

# Convergence analysis of a numerical method to solve the adjoint linearized periodic orbit equations

V. De Witte<sup>\*,a</sup>, W. Govaerts<sup>a</sup>

<sup>a</sup>*Department of Applied Mathematics and Computer Science, Ghent University, Krijgslaan 281-S9, B-9000 Gent, Belgium*

---

## Abstract

The equation adjoint to the linearization of the periodic orbit equations in a dynamical system is fundamental in the study of sensitivity issues for periodic orbits, e.g. in the synchronization of networks of weakly coupled oscillators. It is also fundamental in the computation of normal form coefficients for bifurcations of limit cycles. Numerically, the adjoint equations can be solved in a variety of ways. In the case where the periodic orbit equations are solved as a boundary value problem by collocation at Gauss points, a recent method allows one to compute the solution to the adjoint equations as a byproduct of Newton's method applied to solve the boundary value problem. This method is practically cost-free since it requires only the solution to an already factorized linear system. Moreover, it provides the solution to the adjoint equations in exactly the form needed in the applications.

So far, the method has not been analyzed carefully and no rigorous convergence results have been proved. We prove that the method is equivalent to a collocation method for the adjoint equations so that convergence of order  $h^{m+1}$  holds at all points and of order  $h^{2m}$  at the points of the coarse mesh; here  $h$  is the maximum length of the mesh intervals and  $m$  is the degree of the approximating piecewise polynomials. We support this by extensive numerical tests.

*Key words:* order of convergence, limit cycle, collocation, Gauss point

---

## 1. Introduction

Let a dynamical system be defined by

$$\dot{x}(t) = f(x(t), \alpha), \quad (1)$$

where  $x(t), f(x(t), \alpha) \in \mathbb{R}^n$  and  $\alpha$  is a vector of parameters. In the present paper,  $\alpha$  is fixed and will for simplicity be omitted from the equations.

When studying periodic solutions to (1) it is convenient to introduce the period  $T$  as an explicit unknown by rescaling time to the interval  $[0, 1]$ . Also, to obtain a unique solution it is necessary to fix the phase, e.g. by an integral condition. This leads to the following system:

---

<sup>\*</sup>Corresponding author

Email addresses: [Virginie.DeWitte@UGent.be](mailto:Virginie.DeWitte@UGent.be) (V. De Witte), [Willy.Govaerts@UGent.be](mailto:Willy.Govaerts@UGent.be) (W. Govaerts)

Preprint submitted to *Applied Numerical Mathematics*

May 18, 2010

$$\begin{cases} \dot{x}(t) - Tf(x(t)) = 0 \\ x(0) - x(1) = 0 \\ \int_0^1 \dot{\tilde{x}}(t)^T x(t) dt = 0, \end{cases} \quad (2)$$

where  $\tilde{x}(t)$  is an initial guess for the solution, typically obtained from a previous step in a continuation method. The integral equation minimizes the  $L_2$ -distance between  $x(t)$  and  $\tilde{x}(t)$  over phase shifts and so leads to a robust algorithm. This approach is by now standard in numerical bifurcation software, see [11, 9, 4, 7, 14].

In these packages an adaptive mesh  $\Delta$  and collocation at Gauss points are used to approximate  $x(t)$  by a piecewise polynomial  $x^\Delta(t)$  of degree  $m$  and  $T$  by a scalar  $T^\Delta$ . Collocation software for boundary value ODE's was first implemented in the package COLSYS [1, 2].

The adjoint solution  $v_l(t)$  to the linearization of (2) is the solution to the system

$$\begin{cases} \dot{v}_l(t) + TA(x(t))^T v_l(t) = 0 \\ v_l(0) - v_l(1) = 0 \\ \int_0^1 v_l(t)^T v(t) dt = 1, \end{cases} \quad (3)$$

where  $A = f_x$  and  $v(t) = \dot{x}(t) = Tf(x(t))$ . We note that this implies that  $v_l^T(t)v(t) = 1$  for all  $t$  and that the last, scalar equation in (3) merely scales  $v_l(t)$ . The adjoint solution was first used as a mathematical tool in the study of periodic orbits [20, 21], and references therein. In the applications of dynamical systems theory it quantifies the (linearized) effect of an input pulse to a periodic orbit. In the neural computation community the solution to the adjoint equation is used to compute the so-called phase resetting or phase response curves, which are important for, e.g., the study of synchronization of weakly coupled oscillators [3, 13, 16, 17]. The standard technique for the numerical computation of the solution to (3) is discussed in [23].

In [15] a new way was introduced to compute the solution to the adjoint equations as a byproduct of the computation of the periodic orbit. The same method was used in [18] in the computation of periodic normal form coefficients for codimension 1 bifurcations of cycles.

So far, no rigorous proof for this method has been published. The aim of this paper is to analyze the method, to prove that it is equivalent to a collocation method for the adjoint equations and hence to prove its convergence and obtain the order of convergence.

A word of caution is in order here. At first sight one might think that since the operator in the first group of equations in (3) is adjoint to the operator in the first group of equations in (2), their discretizations, or at least essential parts of them, will be transposed matrices.

This is false in general. It holds in very specific situations, e.g. when the mesh is uniform and the operator in (2) is discretized by forward Euler, while the operator in (3) is discretized by backward Euler. It fails already with the same discretizations if the mesh is non-uniform. In the case of collocation methods the relation between the two discretizations is obscure. It is not discussed in standard references such as [2] and [23]. We note also that our proof depends, surprisingly, on the use of the Gauss points as collocation points, see Proposition 2.

The paper is organized as follows. In sections 2.1-2.3 we discuss the discretization of vector-valued functions and recall the orders of convergence of the collocation approximations  $x^\Delta(t)$ ,  $v^\Delta(t)$  and  $T^\Delta$  to  $x(t)$ ,  $v(t)$  and  $T$ , respectively. In section 2.4, we rewrite (3) as a standard boundary value problem and consider a related problem in which  $x(t)$ ,  $v(t)$  and  $T$  are replaced by  $x^\Delta(t)$ ,  $v^\Delta(t)$  and  $T^\Delta$ , respectively.

We prove that the collocation approximation  $v_l^\Delta(t)$  to the latter problem converges to  $v_l(t)$  with order  $h^{m+1}$  at all points and with order  $h^{2m}$  at the points of the coarse mesh, where  $h$  is the maximal length of the mesh intervals.

In section 2.5, the adjoint system is rewritten so that  $v_l^\Delta(t)$  is obtained as a byproduct of the computation of the periodic orbit. Proposition 4 is the main result of the paper. In section 2.6, we discuss the practical consequences of the method, i.e. how the solution to the adjoint equation is used in the applications. In section 3, the obtained orders of convergence are supported through extensive testing, both in a model case where the exact solution to the adjoint system is analytically known and in the case of periodic orbits of the Lorenz model. Finally, in section 4, we draw some conclusions.

## 2. Discretization by collocation at Gauss points

### 2.1. Solution to a boundary value problem by collocation

We will deal with boundary value problems in which the unknown is a function  $Y(t) \in \mathbb{R}^n$ , defined in  $[0, 1]$  and satisfying

$$\begin{cases} \dot{Y}(t) &= F(Y(t)) \\ aY(0) + bY(1) &= 0, \end{cases} \quad (4)$$

where  $F$  is a sufficiently smooth function and  $a, b$  are constant matrices. The results of [5] apply to this situation (but we remark that [5] also applies to non-autonomous systems and more general boundary conditions).

To discretize (4) by a collocation method, the interval  $[0, 1]$  is first subdivided into  $N$  intervals with grid points:

$$0 = \tau_0 < \tau_1 < \dots < \tau_N = 1.$$

The points  $\tau_0, \tau_1, \dots, \tau_N$  form the coarse mesh  $\Delta$ . We define  $h = \|\Delta\| = \max_i h_i$  where  $h_i = \tau_{i+1} - \tau_i$ .  $Y(t)$  is approximated by a continuous function  $Y^\Delta(t)$  which in each interval  $[\tau_i, \tau_{i+1}]$  is a degree  $m$  polynomial, whose values are represented at equidistant mesh points, namely in

$$\tau_{i,j} = \tau_i + \frac{j}{m} h_i \quad (j = 0, 1, \dots, m).$$

We note that  $\tau_{i,m} = \tau_{i+1} = \tau_{i+1,0}$  for  $0 \leq i \leq N-1$ . These grid points form the fine mesh. In each interval  $[\tau_i, \tau_{i+1}]$  we require the polynomials to satisfy the differential equation in (4) exactly at  $m$  collocation points. The best choice for these collocation points are the Gauss points  $\zeta_{i,j}$ , i.e. the roots of the Legendre polynomial of degree  $m$ , relative to the interval  $[\tau_i, \tau_{i+1}]$  because of the high order of convergence of Gauss-Legendre numerical integration, based on collocation in the Gauss points [5, 6]. We also require the polynomials to satisfy the boundary conditions in (4). Under generic regularity conditions for the system (4) De Boor and Swartz [5] proved that  $Y^\Delta(t)$  converges uniformly over  $[0, 1]$  to  $Y(t)$  with order  $h^{m+1}$  and with order  $h^{2m}$  at the points of the coarse mesh. The regularity conditions are satisfied if both  $F$  and  $Y$  are  $C^{2m+1}$  functions.

### 2.2. Discretizations, weight forms and conversion

For a given vector function  $\eta \in C^1([0, 1], \mathbb{R}^n)$ , we consider two different discretizations [15]:

- $\eta_M \in \mathbb{R}^{(Nm+1)n}$ , the vector of the function values at the fine mesh points,

- $\eta_C \in \mathbb{R}^{Nmn}$ , the vector of the function values at the Gauss points.

We further introduce the structured sparse matrix  $L_{C \times M}$  that converts the vector  $\eta_M$  of function values of a degree  $m$  continuous piecewise polynomial at the mesh points into the vector  $\eta_C$  of its values at the Gauss points:

$$\eta_C = L_{C \times M} \eta_M.$$

To this end we first define  $L \in \mathbb{R}^{mn} \times \mathbb{R}^{(m+1)n}$  by

$$L = \begin{bmatrix} L_{1,1} & L_{1,2} & \cdots & L_{1,m+1} \\ L_{2,1} & L_{2,2} & \cdots & L_{2,m+1} \\ \vdots & \vdots & & \vdots \\ L_{m,1} & L_{m,2} & \cdots & L_{m,m+1} \end{bmatrix}$$

where  $L_{i,j} = l_{j-1}(\zeta_i) I_n$  ( $i = 1, \dots, m, j = 1, \dots, m+1$ ),

$$l_s(z) = \frac{1}{n_s} \prod_{r \neq s, r=0, \dots, m} \left( z - \frac{r}{m} \right)$$

and  $n_s = \prod_{r \neq s; r=0, \dots, m} \left( \frac{s}{m} - \frac{r}{m} \right)$  ( $s = 0, \dots, m$ ).  $I_n$  is the  $n \times n$  identity matrix. We note that  $l_s$  is the Lagrange interpolation polynomial of degree  $m$  which is equal to 1 at  $\frac{s}{m}$  and vanishes at  $\frac{r}{m}$  if  $r \in \{0, \dots, m\}, r \neq s$ ;  $\zeta_i$  is the  $i$ -th Gauss point relative to the interval  $[0, 1]$ . Then  $L_{C \times M} \in \mathbb{R}^{Nmn} \times \mathbb{R}^{(Nm+1)n}$  has the following form:

$$L_{C \times M} = \begin{bmatrix} \boxed{L} & & & & \\ & \boxed{L} & & & \\ & & \boxed{L} & & \\ & & & \ddots & \\ & & & & \boxed{L} \end{bmatrix},$$

where consecutive blocks overlap by  $n$  columns.

We also need the matrix  $D_{C \times M}$  that converts the vector  $\eta_M$  of function values of a degree  $m$  continuous piecewise polynomial at the mesh points into the vector  $\eta'_C$  of its derivative values at

the Gauss points. We first define  $F \in \mathbb{R}^{mn} \times \mathbb{R}^{(m+1)n}$  by

$$F = \begin{bmatrix} F_{1,1} & F_{1,2} & \cdots & F_{1,m+1} \\ F_{2,1} & F_{2,2} & \cdots & F_{2,m+1} \\ \vdots & \vdots & & \vdots \\ F_{m,1} & F_{m,2} & \cdots & F_{m,m+1} \end{bmatrix}$$

where  $F_{i,j} = l'_{j-1}(\xi_i)I_n$  ( $i = 1, \dots, m, j = 1 \dots m+1$ ) and

$$l'_s(z) = \frac{1}{n_s} \sum_{l \neq s, l=0, \dots, m} \prod_{r \neq s, l; r=0, \dots, m} \left( z - \frac{r}{m} \right)$$

( $s = 0, \dots, m$ ). Using this definition, we obtain the following  $(Nmn, (Nm+1)n)$ -matrix  $D_{C \times M}$

$$D_{C \times M} = \begin{bmatrix} \boxed{\frac{F}{h_0}} & & & & \\ & \boxed{\frac{F}{h_1}} & & & \\ & & \boxed{\frac{F}{h_2}} & & \\ & & & \ddots & \\ & & & & \ddots & \\ & & & & & \boxed{\frac{F}{h_{N-1}}} \end{bmatrix},$$

where, again, consecutive blocks overlap by  $n$  columns.

We further consider the weight form  $\eta_{WG} = \begin{bmatrix} \eta_G \\ \eta(0) \end{bmatrix} \in \mathbb{R}^{(Nm+1)n}$ , where  $\eta_G$  is the vector of the function values at the collocation points multiplied by the Gauss-Legendre weights and the lengths of the corresponding mesh intervals. To explain the use of the weight form, we first consider a scalar function  $f \in C^0([0, 1], \mathbb{R})$ . Then the integral  $\int_0^1 f(t)dt$  can be numerically approximated by appropriate linear combinations of function values. This can be done in several ways. (For background on quadrature methods, we refer to [6].)

If the Gauss points are used, then the best approximation has the form

$$\sum_{i=0}^{N-1} \sum_{j=1}^m \omega_{m,j}(f_C)_{i,j} h_i = \sum_{i=0}^{N-1} \sum_{j=1}^m (f_G)_{i,j}, \quad (5)$$

where  $(f_C)_{i,j} = f(\zeta_{i,j})$ , with  $\omega_{m,j}$  the Gauss-Legendre quadrature coefficients with respect to  $[0, 1]$ . Formula (5) delivers the exact integral if  $f(t)$  is a piecewise polynomial of degree  $2m - 1$  or less.

The integral  $\int_0^1 f(t)g(t)dt$  ( $f, g \in C^0([0, 1], \mathbb{R})$ ) is then approximated with Gauss-Legendre by  $f_G^T g_C = f_G^T L_{C \times M} g_M$ . For vector functions  $f, g \in C^0([0, 1], \mathbb{R}^n)$ , the integral  $\int_0^1 f(t)^T g(t)dt$  is formally approximated by the same expression:  $f_G^T g_C = f_G^T L_{C \times M} g_M$ .

We further introduce the matrix  $L_{G \times C} \in \mathbb{R}^{Nmn} \times \mathbb{R}^{Nmn}$  that converts a vector  $\eta_C$  into the vector  $\eta_G$ :

$$\eta_G = L_{G \times C} \eta_C. \quad (6)$$

To this end we define the diagonal matrix  $L_{G \times C}^{(i)} \in \mathbb{R}^{mn} \times \mathbb{R}^{mn}$  ( $i = 0, \dots, N - 1$ ) as

$$L_{G \times C}^{(i)} = h_i \begin{bmatrix} \omega_{m,1} I_n & & & & \\ & \omega_{m,2} I_n & & & \\ & & \omega_{m,3} I_n & & \\ & & & \ddots & \\ & & & & \ddots \\ & & & & & \omega_{m,m} I_n \end{bmatrix}.$$

Then  $L_{G \times C}$  is given by

$$L_{G \times C} = \begin{bmatrix} \boxed{L_{G \times C}^{(0)}} & & & & \\ & \boxed{L_{G \times C}^{(1)}} & & & \\ & & \boxed{L_{G \times C}^{(2)}} & & \\ & & & \ddots & \\ & & & & \boxed{L_{G \times C}^{(N-1)}} \end{bmatrix}.$$

### 2.3. Basic convergence results

We first recall the approach in [2, 10, 11] to the computation of limit cycles by collocation. System (2) can be reformulated as a standard boundary value problem by introducing artificial

scalar variables  $T(t)$  and  $w(t)$  as follows [10]:

$$\begin{cases} \dot{x}(t) - T(t)f(x(t)) = 0 \\ \dot{T}(t) = 0 \\ \dot{w}(t) - \dot{\tilde{x}}(t)^T x(t) = 0, \end{cases} \quad (7)$$

with boundary conditions

$$\begin{cases} x(0) - x(1) = 0 \\ w(0) = 0 \\ w(1) = 0. \end{cases} \quad (8)$$

In each interval  $[\tau_i, \tau_{i+1}]$ ,  $x(t)$ ,  $w(t)$  and  $T(t)$  are approximated by degree  $m$  polynomials  $x^\Delta(t)$ ,  $w^\Delta(t)$ ,  $T^\Delta(t)$ . The results of De Boor and Swartz [5] now apply to (7)-(8) and so the continuous piecewise polynomials  $x^\Delta(t)$ ,  $w^\Delta(t)$  and  $T^\Delta(t)$  converge uniformly over  $[0, 1]$  to  $x(t)$ ,  $w(t)$  and  $T(t)$  with order  $h^{m+1}$  and with order  $h^{2m}$  at the points of the coarse mesh.

By the collocation requirements

$$\begin{cases} \dot{x}^\Delta(\zeta_{i,j}) - T^\Delta(\zeta_{i,j})f(x^\Delta(\zeta_{i,j})) = 0 \\ \dot{T}^\Delta(\zeta_{i,j}) = 0 \\ \dot{w}^\Delta(\zeta_{i,j}) - \dot{\tilde{x}}(\zeta_{i,j})^T x^\Delta(\zeta_{i,j}) = 0 \\ x^\Delta(0) - x^\Delta(1) = 0 \\ w^\Delta(0) = 0 \\ w^\Delta(1) = 0, \end{cases} \quad (9)$$

for  $i = 0, \dots, N-1$ ,  $j = 0, \dots, m$ . Clearly,  $T^\Delta(t)$  has to be a constant  $T^\Delta$  and

$$\begin{aligned} 0 &= w^\Delta(1) - w^\Delta(0) \\ &= \int_0^1 \dot{w}^\Delta(t) dt \\ &= \sum_{i=0}^{N-1} \sum_{j=1}^m \omega_{m,j} \dot{w}^\Delta(\zeta_{i,j}) h_i \\ &= \sum_{i=0}^{N-1} \sum_{j=1}^m \omega_{m,j} \dot{\tilde{x}}(\zeta_{i,j})^T x^\Delta(\zeta_{i,j}) h_i \\ &= \dot{\tilde{x}}_G^T L_{C \times M} x_M^\Delta. \end{aligned}$$

Therefore we can rewrite (9) as

$$\begin{cases} \dot{x}^\Delta(\zeta_{i,j}) - T^\Delta f(x^\Delta(\zeta_{i,j})) = 0 \\ x^\Delta(0) - x^\Delta(1) = 0 \\ \dot{\tilde{x}}_G^T L_{C \times M} x_M^\Delta = 0, \end{cases} \quad (10)$$

for  $i = 0, \dots, N-1$ ,  $j = 0, \dots, m$ . By [5]

$$|x(t) - x^\Delta(t)| = O(h^{m+1}) \quad (11)$$

for all  $t$  and  $|x(t) - x^\Delta(t)| = O(h^{2m})$  at the points of the coarse mesh. Hence

$$|A(x(t)) - A(x^\Delta(t))| = O(h^{m+1}) \quad (12)$$

for all  $t$ . Since the period is a constant, we must have that

$$|T - T^\Delta| = O(h^{2m}). \quad (13)$$

The velocity vector  $v(t) = T f(x(t))$  can be approximated by  $v^\Delta(t) := T^\Delta f(x^\Delta(t))$ ; from the bounds (11) and (13) it follows that

$$|v(t) - v^\Delta(t)| = O(h^{m+1}). \quad (14)$$

System (10) can be considered as a non-linear system in  $x_M^\Delta$  and  $T^\Delta$  if  $x^\Delta(\zeta_{i,j})$  and  $\dot{x}^\Delta(\zeta_{i,j})$  are written in terms of  $x_M^\Delta$  by the use of  $L_{C \times M}$  and  $D_{C \times M}$  respectively. To compute the solution, we use Newton's method where we solve matrix equations with

$$M_h = \begin{bmatrix} D_{C \times M} - T^\Delta A_C(x^\Delta) L_{C \times M} & -f_C^\Delta \\ I_n & O & -I_n & 0_{n \times 1} \\ \dot{x}_G^T L_{C \times M} & & & 0_{1 \times 1} \end{bmatrix}, \quad (15)$$

where  $f^\Delta = f \circ x^\Delta$ , so  $(f_C^\Delta)_{i,j} = f^\Delta(\zeta_{i,j}) = f(x^\Delta(\zeta_{i,j}))$ . Next,  $A_C(x^\Delta) \in \mathbb{R}^{Nmn} \times \mathbb{R}^{Nmn}$  is given by

$$A_C(x^\Delta) = \begin{bmatrix} A(x^\Delta(\zeta_{0,1})) & & & & \\ & A(x^\Delta(\zeta_{0,2})) & & & \\ & & A(x^\Delta(\zeta_{0,3})) & & \\ & & & \ddots & \\ & & & & \ddots \\ & & & & & A(x^\Delta(\zeta_{N-1,m})) \end{bmatrix}.$$

Finally, the matrix  $O$  in (15) is the zero  $(n, (Nm-1)n)$ -matrix. In Proposition 4 we will show that  $M_h$  can also be used to obtain an approximation of order  $h^{m+1}$  to  $v_l(t)$ .

#### 2.4. The adjoint system

To apply the convergence results of De Boor and Swartz [5], we reformulate (3) as a standard boundary value problem. Therefore we introduce two artificial scalar variables  $\lambda(t)$  and  $u(t)$  to obtain the following system:

$$\begin{cases} \dot{v}_l(t) + T A(x(t))^T v_l(t) - \lambda(t) \dot{x}(t) = 0 \\ \dot{\lambda}(t) = 0 \\ \dot{u}(t) - v(t)^T v_l(t) = 0, \end{cases} \quad (16)$$

with boundary conditions

$$\begin{cases} v_l(0) - v_l(1) = 0 \\ u(0) = 0 \\ u(1) = 1. \end{cases} \quad (17)$$

Generically, this standard boundary value problem has an isolated solution with  $\lambda(t) = 0$  and  $u(t) = t$  for all  $t$ , so that the  $v_l(t)$ -component is the solution to (3).

The functions  $x(t)$ ,  $v(t)$  and scalar  $T$  in (16)-(17) are not known a priori and so we can not apply the results of [5] directly. A natural idea is to replace  $x(t)$ ,  $v(t)$  and  $T$  in (16) by their approximations  $x^\Delta(t)$ ,  $v^\Delta(t) = T^\Delta f(x^\Delta(t))$  and  $T^\Delta$  from a discretization as in section 2.3. For this we prove the following proposition:



**Proposition 1.** Let  $v_l^\Delta(t)$ ,  $\lambda^\Delta(t)$  and  $u^\Delta(t)$  be the collocation approximations for the mesh  $\Delta$  and piecewise polynomials of degree  $m$  to the system:

$$\begin{cases} \dot{v}_l(t) + T^\Delta A(x^\Delta(t))^T v_l(t) - \lambda(t) \dot{x}(t) = 0 \\ \dot{\lambda}(t) = 0 \\ \dot{u}(t) - v^\Delta(t)^T v_l(t) = 0 \end{cases} \quad (18)$$

with boundary conditions

$$\begin{cases} v_l(0) - v_l(1) = 0 \\ u(0) = 0 \\ u(1) = 1. \end{cases} \quad (19)$$

Then  $v_l^\Delta(t)$  converges to  $v_l(t)$  at all points of  $[0, 1]$  with order  $h^{m+1}$  and with order  $h^{2m}$  at the coarse mesh points. Also,  $\lambda^\Delta(t)$  is a constant and converges to 0 with order  $h^{2m}$ .

*Proof.* Formally, we can interpret (7)-(16) as one big system with boundary conditions (8)-(17) and then the convergence results of [5] apply. Obviously, the equations decouple and the remaining equations for  $v_l$ ,  $\lambda$  and  $u$  are precisely (18) with boundary conditions (19).

Consequently,  $v_l^\Delta$  converges to  $v_l(t)$  uniformly with order  $h^{m+1}$ , and with order  $h^{2m}$  at the points of the coarse mesh. Furthermore,  $\lambda^\Delta(t)$  is obviously constant. By considering its value at the coarse mesh points, we infer that it converges to 0 with order  $h^{2m}$ .  $\square$

We now come to the numerical computation of  $v_l^\Delta(t)$  and  $\lambda^\Delta(t)$  ( $u^\Delta(t)$  is an auxiliary variable that can be eliminated and plays no further role). By the collocation requirements, the piecewise polynomials  $v_l^\Delta(t)$ ,  $\lambda^\Delta(t)$  and  $u^\Delta(t)$  satisfy the following equations:

$$\begin{cases} \dot{v}_l^\Delta(\zeta_{i,j}) + T^\Delta A(x^\Delta(\zeta_{i,j}))^T v_l^\Delta(\zeta_{i,j}) - \lambda^\Delta(\zeta_{i,j}) \dot{x}(\zeta_{i,j}) = 0 \\ \dot{\lambda}^\Delta(\zeta_{i,j}) = 0 \\ \dot{u}^\Delta(\zeta_{i,j}) - v^\Delta(\zeta_{i,j})^T v_l^\Delta(\zeta_{i,j}) = 0, \end{cases} \quad (20)$$

for  $i = 0, \dots, N-1$ ,  $j = 0, \dots, m$  and with boundary conditions

$$\begin{cases} v_l^\Delta(0) - v_l^\Delta(1) = 0 \\ u^\Delta(0) = 0 \\ u^\Delta(1) = 1. \end{cases} \quad (21)$$

As noted in Proposition 1  $\lambda^\Delta(t) \equiv \lambda^\Delta$ , with  $\lambda^\Delta$  a constant. Further,

$$\begin{aligned} 1 &= u^\Delta(1) - u^\Delta(0) \\ &= \int_0^1 \dot{u}^\Delta(t) dt \\ &= \sum_{i=0}^{N-1} \sum_{j=1}^m \omega_{m,j} \dot{u}^\Delta(\zeta_{i,j}) h_i \\ &= \sum_{i=0}^{N-1} \sum_{j=1}^m \omega_{m,j} v^\Delta(\zeta_{i,j})^T v_l^\Delta(\zeta_{i,j}) h_i \\ &= v_G^{\Delta,T} L_{C \times M} v_{l,M}^\Delta. \end{aligned}$$

Therefore (20)-(21) can be rewritten as

$$\begin{cases} \dot{v}_l^\Delta(\zeta_{i,j}) + T^\Delta A(x^\Delta(\zeta_{i,j}))^T v_l^\Delta(\zeta_{i,j}) - \lambda^\Delta \dot{x}(\zeta_{i,j}) = 0 \\ v_l^\Delta(0) - v_l^\Delta(1) = 0 \\ v_G^{\Delta,T} L_{C \times M} v_{l,M}^\Delta = 1 \end{cases}$$

for  $i = 0, \dots, N-1, j = 0, \dots, m$ . Hence

$$\begin{bmatrix} D_{C \times M} + T^\Delta A_C(x^\Delta)^T L_{C \times M} & -\dot{x}_C \\ I_n & O & -I_n & 0_{n \times 1} \\ f_G^{\Delta, T} L_{C \times M} & 0_{1 \times 1} \end{bmatrix} \begin{bmatrix} v_{l, M}^\Delta \\ \lambda^\Delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{T^\Delta} \end{bmatrix}. \quad (22)$$

### 2.5. The solution to the adjoint equation as a byproduct of Newton's method

In Proposition 4 we will prove that the solution to (22) can be obtained cheaply as a byproduct of Newton's method applied to solve (10). We start with two preliminary results.

**Proposition 2.** *Define*

$$B_1 = \begin{bmatrix} D_{C \times M} - T^\Delta A_C(x^\Delta) L_{C \times M} \\ I_n & O & -I_n \end{bmatrix}, \quad B_2 = \begin{bmatrix} D_{C \times M} + T^\Delta A_C(x^\Delta)^T L_{C \times M} \\ I_n & O & -I_n \end{bmatrix}.$$

Let  $u$  and  $w$  be two continuous piecewise polynomials of degree  $m$  or less. Then, we have

$$[w_G^T \ w(0)^T] B_1 u_M + [u_G^T \ u(1)^T] B_2 w_M = 0. \quad (23)$$

*Proof.* We find that

$$\begin{aligned} w_G^T D_{C \times M} u_M &= w_G^T \dot{u}_C \\ &= \int_0^1 w(t)^T \dot{u}(t) dt \\ &= w(t)^T u(t) \Big|_0^1 - \int_0^1 \dot{w}(t)^T u(t) dt \\ &= -u_G^T D_{C \times M} w_M + w(1)^T u(1) - w(0)^T u(0). \end{aligned} \quad (24)$$

Here we have used the fact that  $w(t)^T \dot{u}(t)$  and  $\dot{w}(t)^T u(t)$  are piecewise polynomials of degree  $2m-1$  or less and  $u, w \in C^1([0, 1])$ .

It is clear that

$$A(x^\Delta(\zeta_{i,j}))^T h_i \omega_{m,j} I_n = h_i \omega_{m,j} I_n A(x^\Delta(\zeta_{i,j}))^T.$$

Therefore, and because of the diagonality of the matrix  $L_{G \times C}$ ,

$$A_C(x^\Delta)^T L_{G \times C} = L_{G \times C} A_C(x^\Delta)^T = L_{G \times C}^T A_C(x^\Delta)^T.$$

Using the previous equality we find that

$$\begin{aligned} w_G^T T^\Delta A_C(x^\Delta) L_{C \times M} u_M &= (w_G^T T^\Delta A_C(x^\Delta) u_C)^T \\ &= u_C^T T^\Delta A_C(x^\Delta)^T L_{G \times C} w_C \\ &= u_C^T L_{G \times C}^T T^\Delta A_C(x^\Delta)^T w_C \\ &= u_G^T T^\Delta A_C(x^\Delta)^T L_{C \times M} w_M. \end{aligned} \quad (25)$$

From (24) and (25), we obtain that

$$\begin{aligned}
& [w_G^T \ w(0)^T] \begin{bmatrix} D_{C \times M} - T^\Delta A_C(x^\Delta) L_{C \times M} \\ I_n & O & -I_n \end{bmatrix} u_M \\
&= -[u_G^T D_{C \times M} w_M - u(1)^T w(1) + w(0)^T u(0) + u_G^T T^\Delta A_C(x^\Delta)^T L_{C \times M} w_M \\
&\quad - w(0)^T (u(0) - u(1))] \\
&= -[u_G^T \ u(1)^T] \begin{bmatrix} D_{C \times M} + T^\Delta A_C(x^\Delta)^T L_{C \times M} \\ I_n & O & -I_n \end{bmatrix} w_M,
\end{aligned}$$

which is equivalent to (23).  $\square$

**Proposition 3.** *Let  $u$  and  $w$  be two continuous piecewise polynomials of degree  $m$  or less and  $\xi$  and  $\alpha$  be two scalars, then*

$$\begin{aligned}
& [w_G^T \ w(0)^T \ \xi] \begin{bmatrix} D_{C \times M} - T^\Delta A_C(x^\Delta) L_{C \times M} & -\eta_C \\ I_n & O & -I_n & 0_{n \times 1} \\ \mu_G^T L_{C \times M} & 0_{1 \times 1} \end{bmatrix} \begin{bmatrix} u_M \\ \alpha \end{bmatrix} + \\
& [u_G^T \ u(1)^T \ \alpha] \begin{bmatrix} D_{C \times M} + T^\Delta A_C(x^\Delta)^T L_{C \times M} & -\mu_C \\ I_n & O & -I_n & 0_{n \times 1} \\ \eta_G^T L_{C \times M} & 0_{1 \times 1} \end{bmatrix} \begin{bmatrix} w_M \\ \xi \end{bmatrix} = 0,
\end{aligned}$$

where  $\mu$  and  $\eta$  are two arbitrary functions in  $\mathbb{R}^n$ .

*Proof.* By Proposition 2 and the fact that  $w_G^T \eta_C = \eta_G^T w_C$  and  $\mu_G^T L_{C \times M} u_M = u_G^T \mu_C$  we find

$$\begin{aligned}
& [w_G^T \ w(0)^T \ \xi] \begin{bmatrix} D_{C \times M} - T^\Delta A_C(x^\Delta) L_{C \times M} & -\eta_C \\ I_n & O & -I_n & 0_{n \times 1} \\ \mu_G^T L_{C \times M} & 0_{1 \times 1} \end{bmatrix} \begin{bmatrix} u_M \\ \alpha \end{bmatrix} \\
&= -[u_G^T \ u(1)^T] \begin{bmatrix} D_{C \times M} + T^\Delta A_C(x^\Delta)^T L_{C \times M} \\ I_n & O & -I_n \end{bmatrix} w_M - \alpha \eta_G^T L_{C \times M} w_M \\
&\quad + \xi u_G^T \mu_C \\
&= -[u_G^T \ u(1)^T \ \alpha] \begin{bmatrix} D_{C \times M} + T^\Delta A_C(x^\Delta)^T L_{C \times M} & -\mu_C \\ I_n & O & -I_n & 0_{n \times 1} \\ \eta_G^T L_{C \times M} & 0_{1 \times 1} \end{bmatrix} \begin{bmatrix} w_M \\ \xi \end{bmatrix}.
\end{aligned}$$

$\square$

Our main result is the following.

**Proposition 4.** *The degree  $m$  piecewise polynomial  $v_l^\Delta(t)$  determined through its values at the Gauss points and in 0 by*

$$\begin{bmatrix} (v_l^\Delta)_G^T & (v_l^\Delta(0))^T & \lambda^M \end{bmatrix} \begin{bmatrix} D_{C \times M} - T^\Delta A_C(x^\Delta) L_{C \times M} & -f_C^\Delta \\ I_n & O & -I_n & 0_{n \times 1} \\ \dot{x}_G^T L_{C \times M} & 0_{1 \times 1} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{T^\Delta} \end{bmatrix} \quad (26)$$

converges to  $v_l(t)$  uniformly with order  $h^{m+1}$  and with order  $h^{2m}$  at the points of the coarse mesh. Moreover,  $\lambda^\Delta$  converges to zero with order  $h^{2m}$ .

*Proof.* We define  $v_l^\Delta$  and  $\lambda^\Delta$  from (22) and apply Proposition 3 with  $\eta = f^\Delta, \mu = \dot{x}, w = v_l^\Delta$  and  $\xi = \lambda^\Delta$  to obtain

$$\begin{aligned} & \left[ (v_l^\Delta)_G^T (v_l^\Delta(0))^T \lambda^\Delta \right] \begin{bmatrix} D_{C \times M} - T^\Delta A_C(x^\Delta) L_{C \times M} & -f_C^\Delta \\ I_n & O & -I_n & 0_{n \times 1} \\ & \dot{x}_G^T L_{C \times M} & & 0_{1 \times 1} \end{bmatrix} \begin{bmatrix} u_M \\ \alpha \end{bmatrix} \\ &= - \begin{bmatrix} u_G^T & u(1)^T & \alpha \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \frac{1}{T^\Delta} \end{bmatrix} \\ &= -\frac{\alpha}{T^\Delta}, \end{aligned}$$

for all continuous piecewise polynomials  $u$  of degree  $m$  or less and all scalars  $\alpha$ . Thus (26) follows. The convergence results follow from Proposition 1.  $\square$

### 2.6. Discussion

We note that from (26) we get  $(v_l^\Delta)_G$  rather than  $(v_l^\Delta)_M$ . This is an advantage since in all known applications we precisely need  $(v_l^\Delta)_G$  to compute integrals of the form

$$I = \int_0^1 v_l(t)^T \zeta(t) dt,$$

where  $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_n)^T \in C^0([0, 1], \mathbb{R}^n)$ . If  $I$  is approximated by

$$I_1 = (v_l)_G^T \zeta_C$$

(which is the best we can do), and

$$I_2 = (v_l^\Delta)_G^T \zeta_C,$$

then

$$\begin{aligned} |I_1 - I_2| &= |(v_l - v_l^\Delta)_G^T \zeta_C| \\ &= |\zeta_G^T (v_l - v_l^\Delta)_C| \\ &\leq O(h^{m+1}) \|\zeta_G\|_1, \end{aligned}$$

where  $\|\zeta_G\|_1$  tends to  $\sum_{i=1}^n \int_0^1 |\zeta_i(t)| dt$  for  $h \rightarrow 0$ .

To obtain  $(v_l^\Delta)_M$  (to our knowledge not needed in any applications) we take into account that  $v_l^\Delta(0) = v_l^\Delta(1)$  and solve the following system:

$$\begin{bmatrix} L_{C \times M} \\ 0.5 I_n & 0_{n \times (Nm-1)n} & 0.5 I_n \end{bmatrix} (v_l^\Delta)_M = \begin{bmatrix} (v_l^\Delta)_C \\ v_l^\Delta(0) \end{bmatrix}, \quad (27)$$

where the matrix in the left-hand side is sparse, square and well-conditioned.

## 3. Test Results

The tests described in this section are done in the framework of MatCont [7, 8]. Before we give test results for the order of convergence, some computational aspects have to be mentioned.

First, the meshes that we use are, like in AUTO [11, 9] and CONTENT [4], adaptive and hence non-uniform. Next, to solve the system that defines the approximation to the limit cycle, namely (10), Newton's method is used with Jacobian matrix  $M_h$  in (15). The stopping criterion consists of two conditions. The norm of the left-hand side of (10) must be smaller than a threshold FunTol, and the last Newton correction has to be smaller than a threshold VarTol.

After convergence we use the Jacobian matrix obtained in the last Newton step, we solve (26) to get  $(v_l^\Delta)_{WG}$ . If desired, then from (6) we get  $(v_l^\Delta)_C$ , and from (27)  $(v_l^\Delta)_M$  follows.

We now discuss upper bounds for the tolerances that in spite of the roundoff error can be met in Newton's method. This will later allow us to decide which errors are caused by truncation. We use the notation and framework on stability and backward stability in [22] and assume that all computations are done in a backward stable way. Let  $x_0$  be the exact solution to  $F(x) = 0$  where  $F$  is the left-hand side of (10). In a step of Newton's method we start with an approximate solution  $\tilde{x}$  with computed value  $\tilde{F}(\tilde{x})$  and computed Jacobian matrix  $B = \tilde{F}_x(\tilde{x})$ . We compute the next approximation as  $\tilde{\tilde{x}} = \tilde{x} - \Delta x$  where  $\Delta x$  is the solution which is obtained by solving the system  $\tilde{F}_x(\tilde{x})\Delta x = \tilde{F}(\tilde{x})$ . Up to higher order terms

$$F(\tilde{\tilde{x}}) = F(\tilde{x}) - F_x(\tilde{x})\widetilde{\Delta x}.$$

In what follows,  $\epsilon$  denotes machine precision and we omit the higher order terms. Since the map  $\tilde{x} \rightarrow \tilde{F}(\tilde{x})$  is backward stable, we have

$$\tilde{F}(\tilde{x}) = F(\tilde{x} + u) = F(\tilde{x}) + F_x(\tilde{x})u = F(\tilde{x}) + w,$$

where  $\|u\| = O(\epsilon)\|\tilde{x}\|$  and  $\|w\| = \|F_x(\tilde{x})u\| \leq \|F_x\|\|u\|$ , so  $\|w\| = O(\epsilon)\|F_x\|\|\tilde{x}\|$ . Now, since the map  $\tilde{x} \rightarrow \tilde{F}_x(\tilde{x})$  is also backward stable,

$$B = \tilde{F}_x(\tilde{x}) = F_x(\tilde{x} + s) = F_x(\tilde{x}) + F_{xx}(\tilde{x})s = F_x(\tilde{x}) + t,$$

where  $\|s\| = O(\epsilon)\|\tilde{x}\|$  and  $\|t\| = \|F_{xx}(\tilde{x})s\| \leq \|F_{xx}\|\|s\|$ , so  $\|t\| = O(\epsilon)\|F_{xx}\|\|\tilde{x}\|$ . The system

$$\tilde{F}_x(\tilde{x})\Delta x = \tilde{F}(\tilde{x})$$

is solved by a backward stable algorithm, hence

$$\tilde{B}\widetilde{\Delta x} = \tilde{F}(\tilde{x}), \tag{28}$$

where  $\tilde{B} = B + \zeta = \tilde{F}_x(\tilde{x}) + \zeta$  with  $\|\zeta\| = O(\epsilon)\|\tilde{F}_x(\tilde{x})\|$ . Now  $\|\tilde{F}_x(\tilde{x})\| = \|F_x(\tilde{x}) + t\| \leq \|F_x\| + \|t\|$  and thus  $\|\zeta\| = O(\epsilon)\|F_x\| + O(\epsilon^2)\|F_{xx}\|\|\tilde{x}\|$ . The second term is small compared with the first one so we can state that  $\|\zeta\| = O(\epsilon)\|F_x\|$ .

From (28) we infer that  $(\tilde{F}_x(\tilde{x}) + \zeta)\widetilde{\Delta x} = (F_x(\tilde{x}) + t + \zeta)\widetilde{\Delta x} = \tilde{F}(\tilde{x}) = F(\tilde{x}) + w$  and thus

$$\begin{aligned} F(\tilde{\tilde{x}}) &= F(\tilde{x}) - F_x(\tilde{x})\widetilde{\Delta x} \\ &= F(\tilde{x}) + (t + \zeta)\widetilde{\Delta x} - (F(\tilde{x}) + w) \\ &= (t + \zeta)\widetilde{\Delta x} - w. \end{aligned}$$

Therefore we obtain that

$$\begin{aligned} \|F(\tilde{\tilde{x}})\| &\leq (\|t\| + \|\zeta\|)\|\widetilde{\Delta x}\| + \|w\| \\ &= O(\epsilon)(\|F_{xx}\|\|\tilde{x}\| + \|F_x\|)\|\widetilde{\Delta x}\| + O(\epsilon)\|F_x\|\|\tilde{x}\|. \end{aligned}$$

Because we assume that  $\|\widetilde{\Delta x}\|$  is small, the first term will be small in comparison with the second one. Now, in the limit  $\|\tilde{x}\| \approx \|x_0\|$ , so we can state that

$$\|F(\tilde{x})\| = O(\epsilon) \|F_x\| \|x_0\|, \quad (29)$$

and thus in the test results we expect to get FunTol smaller than  $O(\epsilon) \|F_x\| \|x_0\|$ .

Now we determine an upper bound for the value of VarTol that we expect to achieve. Using the approximation  $F(\tilde{x}) = F_x(x_0)(\tilde{x} - x_0)$ , we obtain that

$$\|x_0 - \tilde{x}\| = \|F_x(x_0)^{-1} F(\tilde{x})\| \leq \|F_x^{-1}\| \|F(\tilde{x})\|.$$

Taking (29) into account, we conclude that

$$\|x_0 - \tilde{x}\| = O(\epsilon) \|F_x\| \|F_x^{-1}\| \|x_0\|. \quad (30)$$

We hence expect that we can reduce VarTol to be at least  $\|\tilde{x} - \tilde{x}\| = O(\epsilon) \|F_x\| \|F_x^{-1}\| \|x_0\|$ .

We now study two examples numerically and illustrate numerically and graphically the convergence of  $v_l^\Delta(t)$  to  $v_l(t)$  with order  $h^{m+1}$  in the cases where roundoff error can be ignored. This requires some careful computations in which (29)-(30) are used to ensure that roundoff is negligible.

### 3.1. Example: canonical unfolding of the Hopf bifurcation

We consider the following example

$$\begin{cases} x' = \alpha x - y - x(x^2 + y^2) \\ y' = x + \alpha y - y(x^2 + y^2). \end{cases} \quad (31)$$

For  $\alpha \leq 0$  the equilibrium  $(x \ y)^T = (0 \ 0)^T$  is stable. At  $\alpha = 0$  there is a Hopf bifurcation and for  $\alpha = 1$  we get the following stable periodic orbit:

$$\begin{cases} x = \cos(\tau) \\ y = \sin(\tau), \end{cases}$$

with  $\tau \in [0, 2\pi]$ , which is unique up to a phase shift. After time-rescaling we obtain

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \end{pmatrix} \text{ and } v(t) = \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = 2\pi \begin{pmatrix} -\sin(2\pi t) \\ \cos(2\pi t) \end{pmatrix}.$$

We now prove that

$$v_l(t) = \frac{1}{2\pi} \begin{pmatrix} -\sin(2\pi t) \\ \cos(2\pi t) \end{pmatrix} \quad (32)$$

by checking that (3) holds. First, we note that

$$A(x(t)) = \begin{pmatrix} -2 \cos^2(2\pi t) & -1 - 2 \sin(2\pi t) \cos(2\pi t) \\ 1 - 2 \sin(2\pi t) \cos(2\pi t) & -2 \sin^2(2\pi t) \end{pmatrix}$$

so that

$$\begin{aligned} & \dot{v}_l(t) + TA(x(t))^T v_l(t) \\ &= \begin{pmatrix} -\cos(2\pi t) \\ -\sin(2\pi t) \end{pmatrix} + \begin{pmatrix} -2\cos^2(2\pi t) & -1 - 2\sin(2\pi t)\cos(2\pi t) \\ 1 - 2\sin(2\pi t)\cos(2\pi t) & -2\sin^2(2\pi t) \end{pmatrix} \begin{pmatrix} -\sin(2\pi t) \\ \cos(2\pi t) \end{pmatrix} \\ &= 0. \end{aligned}$$

The boundary and integral condition in (3) hold obviously.

We consider discretizations with 10 to 40 mesh intervals in steps of 5;  $m$  always equals 4. In Table 1 we list  $N$ ,  $h$ , the largest and smallest singular values of the Jacobian matrix  $M_h$  given in (15), and its spectral condition number. In Table 2 the reached values of VarTol and FunTol

$N$	$h$	$\sigma_1$	$\sigma_n$	$\frac{\sigma_1}{\sigma_n}$
10	0.115090	156.0946	0.4668	334.3794
15	0.079341	232.0605	0.4015	577.9155
20	0.059620	315.4630	0.3554	887.6030
25	0.049054	389.1364	0.3241	1200.8145
30	0.041290	468.1366	0.2986	1567.7098
35	0.035811	545.2071	0.2791	1953.4180
40	0.031344	628.6388	0.2621	2398.1686

Table 1: Number of mesh intervals, maximum length of mesh intervals, largest and smallest singular value of  $M_h$ , spectral condition number of  $M_h$ .

$N$	VarTol	FunTol	$\epsilon \cdot \frac{\sigma_1}{\sigma_n} \cdot \ x\ $	$\epsilon \cdot \sigma_1 \cdot \ x\ $
10	1.6664e-15	5.4934e-14	6.7019e-13	3.1286e-13
15	1.8915e-15	7.7278e-14	1.2927e-12	5.1907e-13
20	3.5076e-15	1.2384e-13	2.1722e-12	7.7204e-13
25	2.6654e-15	1.8080e-13	3.1715e-12	1.0278e-12
30	5.5659e-15	2.5885e-13	4.4235e-12	1.3209e-12
35	6.8751e-15	3.1728e-13	5.8432e-12	1.6309e-12
40	1.3415e-14	4.0415e-13	7.5585e-11	1.9813e-12

Table 2: Number of mesh intervals, threshold VarTol, threshold FunTol, theoretical upper bound for VarTol, theoretical upper bound for FunTol.

are compared with the theoretical upper bounds (29) and (30). We note that the extremely small values for the obtained VarTol do not imply that the computed solutions are accurate up to these values.

Now we examine the superconvergence of the period. Table 3 compares the computed values  $T^\Delta$  with the exact value of the period, i.e.  $2\pi$ . According to (13) we must have that  $|2\pi - T^\Delta| = O(h^8)$ . Therefore  $\log_{10}(|2\pi - T^\Delta|) \approx 8 \log_{10}(h) + C$ . For 25 mesh intervals we find that  $|2\pi - T^\Delta| \approx \epsilon \|F_x\| \|F_x^{-1}\| T^\Delta = 1.6753e-12$ , i.e. the error is of the order of the round-off error. Thus it is reasonable to only consider the points  $(\log_{10}(h), \log_{10}(|2\pi - T^\Delta|))$  for  $N \leq 25$ . The slope of the least squares linear fit through these points is given by 8.458 which confirms the superconvergence of the period, see Figure 1.

$N$	$T^\Delta$	$ 2\pi - T^\Delta $	$\frac{ 2\pi - T^\Delta }{h^8}$
10	6.283185314816388	7.6368e-9	0.2481
15	6.283185307461999	2.8241e-10	0.1799
20	6.283185307214563	3.4976e-11	0.2191
25	6.283185307184551	4.9649e-12	0.1481
30	6.283185307180794	1.2079e-12	0.1430
35	6.283185307179926	3.4017e-13	0.1258
40	6.283185307179727	1.4033e-13	0.1506

Table 3: Number of mesh intervals, computed period, absolute error for the period, absolute error for the period compared with the expected order of convergence.

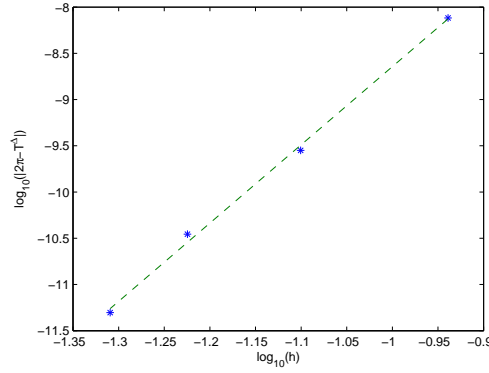


Figure 1: Least squares linear fit:  $\log_{10}(|2\pi - T^\Delta|) = 8.458 \log_{10}(h) - 0.186$ .

Now, we consider the order of convergence of the solution to the adjoint equations. By (32),  $\|v_l\|^2 = \frac{1}{4\pi^2}$ . Therefore, for every  $N$  we compute the square of the norm of the computed solution in the fine mesh points, compare this with  $\frac{1}{4\pi^2}$  and take the maximal difference over the fine mesh points, see Table 4. For  $N > 25$ , the maximal error doesn't decrease anymore. This can be understood as follows. For  $N > 25$  the truncation error at the points of the coarse mesh and in the period is of the order of the round-off error in the linear algebra and so the round-off error can no longer be ignored. We note that the uniform error bound used in Table 4 is a very strict bound; for the purpose of the applications the average error over the orbit is probably more relevant. We deal with this in the next example.

To visualize the order of convergence, we plot in Figure 2 on the horizontal axis  $\log_{10}(h)$  and on the vertical axis  $\log_{10}(\max(|\frac{1}{4\pi^2} - \|v_l^\Delta\|^2|))$  for  $N \leq 25$  and we consider the least squares linear fit through the corresponding points. The slope of the least squares linear fit is 6.308 which is even better than the expected slope of 5. We hypothesize that this is due to the fact that (31) is a very simple system with a rotational symmetry and a very smooth periodic orbit.



$N$	$\max\left(\left \frac{1}{4\pi^2} - \ v_l^\Delta\ ^2\right \right)$	$\frac{\max\left(\left \frac{1}{4\pi^2} - \ v_l^\Delta\ ^2\right \right)}{h^5}$
10	3.6072e-8	0.001786
15	3.2994e-9	0.001049
20	4.1821e-10	0.000555
25	1.9118e-10	0.000673
30	1.9227e-10	0.001602
35	1.9223e-10	0.003264
40	1.9257e-10	0.0063652

Table 4: Number of mesh intervals, error for the adjoint solution, error for the adjoint solution compared with the expected order of convergence.

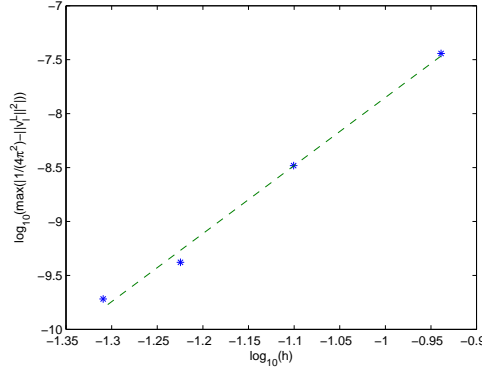


Figure 2: Least squares linear fit:  $\log_{10}(\max(|\frac{1}{4\pi^2} - \|v_l^\Delta\|^2|)) = 6.308 \log_{10}(h) - 1.543$ .

### 3.2. Example: the Lorenz84 model

As a second example we consider the following system:

$$\begin{cases} \dot{x} &= -y^2 - z^2 - ax + aF \\ \dot{y} &= xy - bxz - y + G \\ \dot{z} &= bxy + xz - z. \end{cases}$$

This Lorenz model [19] has a stable periodic orbit for the parameters values  $a = 0.25, b = 4, F = 4$  and  $G = 0.5$ . The solution to the adjoint problem is not known analytically. Therefore, we compute  $v_l^\Delta$  for 300 mesh intervals and consider this as the exact  $v_l$ . Again, we consider meshes  $\Delta$  with from 10 to 40 intervals in steps of 5 and piecewise polynomials of degree 4.

Table 5 lists  $N, h$ , the 2-norm of the Jacobian matrix (15), its smallest singular value and its spectral condition number. In Table 6 we give the obtained values of VarTol and FunTol and their theoretical upper bounds.

Next, the approximated periods are compared with the period found in the discretization by 300 mesh intervals, namely  $T^{300} = 1.544168236465894$ . As in the first example we expect that

$N$	$h$	$\sigma_1$	$\sigma_n$	$\frac{\sigma_1}{\sigma_n}$
10	0.1313	171.5856	0.1835	934.8931
15	0.0833	254.7963	0.1519	1676.3931
20	0.0719	345.8800	0.1326	2608.2081
25	0.0520	429.2663	0.1191	3604.5458
30	0.0518	519.0071	0.1090	4759.5223
35	0.0380	611.1010	0.1011	6042.3979
40	0.0370	694.9195	0.0948	7333.2397
300	0.0049	5240.0547	0.0349	150112.9250

Table 5: Number of mesh intervals, maximum length of mesh intervals, largest and smallest singular value of  $M_h$ , spectral condition number of  $M_h$ .

$N$	VarTol	FunTol	$\epsilon \cdot \frac{\sigma_1}{\sigma_n} \cdot \ x\ $	$\epsilon \cdot \sigma_1 \cdot \ x\ $
10	5.8140e-15	7.8167e-14	1.9870e-12	3.6468e-13
15	7.4450e-15	1.4774e-13	4.2013e-12	6.3856e-13
20	2.5826e-14	2.0101e-13	7.4010e-12	9.8147e-13
25	1.5716e-14	2.9385e-13	1.1293e-11	1.3449e-12
30	2.8568e-14	3.5105e-13	1.6202e-11	1.7668e-12
35	8.9663e-15	5.1049e-13	2.2073e-11	2.2324e-12
40	1.2868e-14	5.4380e-13	2.8508e-11	2.7015e-12
300	1.0192e-12	1.2044e-11	1.5532e-9	5.4219e-11

Table 6: Number of mesh intervals, threshold VarTol, threshold FunTol, theoretical upper bound for VarTol, theoretical upper bound for FunTol.

$$\log_{10}(|T^{300} - T^\Delta|) \approx 8 \log_{10}(h) + C$$

for a constant  $C$ , as long as the round-off error can be ignored. For  $N = 25$  we find that  $|T^{300} - T^\Delta| \approx \epsilon \|F_x\| \|F_x^{-1}\| T^\Delta = 1.6319e-12$ . Thus, when making a plot to visualize the order of convergence we only consider the points  $(\log_{10}(h), \log_{10}(|T^{300} - T^\Delta|))$  for  $N \leq 25$ . Figure 3 shows that the slope of the least squares linear fit is given by 8.279, which confirms the order of convergence  $h^{2m}$  of the period.

To study the order of convergence of  $v_l^\Delta$ , we use the following method. When we compare the exact periodic orbit  $x$  and an approximated solution  $x^\Delta$ , we have to take into account that the phase shifts, which depend on various details of the way in which the orbits are computed, are not necessarily the same. Therefore, we compute a shift  $\tau$  such that the following integral is minimized:

$$\int_0^1 \|x(t) - x^\Delta(t + \tau)\|^2 dt.$$

This is equivalent to

$$\begin{aligned} \frac{d}{d\tau} \langle x(t) - x^\Delta(t + \tau), x(t) - x^\Delta(t + \tau) \rangle &= 0 \quad \forall t \in [0, 1] \\ \Leftrightarrow 2 \left\langle x(t) - x^\Delta(t + \tau), -\frac{d}{d\tau} x^\Delta(t + \tau) \right\rangle &= 0 \quad \forall t \in [0, 1] \\ \Leftrightarrow \int_0^1 \left\langle x(t) - x^\Delta(t + \tau), -\frac{d}{d\tau} x^\Delta(t + \tau) \right\rangle dt &= 0. \end{aligned}$$

If we define  $\zeta_1(\tau, t) = \langle x(t) - x^\Delta(t + \tau), -\frac{d}{d\tau} x^\Delta(t + \tau) \rangle$ , we then have to find a zero  $\tau_0$  of the

$N$	$T^\Delta$	$ T^{300} - T^\Delta $	$\frac{ T^{300} - T^\Delta }{h^8}$
10	1.544168239138914	2.6730e-9	0.0302
15	1.544168236548486	8.2591e-11	0.0356
20	1.544168236478096	1.2202e-11	0.0170
25	1.544168236467308	1.4131e-12	0.0265
30	1.544168236466391	4.9694e-13	0.0096
35	1.544168236465983	8.8818e-14	0.0205
40	1.544168236465911	1.6209e-14	0.0046

Table 7: Number of mesh intervals, computed period, absolute error for the period, absolute error for the period compared with the expected order of convergence.

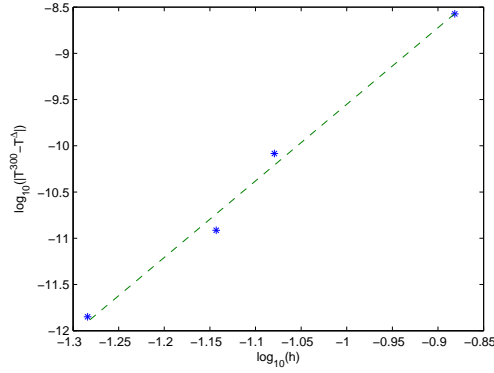


Figure 3: Least squares linear fit:  $\log(|T^{300} - T^\Delta|) = 8.279 \log_{10}(h) - 1.272$ .

function

$$G_1(\tau) = \int_0^1 \zeta_1(\tau, t) dt.$$

We compute this zero by a straightforward bisection method, whereby the integrals  $G_1(\tau)$  are computed by Gauss-Legendre quadrature over the 300-mesh.

Since the solution to the adjoint equation is obtained as a byproduct of the computation of the solution to (10), both have the same phase shift. So to compare the exact  $v_l$  to the computed solution we calculate

$$H(\tau_0) = \sqrt{\int_0^1 \|v_l(t) - v_l^\Delta(t + \tau_0)\|^2 dt},$$

where  $\tau_0$  is the obtained phase shift.

Table 8 gives the phase shift and the value of  $H$ . In Figure 4 and Figure 5 we plot the points  $(\log_{10}(h), \log_{10}(H))$  for the discretizations with  $10 \leq N \leq 25$ , and with  $10 \leq N \leq 40$ , respectively. The slope of the least squares linear fit through the points is given by 5.189 in the first case and by 5.109 in the second case. This again confirms the order of convergence  $h^{m+1}$  of  $v_l^\Delta$ ; we note that in this example we have used a bound on the average error.

$N$	$\tau_0$	$H$
10	0.014943	1.1974e-6
15	0.013715	1.4119e-7
20	0.011749	3.8061e-8
25	0.010210	1.0805e-8
30	0.009844	5.0516e-9
35	0.008517	2.5690e-9
40	0.008983	1.9084e-9

Table 8: Number of mesh intervals, phase shift for the adjoint solution, error for the adjoint solution.

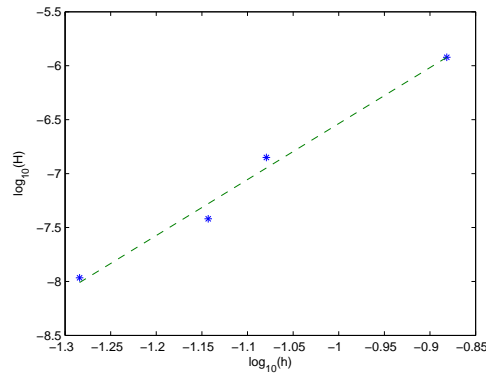


Figure 4: Least squares linear fit:  $\log_{10}(H) = 5.189 \log_{10}(h) - 1.347$ .

#### 4. Conclusion

We have proved that the approximation by collocation at Gauss points for the system adjoint to the linearization of the periodic orbit equations can be obtained with very little computational cost from the solution of the periodic orbit equations. Moreover, the solution is obtained in a form which is most suitable for the applications, i.e. to compute integrals of the form  $\int v_l(t)^T \zeta(t) dt$ , where  $v_l$  is the solution to the adjoint equations and  $\zeta$  is a continuous function. The (high) order of convergence to the exact solution to the adjoint equation is inherited from a standard collocation method.

We have considered two model cases to support our results. In the first example, the solution to the adjoint system is analytically known, in the second it is not. In both examples, we have investigated the orders of convergence for the period and the solution to the adjoint system. In all cases we have given numerical values for the errors and demonstrated the order of convergence by log-log linear fits.

#### References

- [1] U. Ascher, J. Christiansen, R.D. Russell, Collocation software for boundary value ODEs, ACM TOMS 7 (2) (1981) 209-222.
- [2] U. Ascher, R.D. Russell, Reformulation of boundary value problems into "standard" form, SIAM Review 23 (2) (1981) 238-254.

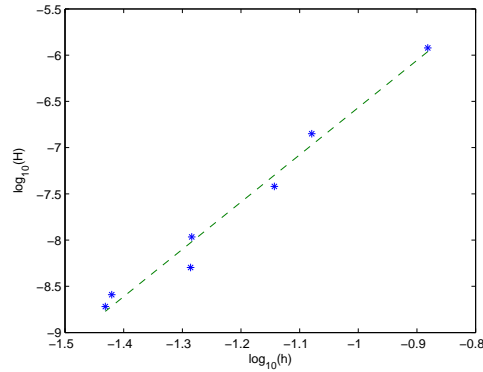


Figure 5: Least squares linear fit:  $\log_{10}(H) = 5.109 \log_{10}(h) - 1.458$ .

- [3] E. Brown, J. Moehlis, P. Holmes, On the phase reduction and response dynamics of neural oscillator populations, *Neural Comput.* 16 (2004) 673-715.
- [4] CONTENT is available via <http://www.math.uu.nl/people/kuznet/CONTENT>.
- [5] C. De Boor, B. Swartz, Collocation at Gaussian points, *SIAM J. Numer. Anal.* 10 (1973) 582-606.
- [6] P. Deuffhard, A. Hohmann, *Numerische Mathematik: Eine Algorithmisch Orientierte Einföhrung*, De Gruyter, Berlin, 1991.
- [7] A. Dhooge, W. Govaerts, Yu.A. Kuznetsov, MATCONT: A Matlab package for numerical bifurcation analysis of ODEs, *ACM Trans. Math. Software* 29 (2003) 141-164. Available via <http://sourceforge.net/projects/matcont/>.
- [8] A. Dhooge, W. Govaerts, Yu.A. Kuznetsov, H.G.E. Meijer, B. Sautois, New features of the software MatCont for bifurcation analysis of dynamical systems, *Math. Comput. Model. Dyn. Syst.* 14 (2008) 147-175.
- [9] E.J. Doedel, A.R. Champneys, T.F. Fairgrieve, Yu.A. Kuznetsov, B. Sandstede, X.J. Wang, AUTO97: Continuation and Bifurcation Software for Ordinary Differential equations (with HomCont), Concordia University, Montreal, 1997, <http://cmvl.cs.concordia.ca/>.
- [10] E. Doedel, H.B. Keller, J.P. Kernevez, Numerical analysis and control of bifurcation problems, Part II: Bifurcation in infinite dimensions, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* 1 (1991) 745-772.
- [11] E.J. Doedel, J.P. Kernevez, AUTO: Software for continuation problems in ordinary differential equations with applications, Applied Mathematics, California Institute of Technology, Pasadena, 1986.
- [12] B. Ermentrout, *Simulating, analyzing and animating dynamical systems: A guide to XPPAUT for researchers and students*, SIAM, Philadelphia, 2002.
- [13] B. Ermentrout, Type I membranes, phase resetting curves and synchrony, *Neural Comput.* 8 (1996) 979-1001.
- [14] W. Govaerts, Yu.A. Kuznetsov, Interactive Continuation Tools, in: B. Krauskopf, H.M. Osinga, J. Galan-Vioque (Eds.), *Numerical Continuation Methods for Dynamical Systems*, Springer, 2007, pp. 51-75.
- [15] W. Govaerts, B. Sautois, Computation of the phase response curve: A direct numerical approach, *Neural Comput.* 18 (2006) 817-847.
- [16] D. Hansel, G. Mato, C. Meunier, Synchrony in excitatory neural networks, *Neural Comput.* 7 (1995) 307-337.
- [17] E. M. Izhikevich, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*, The MIT press, 2007.
- [18] Yu.A. Kuznetsov, W. Govaerts, E.J. Doedel, A. Dhooge, Numerical periodic normalization for codim 1 bifurcations of limit cycles, *SIAM J. Numer. Anal.* 43 (2005) 1407-1435.
- [19] E.N. Lorenz, Irregularity: A fundamental property of the atmosphere, *Tellus*, 36A (1984) 98-110.
- [20] I.G. Malkin, *Metodi Puankare i Liapunova v teorii nelineinix kolebanii* (Methods of Poincaré and Lyapunov in the theory of nonlinear oscillations), Gostexizdat, Moscow, 1949.
- [21] I.G. Malkin, *Nekotorye zadachi teorii nelineinix kolebanii* (Some problems in nonlinear oscillation theory), Gostexizdat, Moscow, 1956.
- [22] L.N. Trefethen, D. Bau, *Numerical Linear Algebra*, SIAM, 1997.
- [23] Z. Yang, H.B. Keller, A direct method for computing higher order folds, *SIAM J. Sci. Stat. Comput.* 7 (2) (1986) 351-361.