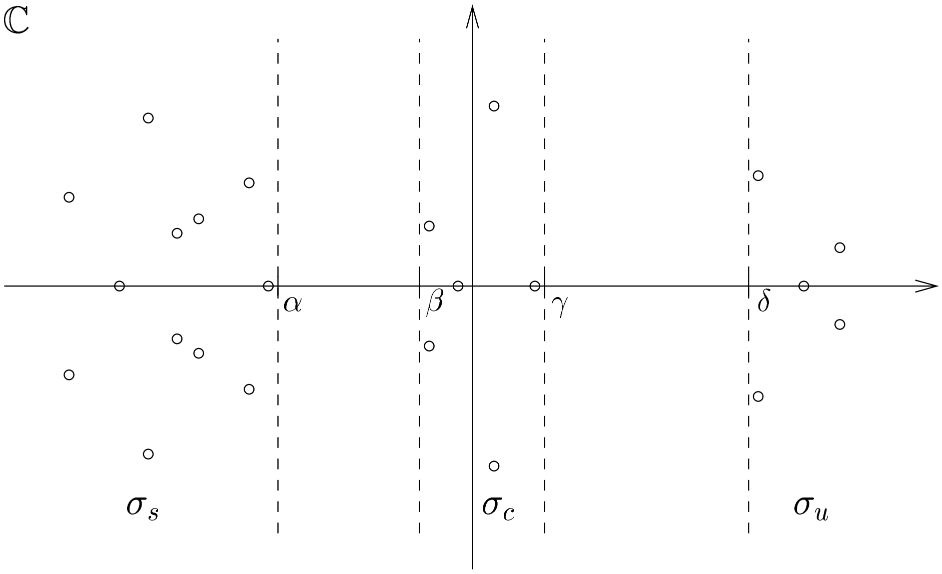

Figure 1: $\sigma(A) = \sigma_s \cup \sigma_c \cup \sigma_u$

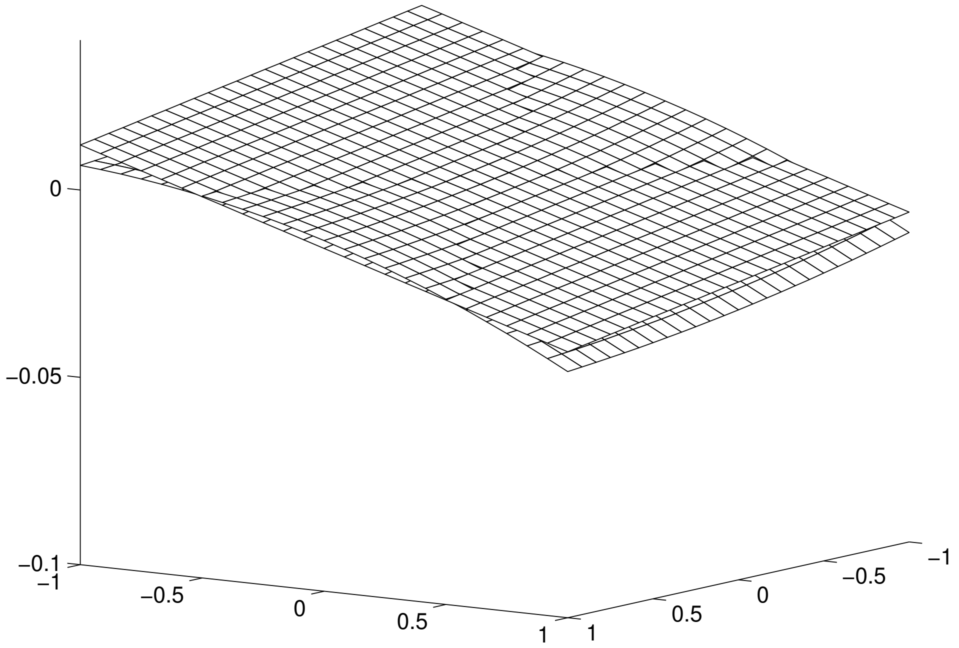

Figure 2: Linear and quadratic approximation at $\lambda = 0.01$

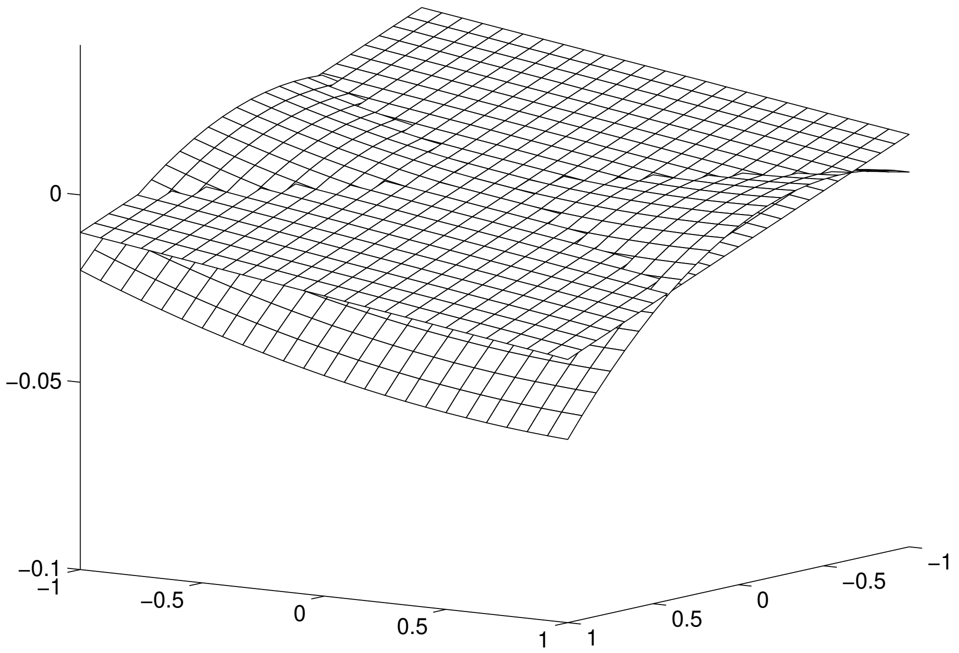

Figure 3: Linear and quadratic approximation at $\lambda = \lambda_0$ (Hopf–point)

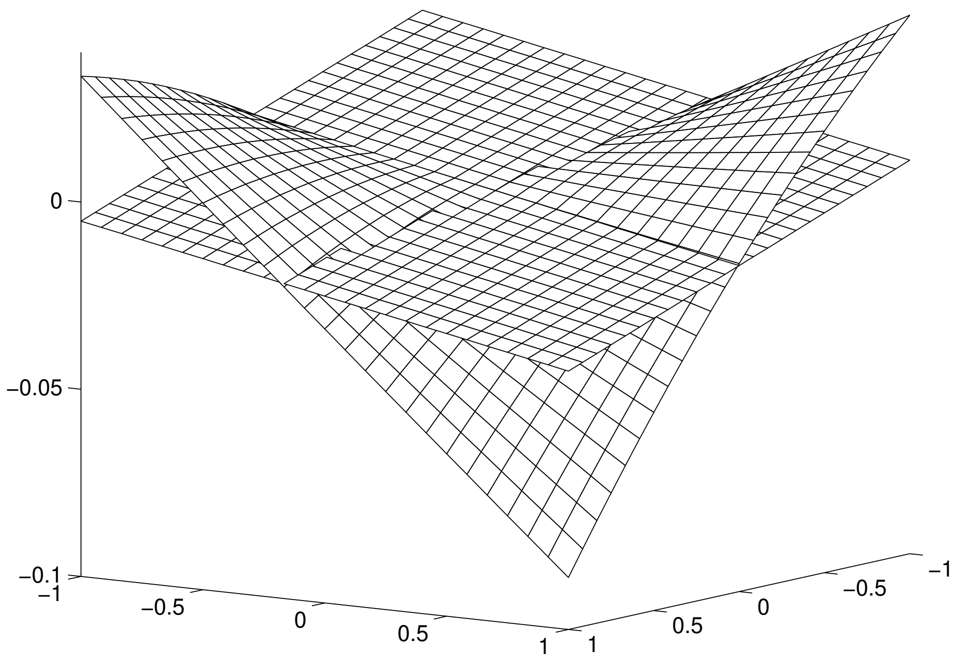

Figure 4: Linear and quadratic approximation at $\lambda = 0.08$

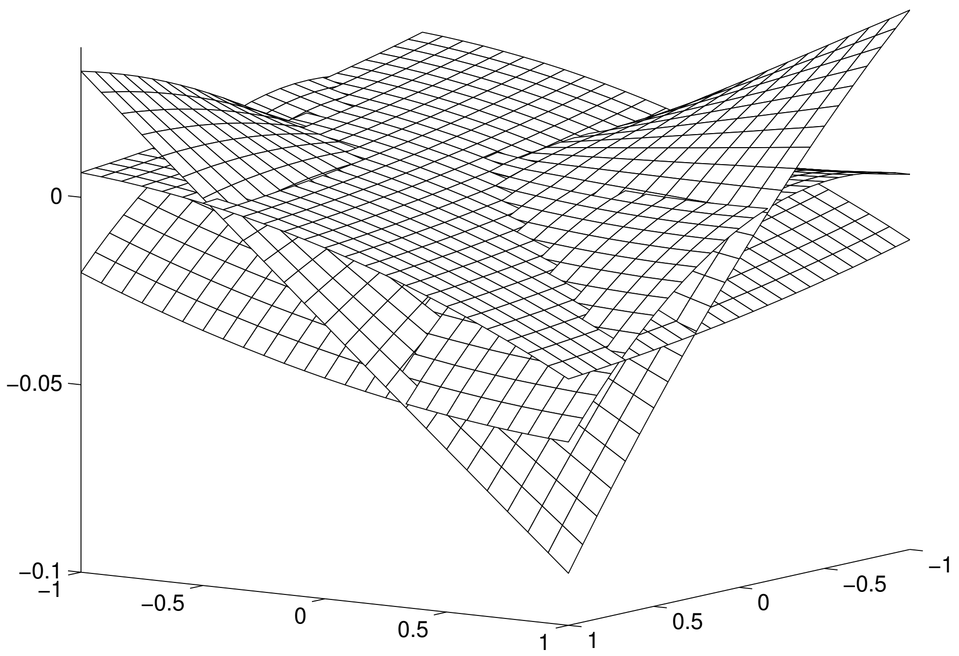

Figure 5: Quadratic approximations at $\lambda = 0.01$, λ_0, 0.08