Dimension reduction for periodic boundary value problems of functional differential equations

Jan Sieber

December 15, 2010

Periodic boundary-value problems for functional differential equations can be reduced to finite-dimensional algebraic systems of equations. The smoothness assumptions on the right-hand side follow those of the review by Hartung *et al.* (2006) and are set up such that the result can be applied to differential equations with state-dependent delays. For example, this reduction can be used to prove the Hopf bifurcation theorem for functional-differential equations with state-dependent delays. Thus, it provides an alternative to the original proof by Eichmann (2006).

1. Introduction and statement of main result

If a dynamical system is described by a differential equation where the derivative at the current time does not only depend on the state of the current time but on states in the past one speaks of delay-differential or, more generally, functional differential equations (FDEs). A reasonably general formulation of an autonomous dynamical system of this type looks like this:

$$\dot{x}(t) = f(\Delta_t x) \tag{1}$$

where f is a functional, mapping $C^0([-\tau,0];\mathbb{R}^n)$ (the space of continuous functions on the interval $[-\tau,0]$ with values in $\mathbb{R}^n$) into $\mathbb{R}^n$, x is a function on $[-\tau,T_{\max})$ and Δ_t is the time shift by time t: $[\Delta_t x](s) = x(t+s)$. For a system of the form (1) one would have to prescribe a continuous function x on the interval $[-\tau,0]$ as the initial value and then extend x toward time $T_{\max}$ (see textbooks on functional differential equations such as [3, 5, 11]). In applications one is often interested in the long-time behaviour and aims to find attractors or invariant sets.

A long-standing problem with a certain type of FDEs is that they do not fit well into the general framework of smooth infinite-dimensional dynamical system theory. The problem occurs whenever the functional f invokes the evaluation operation in a non-trivial way, that is, for example, if one has a state-dependent delay. A prototypical toy example would be the functional

$$f : C^0(\mathbb{R};\mathbb{R}) \mapsto \mathbb{R}, \qquad f(x) = \mu - x(x(0)), \tag{2}$$

which corresponds to the FDE

$$\dot{x}(t) = \mu - x(t + x(t)). \tag{3}$$

Here, f evaluates its argument x at a point that itself depends on x ($\mu \in \mathbb{R}$ is a fixed parameter). The difficulty stems from the fact that f as a map is in general only as smooth as its argument x. Specifically, the formal derivative of f in this example is

$$\partial_1 f(u)v = -u'(u(0))v(0) - v(u(0)).$$

So, if we choose C^0 as our space in the example as done in (2) then the right-hand side is not differentiable. At first sight the appropriate assumption on f to make is that f is k times continuously differentiable as a map from $C^k(\mathbb{R};\mathbb{R}^n)$ into $\mathbb{R}^n$. However, the review by Hartung *et al* [6] notes that the functionals f typically satisfy a slightly stronger assumption than this. As observed by Walther [14], in a typical derivative only the time derivative of the point in which we differentiate is needed but not the derivative of the linear deviation ($\partial_1 f(u)v$ in the example (2) depends on u' but not on v'). Making this property an explicit requirement, together with the continuity of the derivative $\partial_1 f(u)v$ as a map from $C^1 \times C^0$ into $\mathbb{R}^n$, enabled Walther [14] to prove the existence of a continuously differentiable semifow on a closed submanifold of C^1 (see also review [6]).

This paper focusses on periodic boundary value problems. Periodic orbits are one of the most common invariant sets in dynamical systems. They arise typically when equilibria change their stability and they form the backbone (that is, they form a dense subset) of more complicated attractors. Periodic boundary value problems are formulated for functions x on the unit circle, which we write as $\mathbb{T}$, that is, the interval $[-\pi, \pi]$ with periodic boundary conditions. We pose the FDE on the unit circle $\mathbb{T}$, requiring

$$\dot{x}(t) = f(\Delta_t x) \tag{4}$$

for all $t \in \mathbb{T}$. Choosing the unit circle (that is, fixing the period of x as 2π) is not a restriction. We show in Section 8 that periodic solution of FDEs of arbitrary period satisfy (4) on the unit circle after a re-scaling of time (this is identical the approach taken for periodic orbits of ordinary differential equations). Generally it is impossible to formulate universal existence or uniqueness theorems for periodic boundary value problems except for equations with special structure (for example, equations with negative feedback, see [2]).

However, we can make the following statement in the spirit of the implicit function theorem (and this is the main result of the paper):

Main Result (Reduction to algebraic system)

> Assume that $x_0 \in C^1(\mathbb{T};\mathbb{R}^n)$ has a Lipschitz continuous derivative and that $f : C^0(\mathbb{T};\mathbb{R}^n) \mapsto \mathbb{R}^n$ is smooth to degree k in the extendable sense of Walther [14] (then f is called EC^k smooth in this paper). Then
>
> $$\dot{x} = f(\Delta_t x) \quad \text{if and only if} \quad 0 = g(p) \text{ and } x = X(p)$$
>
> for all x in a sufficiently small neighborhood $U \subset C^1(\mathbb{T};\mathbb{R}^n)$ of x_0, where $p \in D(X) \subset \mathbb{R}^{n_p}$ is finite-dimensional, and the right-hand side $g : D(X) \subset \mathbb{R}^{n_p} \mapsto \mathbb{R}^{n_p}$ and the map $X : D(X) \subset \mathbb{R}^{n_p} \mapsto C^1(\mathbb{T};\mathbb{R}^n)$ are k times continuously differentiable.

Definition 2.1 in Section 2 will state precisely what the notion of extendable continuous (EC^1) differentiability refers to and how it generalizes naturally to higher degrees. The differentiability of the right-hand side of the algebraic system, g, and the map X is meant in the classical sense (their domain finite-dimensional).

This result reduces statements about existence and smooth dependence of periodic solutions to FDEs to root-finding problems of the smooth algebraic equation $0 = g(p)$. For example, the Hopf bifurcation theorem (first proved by Eichmann [4] for FDEs with state-dependent delays) is reduced to a standard smooth algebraic branching problem. We show this as an illustrative example in Section 9. Also other scenarios that are based on branching of periodic solutions (such as the scenarios considered in [7], period doublings or Arnol'd tongues) can be transformed into smooth algebraic problems using the equivalence.

Outline of construction of X and g The techniques in the proof of the reduction theorem are a modification of the construction introduced by Szalai *et al.* [10, 12] that reduced *linear* periodic boundary-value problems to linear systems of algebraic equations to define characteristic matrices for linear periodic FDEs. However, instead of the multiple shooting approach employed in [10] the present construction relies on Fourier modes (which is similar in spirit to the techniques employed in [7]). We introduce the projection of a function x onto its lowest N Fourier modes as P_N and the complementary projection as $Q_N = I - P_N$ (so, for example, $P_0 x$ would be the average of x), and the linear map

$$L : y \in C^k(\mathbb{T};\mathbb{R}^n) \mapsto \int_0^t y(s) - P_0 y \,\mathrm{d}s \in C^k(\mathbb{T};\mathbb{R}^n).$$

Then the linear map $Q_N L$ has the norm $CN^{-1}\log N$ (see Section 3), and the fixed point problem

$$x = p + Q_N L f(\Delta_t x)$$

has a unique fixed point for every p in the (finite-dimensional) image of P_N if the map $F : x \mapsto f(\Delta_t x)$ is Lipschitz continuous (with Lipschitz constant K), and N is such that $CK \log N < N$ (see Section 4). The map $X(p)$ is defined as this fixed point, which depends on p as a parameter. The algebraic equation is then formulated for the finitely many Fourier coefficients of p as

$$0 = P_0 F(X(p)) + Q_0 \left[p - P_N L F(X(p))\right].$$

If p satisfies this algebraic system of equations then $x = X(p)$, which is periodic by construction, satisfies

$$x(t) = c + \int_0^t f(\Delta_s x)\,\mathrm{d}s \quad \text{for some } c \in \mathbb{R}^n,$$

which is equivalent to the FDE $\dot{x}(t) = f(\Delta_t x)$ (see Section 5). A technical difficulty comes from the fact that the map $F : x \mapsto f(\Delta_t x)$ cannot be expected to be locally Lipschitz continuous in any space $C^k(\mathbb{T};\mathbb{R}^n)$. However, extendable continuous differentiability implies a relaxed form of Lipschitz continuity for F [14], which one can use to apply the Banach contraction principle in a bounded closed subset of the space $C^{0,1}(\mathbb{T};\mathbb{R}^n)$ of Lipschitz continuous functions (sets of this type are complete with respect to the C^0-norm). The proof of first-order differentiability in Section 6 repeats this argument for the derivative $\partial X(p)/\partial p$. The proof for higher order degrees of differentiability in Section 7 relies on linear arguments only.

A purely algebraic proof of the Hopf bifurcation theorem, stated in Section 9, demonstrates how one can apply the dimension reduction to make statements about families of periodic orbits in dynamical systems (this is an alternative to the approach used in [4, 7]). The conclusion (Section 10) gives further possible generalizations and applications. To keep the paper self-contained, a separate appendix contains facts and proofs that are required at intermediate steps but are not specific to the problem (they would have to be picked from various parts of the literature, see review [6]).

2. Periodic BVPs

This section gives a precise definition of extendable continuous differentiability as introduced by Walther [14] and applies it recursively to define higher degrees of smoothness.

Let f be a nonlinear functional on the space of periodic functions, that is,

$$f : C^j(\mathbb{T};\mathbb{R}^n) \mapsto \mathbb{R}^n.$$

For $j \geq 0$, $C^j(\mathbb{T};\mathbb{R}^n)$ is the space of all functions x on the unit circle $\mathbb{T}$ with bounded derivatives up to order j (including order 0 and j) satisfying periodic boundary conditions $x^{(l)}(-\pi) = x^{(l)}(\pi)$ for $l = 0 \ldots j$. (Elements of $C^0(\mathbb{T};\mathbb{R}^n)$ are just continuous). The norm in $C^j(\mathbb{T};\mathbb{R}^n)$ is as usual

$$\|x\|_j = \max_{t\in\mathbb{T}} \left\{ |x(t)|, |x'(t)|, \ldots, |x^{(j)}(t)| \right\}.$$

The notation $\mathbb{T}$ indicates that the arguments of x are from the unit circle $\mathbb{T}$, that is, the interval $[-\pi, \pi]$ with periodic boundary conditions. For arbitrary arguments $t \in \mathbb{R}$ we define $x(t)$ as $x(t - 2k\pi)$ choosing k such that $t - 2k\pi \in [-\pi, \pi)$. We permit for f to use the evaluation operator

$$\mathrm{ev} : C^j(\mathbb{T};\mathbb{R}^n) \times \mathbb{T} \mapsto \mathbb{R}^n \quad (j \geq 0) \text{ defined by} \quad \mathrm{ev}(x,s) = x(s)$$

with a second argument that depends on x. For example, in the scalar example (2) (where $n = 1$), the functional f with a fixed parameter μ can be written as

$$f(x) = \mu - \mathrm{ev}(x, \mathrm{ev}(x, 0))$$

such that the second argument of the outer evaluation depends on x. Only the restriction of the operator ev to $C^j(\mathbb{T};\mathbb{R}^n)$ is j times continuously differentiable with respect to its second argument such that the same restriction applies to f: one can only assume that f restricted to $C^j(\mathbb{T};\mathbb{R}^n)$ is j times continuously differentiable for all $j \in \{1, \ldots, j_{\max}\}$.

In a slight deviation from the notation of Hartung *et al* [6] we introduce the time-shift expressly as an operator $\Delta_t : C^j(\mathbb{T};\mathbb{R}^n) \mapsto C^j(\mathbb{T};\mathbb{R}^n)$:

$$[\Delta_t x](s) = x(t + s).$$

The operator Δ_t is linear and has norm 1 in all spaces $C^j(\mathbb{T};\mathbb{R}^n)$. We consider autonomous periodic boundary-value problems for differential equations where f is the right-hand side:

$$\dot{x}(t) = f(\Delta_t x). \tag{5}$$

A function $x \in \mathbb{C}^1(\mathbb{T};\mathbb{R}^n)$ is a solution of (5) if x satisfies equation (5) for all $t \in \mathbb{T}$ (for each $t \in \mathbb{T}$ equation (5) is an equation in $\mathbb{R}^n$).

Additional useful function spaces are the space of Lipschitz continuous functions and, correspondingly, spaces with Lipschitz continuous derivatives, denoted by $C^{j,1}(\mathbb{T};\mathbb{R}^n)$, which are equipped with the norm

$$\|x\|_{j,1} = \max\left\{\|x\|_j, \sup_{t\neq s} \frac{|x^{(j)}(s) - x^{(j)}(t)|}{|s - t|}\right\}$$

($x^{(0)}(t)$ refers to $x(t)$).

The review [6] observed the following typical property of functionals f appearing in equations of type (5): the derivative $\partial_1 f(x)$ of f in x as a linear map from $C^1(\mathbb{T};\mathbb{R}^n)$ into $\mathbb{R}^n$ can be extended to a bounded linear map from $C^0(\mathbb{T};\mathbb{R}^n)$ into $\mathbb{R}^n$, and the mapping

$$\partial_1 f : C^1(\mathbb{T};\mathbb{R}^n) \times C^0(\mathbb{T};\mathbb{R}^n) \mapsto \mathbb{R}^n \quad \text{defined by} \quad (x, y) \mapsto \partial f(x) y$$

is continuous as a function of both arguments.

This paper relies strongly on this notion of *extendable* continuous differentiability, which is a more restrictive condition than merely requiring that f is j times continuously differentiable as a map from $C^j(\mathbb{T};\mathbb{R}^n)$. The precise definition is:

Definition 2.1 (Extendable continuous differentiability EC^k [6])

> Let $J \subset \mathbb{R}$ be a closed interval or $J = \mathbb{T}$, and let D be a function space of the form $D = C^{k_1}(J;\mathbb{R}^{m_1}) \times \ldots \times C^{k_l}(J;\mathbb{R}^{m_l})$ (where $l \geq 1$, $k_j \geq 0$ and $m_j \geq 1$ are integers). We say that $f : D \mapsto \mathbb{R}^n$ has an extendable continuous derivative if there exists a

> map $\partial_1 f$
> $$\partial_1 f(u)\,v : D^1 \times D \mapsto \mathbb{R}^n$$
> that is continuous in both arguments (u,v) and linear in its second argument v such that for all $u \in D^1$
> $$\lim_{\substack{v\in D^1\\ \|v\|_{D,1}\to 0}} \frac{|f(u+v)-f(x)-\partial_1 f(u)\,v|}{\|v\|_{D,1}} = 0. \tag{6}$$
> We say that f is k times continuously differentiable in this extendable sense if the map $\partial_k f$, recursively defined as $\partial_k f = \partial_1[\partial_{k-1} f]$, exists and satisfies the limit condition (6) for $\partial_{k-1} f$. We abbreviate this notion by saying that f is EC^k smooth in D.

In Definition 2.1 we use the notation

$$D^k = \{x : x^{(k)} \in D\} \quad \text{with the norm} \quad \|x\|_{D,k} = \max_{0\le j\le k} \|x^{(j)}\|_D$$

(for example, D^1 is simply D with all differentiability degrees raised by 1). The limit in (6) is a limit in $\mathbb{R}$. The second argument of $\partial_1 f$, v, is not enclosed in the bracket to emphasize that $\partial_1 f$ is linear with respect to v. Extendable continuous differentiability requires the derivative to exist only in points in D^1 but that the derivative as a linear map must extend to D. Also, it does not require that the map $x \in D^1 \mapsto \partial_1 f(x) \in L(D;D)$ is continuous (this would not be true even for example (2)). The functional f of the example (2) is EC^k smooth in $D = C^0(\mathbb{T};\mathbb{R}^n)$ to arbitrary degree k. Its first two derivatives are (putting all arguments into brackets)

$$\begin{aligned}
&\partial_1 f : C^1(\mathbb{T};\mathbb{R}) \times C^0(\mathbb{T};\mathbb{R}),\\
&\partial_1 f(u,v) = -u'(u(0))\,v(0) - v(u(0)), \text{ and}\\
&\partial_2 f : \left[C^2(\mathbb{T};\mathbb{R}) \times C^1(\mathbb{T};\mathbb{R})\right] \times \left[C^1(\mathbb{T};\mathbb{R}) \times C^0(\mathbb{T};\mathbb{R})\right],\\
&\partial_2 f(u,v,w,x) = -\,u''(u(0))\,w(0)\,v(0) - u'(u(0))\,x(0)\\
&\qquad\qquad - w'(u(0))\,v(0) - v'(u(0))\,w(0) - x(u(0)).
\end{aligned}$$

As one can see, the first derivative $\partial_1 f$ has the same structure as f itself if we replace D by $D^1 \times D$. So, it is natural to apply the definition again to $\partial_1 f$ on the space $D^1 \times D$.

Assuming that f is EC^1 smooth on $C^0(\mathbb{T};\mathbb{R}^n)$ implies classical continuous differentiability of f as a map from $C^1(\mathbb{T};\mathbb{R}^n)$ into $\mathbb{R}^n$ and is, thus, strictly stronger than assuming that f is continuously differentiable on $C^1(\mathbb{T};\mathbb{R}^n)$.

If f is EC^1 smooth then it automatically satisfies a relaxed form of local Lipschitz continuity [6], which we call local EC Lipschitz continuity:

Definition 2.2 (Relaxed (EC) Lipschitz continuity)

> We say that $f : C^0(\mathbb{T};\mathbb{R}^n) \mapsto \mathbb{R}^n$ is locally EC Lipschitz continous if for every $x_0 \in C^1(\mathbb{T};\mathbb{R}^n)$ there exists a neighborhood $U(x_0) \subset C^1(\mathbb{T};\mathbb{R}^n)$ and a constant K

such that

$$|f(y)-f(z)|\le K\|y-z\|_0 \tag{7}$$

holds for all y and z in $U(x_0)$.

That EC^1 smoothness implies local EC Lipschitz continuity has been shown, for example, in [14] (but see also Lemma B.3 in Appendix B). Note that the estimate (7) uses the $\|\cdot\|_0$-norm for the upper bound. This is a sharper estimate than one would obtain using the expected $\|\cdot\|_1$-norm. The constant K may depend on the derivatives of the elements in $U(x_0)$ though. For example, for $f(x)=\mu-x(x(0))$ as in (2) one would have the estimate

$$\left|f(x+y)-f(x)\right|\le\left[1+\|x'\|_0\right]\|y\|_0 \quad \text{such that} \quad K\le 1+\max_{x\in U(x_0)}\|x\|_1.$$

For the dimension reduction in the following and in Section 4 and Section 5 only local EC Lipschitz continuity in the sense of Definition 2.2 is necessary. We need that f is EC^1 smooth only during the proof of continuous differentiability of the right-hand side of the algebraic system in Section 6. For higher order differentiability we assume that f is EC^k smooth for degrees k up to $j_{\max}$.

The following corollary states that we can extend the neighborhood $U(x)$ in Definition 2.2 into the space of Lipschitz continuous functions ($C^{0,1}$ instead of C^1) and include time shifts (which possibly increases the bound K).

Corollary 2.3 (EC Lipschitz continuity uniform in time)

Let f be locally EC Lipschitz continuous. Then for every $x_0\in C^{0,1}(\mathbb{T};\mathbb{R}^n)$ there exists a neighborhood $U(x_0)\subset C^{0,1}(\mathbb{T};\mathbb{R}^n)$ and a constant K such that

$$|f(\Delta_t y)-f(\Delta_t z)|\le K\|\Delta_t y-\Delta_t z\|_0=K\|y-z\|_0 \tag{8}$$

holds for all y and z in $U(x_0)$, and for all $t\in\mathbb{T}$.

See Lemma B.3 and Lemma B.4 in Appendix B for the proof of Corollary 2.3.

A consequence of Corollary 2.3 is that the time derivative of a solution is also Lipschitz continuous (in time): if x_0 is a solution of (5) then there exists a constant K such that

$$\|x_0'(t)-x_0'(s)\|_0\le K|t-s| \tag{9}$$

Thus, $x_0\in C^{1,1}(\mathbb{T};\mathbb{R}^n)$. This follows from Corollary 2.3 by inserting $\Delta_t x_0$ and $\Delta_s x_0$ for y and z and using that $x_0'(t)=f(\Delta_t x_0)$ (it is enough to show (9) for $|t-s|$ small).

3. Fourier subspace projection

Consider the functions on $\mathbb{T}$

$$b_0=t\mapsto\frac{1}{2},\quad b_k=t\mapsto\cos(kt),\quad b_{-k}=t\mapsto\sin(kt)$$

for $k = 1, \ldots, \infty$ (which is the classical Fourier basis of $\mathbb{L}^2(\mathbb{T};\mathbb{R})$), and define the projectors and maps

$$\begin{aligned}
P_N &: C^j(\mathbb{T};\mathbb{R}^n) \mapsto C^j(\mathbb{T};\mathbb{R}^n) && [P_N x](t)_i = \sum_{k=-N}^{N} \left[\frac{1}{\pi}\int_{-\pi}^{\pi} b_k(s)x_i(s)\,\mathrm{d}s\right] b_k(t) \\
Q_N &= I - P_N \\
E_N &: \mathbb{R}^{n\times(2N+1)} \mapsto C^j(\mathbb{T};\mathbb{R}^n) && [E_N p](t)_i = \sum_{k=-N}^{N} p_{i,k} b_k(t) \\
R_N &: C^j(\mathbb{T};\mathbb{R}^n) \mapsto \mathbb{R}^{n\times(2N+1)} && [R_N x]_{i,k} = \frac{1}{\pi}\int_{-\pi}^{\pi} b_k(s)x_i(s)\,\mathrm{d}s \\
L &: C^j(\mathbb{T};\mathbb{R}^n) \mapsto C^j(\mathbb{T};\mathbb{R}^n) && [Lx](t) = \int_0^t x(s) - R_0 x\ \mathrm{d}s = \int_0^t Q_0[x](s)\,\mathrm{d}s.
\end{aligned}$$

The projector P_N projects a periodic function onto the subspace spanned by the first $2N+1$ Fourier modes, and Q_N is its complement. The map E_N maps a vector p of $2N+1$ Fourier coefficients (which are each vectors of length n themselves) to the periodic function that has these Fourier coefficients. The map R_N extracts the first $2N+1$ Fourier coefficients from a function. The simple relation $P_N = E_N R_N$ holds. $R_0 x$ is the average of a function x, and Q_0 subtracts the average from a periodic function. The operator L takes the anti-derivative of a periodic function after subtracting its average (to ensure that L maps back into the space of periodic functions). In all of the definitions the degree of smoothness, j, of the vector space C^j can be any non-negative integer. The operator L maps not only C^j back into itself but it maps $C^j(\mathbb{T};\mathbb{R}^n)$ into $C^{j+1}(\mathbb{T};\mathbb{R}^n)$.

The following Lemma states the well-known fact that, roughly, integrating a function makes its high-frequency Fourier coefficients smaller.

Lemma 3.1 (Decay of Fourier coefficients of integrals)

> The norm of the linear operator $Q_N L$, mapping the space $C^j(\mathbb{T};\mathbb{R}^n)$ back into itself, is bounded by
> $$\|Q_N L\|_j \le C\frac{\log N}{N}$$
> where C is a constant. The same holds in the Lipschitz norm (with the same constant C):
> $$\|Q_N L\|_{0,1} \le C\frac{\log N}{N}.$$

See Appendix A for proof.

A direct consequence of Lemma 3.1 is that the Lipschitz norm of $Q_N x$, $\|Q_N x\|_{0,1}$, goes to zero for $N \to \infty$ for elements of $C^{1,1}(\mathbb{T}\,\mathbb{R}^n)$, so, for example, for solutions x_0 of periodic boundary value problems:

$$\|Q_N x\|_{0,1} = \|Q_N L x'\|_{0,1} \le C\frac{\log N}{N}\|x'\|_{0,1} \le C\frac{\log N}{N}\|x\|_{1,1}. \tag{10}$$

4. Fixed point problem

We plan to construct an equivalent system of algebraic equations for the original periodic boundary value problem (5). The nonlinearity f, together with the shift Δ_t, creates a nonlinear operator in $C^0(\mathbb{T};\mathbb{R}^n)$, defined as

$$F : C^0(\mathbb{T};\mathbb{R}^n) \mapsto C^0(\mathbb{T};\mathbb{R}^n) \qquad\qquad [F(x)](t) = f(\Delta_t x). \tag{11}$$

Let x_0 be an element of $C^{1,1}(\mathbb{T};\mathbb{R}^n)$, for example, a solution of the periodic boundary value problem (5), $\dot{x}(t) = f(\Delta_t x) = F(x)(t)$. Consider a closed ball $B_\delta^{0,1}(x_0)$ of radius δ around x_0 in the Lipschitz norm:

$$B_\delta^{0,1}(x_0) = \{x \in C^{0,1}(\mathbb{T};\mathbb{R}^n) : \|x - x_0\|_{0,1} \le \delta\}.$$

The superscript "0, 1" indicates which norm is used to measure the distance from x_0 and that only elements of $C^{0,1}(\mathbb{T};\mathbb{R}^n)$ are included. The Lipschitz norm of $F(x)$ for $x \in B_\delta^{0,1}(x_0)$ is bounded:

Corollary 4.1 (Lipschitz boundedness of F)

> Let f be locally EC Lipschitz continuous. For sufficiently small δ, the nonlinearity F maps $B_\delta^{0,1}(x_0)$ into a bounded ball $B_R^{0,1}(0)$ in the Lipschitz norm:
>
> $$\|F(x)\|_{0,1} \le R \quad \text{for all } x \in B_\delta^{0,1}(x_0).$$

See Lemma B.4 in Appendix B for the proof.

Corollary 2.3 implies that F is Lipschitz continuous with respect to the $\|\cdot\|_0$-norm in $B_\delta^{0,1}(x_0)$ if we choose δ sufficiently small:

$$\|F(x) - F(y)\|_0 \le K\|x - y\|_0 \tag{12}$$

for all x and y in $B_\delta^{0,1}(x_0)$ and a fixed $K > 0$.

We can now formulate a lemma about the unique solvability of the fixed point problem

$$x = E_N p + Q_N L F(x).$$

This unique solvability allows us to reduce the periodic boundary value problem to a system of algebraic equations. Remember that $E_N p$ takes a vector p of $2N+1$ Fourier coefficients and maps it to the periodic function having these Fourier coefficients, $R_N x$ extracts the first $2N+1$ Fourier coefficients from a periodic function x, $P_N x$ projects the periodic function x onto the space spanned by the basis $b_{-N}, \ldots, b_N$ and $Q_N = I - P_N$ sets the first Fourier modes of a function to zero. (P_N and Q_N are projections in the function space, and R_N and E_N map between the finite-dimensional subspace $\operatorname{rg} P_N$ and $\mathbb{R}^{n\times(2N+1)}$.)

Lemma 4.2 (Unique solvability of fixed point problem)

> Let x_0 be in $C^{1,1}(\mathbb{T};\mathbb{R}^n)$, and let $\delta > 0$ be sufficiently small such that
>
> $$\|F(x)\|_{0,1} \le R \quad \text{and} \quad \|F(x) - F(y)\|_0 \le K\|x - y\|_0 \tag{13}$$

for all x and $y \in B^{0,1}_\delta(x_0)$ and for some constants $K > 0$ and $R > 0$ depending on δ. Then for any sufficiently large N the fixed point problem

$$x = E_N p + Q_N LF(x) \tag{14}$$

has a unique solution $x \in B^{0,1}_\delta(x_0)$ for all vectors $p \in \mathbb{R}^{n\times(2N+1)}$ sufficiently close to $R_N x_0$.

Proof The idea is, of course, that the function

$$x \mapsto E_N p + Q_N LF(x)$$

maps the closed ball $B^{0,1}_\delta(x_0)$ back into itself and is uniformly contracting for suitably large N and suitable vectors $p \in \mathbb{R}^{n\times(2N+1)}$.

First, the closed ball $B^{0,1}_\delta(x_0)$ is closed (and, thus, forms a complete metric space) with respect to the $\|\cdot\|_0$-norm. This completeness is a simple continuity argument: let $y_n = x_0 + z_n$ be a fundamental sequence in $B^{0,1}_\delta(x_0)$ with respect to the $\|\cdot\|_0$-norm. Then z_n converges to a continuous function z, and, since $\|z_n\|_0 \le \|z_n\|_{0,1} \le \delta$, for all n, the maximum norm of z is also bounded by δ: $\|z\|_0 \le \delta$. We only have to show that the Lipschitz constant of z is bounded by δ, too. Let $\epsilon > 0$ be arbitrary and let $t \ne s$ be arbitrary in $\mathbb{T}$. We select some n such that $\|z - z_n\|_0 < \epsilon|t-s|/2$. Then

$$\begin{aligned} |z(t) - z(s)| &\le |z(t) - z_n(t)| + |z_n(t) - z_n(s)| + |z_n(s) - z(s)| \\ &< \epsilon|t-s| + \delta|t-s| \le (\delta+\epsilon)|t-s|. \end{aligned}$$

Thus, the Lipschitz constant of z is less than $\delta + \epsilon$ for arbitrary $\epsilon > 0$. Hence, $\|z\|_{0,1} \le \delta$, completing the argument for completeness of $B^{0,1}_\delta(x_0)$ with respect to the $\|\cdot\|_0$-norm.

Let us denote the closed ball $B^{0,1}_\delta(x_0)$, equipped with the metric induced by the $\|\cdot\|_0$-norm, by B. We choose N large enough, such that

$$\|Q_N x_0\|_{0,1} \le \frac{\delta}{3}, \quad \|Q_N L\|_{0,1} \le \frac{\delta}{3R}, \quad \text{and} \quad \|Q_N L\|_0 \le \frac{1}{2K} \tag{15}$$

where R and K are the bounds on F given in (13), which we know exist due to Corollary 2.3 (see Equation (12)) and Corollary 4.1. We know that choosing N according to (15) is possible from Lemma 3.1 and and estimate (10).

Finally, after choosing N, we choose the neighborhood of $R_N x_0 \in \mathbb{R}^{n\times(2N+1)}$ from which the vectors p in the fixed point problem (14) are permitted to come (these vectors p act as parameters for the fixed point problem). We choose the neighborhood $U(R_N x_0) \subset \mathbb{R}^{n\times(2N+1)}$ such that

$$\left\| E_N \left[p - R_N x_0 \right] \right\|_{0,1} \le \frac{\delta}{3} \quad \text{for all } p \in U(R_N x_0). \tag{16}$$

This contains an open set of $\mathbb{R}^{n\times(2N+1)}$ since E_N is an isomorphism between $\operatorname{rg} P_N$, equipped with the $\|\cdot\|_{0,1}$-norm, and $\mathbb{R}^{n\times(2N+1)}$.

Let us check first that $x \mapsto E_N p + Q_N LF(x)$ maps B (the closed ball $B_\delta^{0,1}(x_0)$) back into itself:

$$\begin{aligned}\left\|E_N p + Q_N LF(x) - x_0\right\|_{0,1} &\leq \left\|E_N\left[p - R_N x_0\right] - Q_N x_0 + Q_N LF(x)\right\|_{0,1}\\ &\leq \left\|E_N\left[p - R_N x_0\right]\right\|_{0,1} + \left\|Q_N x_0\right\|_{0,1} + \left\|Q_N L\right\|_{0,1}\|F(x)\|_{0,1}\\ &\leq \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3R}R = \delta.\end{aligned}$$

Here we used the bounds (15) implied by our choice of N and the set of permitted p, given by (16), and the bound on $\|F(x)\|_{0,1}$, which is determined in (13) by our choice of δ.

Second, let us check that $x \mapsto E_N p + Q_N LF(x)$ is a uniform contraction in B with respect to the metric in B, $(x, y) \mapsto \|x - y\|_0$:

$$\left\|Q_N L\left[F(x) - F(y)\right]\right\|_0 \leq \left\|Q_N L\right\|_0 \left\|F(x) - F(y)\right\|_0 \leq \frac{1}{2K} K\|x - y\|_0 \leq \frac{1}{2}\|x - y\|_0.$$

Again, we exploited the bounds (15), implied by our choice of N, and the local Lipschitz constant of F determined in (13) by our choice of δ.

Since B is complete the Banach contraction mapping principle implies that the fixed point problem (14) has a unique solution $x \in B_\delta^{0,1}(x_0)$. □

5. Reduction to algebraic system of equations

Let x_0 be an element of $C^{1,1}(\mathbb{T};\mathbb{R}^n)$, f be locally EC Lipschitz continuous and δ be such that we can apply Lemma 4.2 to establish a unique solution of the fixed point problem (14). This unique solution defines a map

$$X : U(R_N x_0) \subset \mathbb{R}^{n\times(2N+1)} \mapsto B_\delta^{0,1}(x_0) \subset C^{0,1}(\mathbb{T};\mathbb{R}^n) \tag{17}$$

where $X(p)$ is the unique solution x of

$$x = E_N p + Q_N LF(x). \tag{18}$$

This map X maps a vector p into the space of Lipschitz continuous periodic functions, $C^{0,1}(\mathbb{T};\mathbb{R}^n)$. At the moment X is only defined for the local neighborhoods $U(R_N x_0)$ and $B_\delta^{0,1}(x_0)$.

The following theorem is the first part of the central statement of the paper. It claims that the periodic boundary value problem is equivalent to an algebraic system of equations:

Theorem 5.1 (Dimension reduction)

> A function $x \in \left(R_N^{-1}U(R_N x_0)\right) \cap B_\delta^{0,1}(x_0)$ is a solution of $\dot{x}(t) = f(\Delta_t x)$ if and only if $x = X(p)$ where $p = R_N x \in U(R_N x_0) \subset \mathbb{R}^{n\times(2N+1)}$, and the vector p satisfies the

system of $n\times(2N+1)$ algebraic equations

$$0 = R_N\left(P_0F(X(p)) + Q_0\left[E_Np - P_NLF(X(p))\right]\right). \tag{19}$$

Equation (19) consists of $n\times(2N+1)$ equations since the operation R_N extracts the lowest $2N+1$ Fourier coefficients from a function. The expression inside the parentheses is a sum of two parts that each have to be zero: P_0 extracts the average (Fourier coefficient with index 0) of an expression, and Q_0 projects onto all other Fourier modes, discarding the average. Hence, $p\in\mathbb{R}^{n\times(2N+1)}$ satisfies (19) if and only if there exists a constant $c\in\mathbb{R}^n$ such that the pair (c,p) satisfies the system of equations

$$0 = R_0F(X(p)) \tag{20}$$
$$c = E_Np - P_NLF(X(p)). \tag{21}$$

In this system, (20) ensures that the average of the time derivative of $X(p)$ is zero. Equation (21) is an equation in the finite-dimensional space $\operatorname{rg}P_N$. In combination with the defining fixed point problem for $X(p)$, (18), it ensures that

$$X(p) = c + LF(X(p)).$$

This is the integral formulation of the differential equation $X(p)' = F(X(p))$, keeping in mind that $[Ly](t) = \int_0^t y(s) - R_0y\,\mathrm{d}s$ and $R_0F(X(p)) = 0$.

Proof of Theorem 5.1 Theorem 5.1 is a simple consequence of the above mentioned arguments.

"$\Rightarrow$": Let $x\in B_\delta^{0,1}(x_0)$ be such that $p := R_Nx\in U(R_Nx_0)$, and assume that x satisfies $x'(t) = F(x)(t)$. Then x satisfies the integral formulation

$$x(t) = x(0) + \int_0^t F(x)(s)\,\mathrm{d}s \quad \text{for all } t\in\mathbb{T}. \tag{22}$$

Furthermore, the average of x' must be zero, that is,

$$R_0x' = R_0F(x) = 0.. \tag{23}$$

Since $Ly = \int_0^t y(s) - R_0y\,\mathrm{d}s$, the identity (22) has the form

$$x(t) = x(0) + [LF(x)](t). \tag{24}$$

Applying projection Q_N to this identity we obtain $Q_Nx = Q_NLF(x)$, and, since $Q_Nx = x - P_Nx = x - E_Np$,

$$x = E_Np + Q_NLF(x).$$

Consequently, $x = X(p)$ where p was defined as $p = R_Nx$. This implies $0 = R_0F(X(p))$ due to (23). Applying projection P_N to identity (24) we obtain (using $P_Nx = E_Np$)

$$E_Np = x(0) + P_NLF(x),$$

which is equation (21) with the constant $c = x(0)$.

"$\Leftarrow$": Let $p \in U(R_N x_0)$, and assume that p satisfies (19). Then $X(p)$ is well defined and $x = X(p)$ satisfies

$$\begin{aligned} x &= E_N p + Q_N LF(x) && \text{by definition of } X(p), \\ c &= E_N p - P_N LF(x) && \text{for some } c \in \mathbb{R}^n \text{ due to (21), and} \\ 0 &= R_0 F(x) && \text{due to (20),} \end{aligned}$$

Combining the first two equations gives $x = c + LF(x)$, which reads in full

$$x(t) = c + \int_0^t F(x)(s) - R_0 F(x)\,\mathrm{d}s.$$

Inserting the last equation $0 = R_0 F(x)$ we obtain the integral formulation for the differential equation $x' = F(x)$. Also, $x = X(p)$ is periodic by construction. □

We cannot expect that the algebraic system (19) has isolated solutions (except if $X(p)$ is constant) since, whenever x is a solution of the periodic boundary value problem $\dot{x}(t) = f(\Delta_t x)$, so is $\Delta_t x$ for any $t \in \mathbb{T}$ (this is not specific to DDEs but to autonomous problems). We will introduce additional parameters in Section 8 (in particular, the unknown frequency ω) into the right-hand side in order to get regular branches of periodic orbits. This part is analogous to the theory of periodic orbits in ODEs.

Theorem 5.1 reduces the problem of finding periodic solutions to a differential equation to the problem of finding solutions p of the algebraic system (19). For strong statements about existence of these solutions p we need a certain degree of smoothness of the nonlinearities in the right-hand side: $P_0 F(X(p))$ and $P_N LF(X(p))$. These can be guaranteed only using more restrictive conditions on the right-hand side f (or F) of the differential equation. We will discuss this in Section 6. Using only local EC Lipschitz continuity (Definition 2.2) we can only state local Lipschitz continuity of the right-hand side of the algebraic system:

Lemma 5.2 (Regularity of X and the algebraic system)

For p and q in the domain of definition of X the following holds:

1. the image $X(p)$ is in $C^{1,1}(\mathbb{T};\mathbb{R}^n)$ (that is, $X(p) \in C^1(\mathbb{T};\mathbb{R}^n)$ and its time derivative is Lipschitz continuous),

2. X is locally Lipschitz continuous with respect to the $\|\cdot\|_1$-norm: there exists a constant C_N such that

$$\|X(p) - X(q)\|_1 \le C_N |p - q|,$$

3. the nonlinearity $p \mapsto \left[R_0 F(X(p)), P_N LF(X(p))\right]$ is locally Lipschitz continuous.

Proof For a function $y \in \operatorname{rg} P_N$, differentiation is a bounded operator: $y' = D_N y$: the vector $R_N y$ of the first $2N+1$ Fourier coefficients of a function y and the vector $R_N[y']$ satisfy $R_N[y'] = M_N R_N y$ where M_N is a matrix (independent of y). Hence, $y' = E_N M_N R_N y$ for all $y \in \operatorname{rg} P_N$ such that we can define $D_N = E_N M_N R_N$. By definition of the map X the time derivative of $x = X(p)$ satisfies

$$x' = D_N E_N p + Q_0 F(x) - D_N P_N L F(x). \tag{25}$$

This guarantees that $x \in C^1(\mathbb{T};\mathbb{R}^n)$. Lemma 4.1 ensures that $\|F(x)\|_{0,1} \le R$, which in turn implies that

$$\|x'\|_{0,1} \le \|D_N E_N\|_{0,1}|p| + \|Q_0\|_{0,1}R + \|D_N P_N L\|_{0,1}R,$$

and, thus, $x \in C^{1,1}(\mathbb{T};\mathbb{R}^n)$.

Representation (25) also implies point 2: let $x = X(p)$ and $y = X(q)$ be two functions in the image of X:

$$\|x' - y'\|_0 \le \|D_N E_N\|_0 |p-q| + (\|Q_0\|_0 + \|D_N P_N L\|_0) K \|x-y\|_0, \tag{26}$$

where K was the local Lipschitz constant of f (and F) in $B_\delta^{0,1}(x_0)$. The difference $x-y$ in the $\|\cdot\|_0$-norm is bounded due to the contractivity of the right-hand side in fixed point problem (18) defining X (the $\|\cdot\|_0$-norm was the metric used to apply the contraction mapping principle):

$$\|x-y\|_0 \le \|E_N\|_0|p-q| + \|Q_N L[F(x)-F(y)\|_0 \le \|E_N\|_0|p-q| + \frac{1}{2}\|x-y\|_0.$$

Thus,

$$\|x-y\|_0 \le 2\|E_N\|_0|p-q|,$$

which, combined with (26), gives Lipschitz continuity of X as a map into $C^1(\mathbb{T};\mathbb{R}^n)$:

$$\|x'-y'\|_0 \le \left[\|D_N E_N\|_0 + (\|Q_0\|_0 + \|D_N P_N L\|_0) 2K \|E_N\|_0\right] |p-q| =: C_N |p-q|. \tag{27}$$

Point 3 is a direct consequence of the Lipschitz continuity of F with respect to $\|\cdot\|_0$-norm in $B_\delta^{0,1}(x_0)$, the Lipschitz continuity of X in the $\|\cdot\|_0$-norm, and the fact that X maps into $B_\delta^{0,1}(x_0)$. □

6. First-order differentiability of the algebraic system

Until now we have only used the local EC Lipschitz continuity (in the sense of Definition 2.2) of the nonlinearity F with respect to the $\|\cdot\|_0$-norm. We can expect that the algebraic system (19) has smooth coefficients only if we require more smoothness of the right-hand side f (which enters the nonlinearity F in the algebraic system).

We first discuss first-order differentiability of the map X and the algebraic system (19). For this we assume EC^1 smoothness of f as defined in Definition 2.1. For $x \in C^1(\mathbb{T};\mathbb{R}^n)$

the norm of $\partial_1 f(x)$ as an element of $L(C^0(\mathbb{T};\mathbb{R}^n);\mathbb{R}^n)$ is less than or equal to the local EC Lipschitz constant of f in x.

Let us define the map

$$\partial_1 F : C^1(\mathbb{T};\mathbb{R}^n) \times C^0(\mathbb{T};\mathbb{R}^n) \mapsto C^0(\mathbb{T};\mathbb{R}^n), \qquad \left[\partial_1 F(v)w\right](t) = \partial_1 f(\Delta_t v)\Delta_t w.$$

If $v \in C^1(\mathbb{T};\mathbb{R}^n)$ then $\partial_1 F$ is the derivative of F if its second argument is restricted to $C^{0,1}(\mathbb{T};\mathbb{R}^n)$ (see Corollary B.5 in Appendix B):

$$\lim_{\substack{w \in C^{0,1}(\mathbb{T};\mathbb{R}^n) \\ \|w\|_{0,1} \to 0}} \frac{\|F(v+w) - F(v) - \partial_1 F(v)w\|_0}{\|w\|_{0,1}} = 0. \tag{28}$$

Due to the EC^1 smoothness of f and the compactness of $\mathbb{T}$ the map $\partial_1 F$ is continuous with respect to the $\|\cdot\|_0$-norm as a map of both arguments (in their respective norm), too. Again, we do not enclose the second argument w with the bracket to emphasize that $\partial_1 F$ is linear with respect to w. The norm of the linear map $\partial_1 F(v) \in L(C^0(\mathbb{T};\mathbb{R}^n); C^0(\mathbb{T};\mathbb{R}^n))$ is bounded by the local C^0 Lipschitz constant of F, which is known to exist due to Lemma 2.3.

The additional regularity assumption on f permits us to improve our statements about regularity of X and the algebraic system:

Lemma 6.1 (Continuous differentiability of X and the algebraic system)

> Assume that the right-hand side f is EC^1 smooth in the sense of Definition 2.1. Then the regularity statements about the map X, defined in (18), and the right-hand side of the algebraic system, defined in (19), can be extended:
>
> 1. $X(p)$ is in $C^2(\mathbb{T};\mathbb{R}^n)$ for all p in the domain of definition of X, $D(X)$.
> 2. The map X is continuously differentiable in $D(X)$ as a map into $C^1(\mathbb{T};\mathbb{R}^n)$.
> 3. The nonlinearity $p \mapsto \left[R_0 F(X(p)), P_N L F(X(p))\right]$ of the algebraic system (19) is continuously differentiable.

Proof Let $p \in \mathbb{R}^{n\times(2N+1)}$ be in the interior of the domain of definition of X, $D(X) \subset B_\delta^{0,1}(x_0)$. Lemma B.6 in Appendix B proves that $F(x) \in C^1(\mathbb{T};\mathbb{R}^n)$ for $x \in C^1(\mathbb{T};\mathbb{R}^n)$ (setting $D = C^0(\mathbb{T};\mathbb{R}^n)$ and $k = 0$ in its formulation). This implies the first statement, that $X(p) \in C^2(\mathbb{T};\mathbb{R}^n)$: since

$$X(p) = E_N p + Q_N L F(X(p))$$

and $X(p) \in C^{1,1}(\mathbb{T};\mathbb{R}^n)$ (see Lemma 5.2), $F(X(p))$ is in $C^1(\mathbb{T};\mathbb{R}^n)$, and, thus, $LF(X(p))$ is in $C^2(\mathbb{T};\mathbb{R}^n)$. Hence, $X(p)$ is an element of $C^2(\mathbb{T};\mathbb{R}^n)$, too.

Concerning the second statement: denote $X(p)$ as x, let $q \in \mathbb{R}^{n\times(2N+1)}$ be a direction, and let ϵ be sufficiently small such that $p + hq$ is still in $D(X)$ for $h \in (-\epsilon, \epsilon)$. Let us introduce the difference quotient for $h \in (-\epsilon, \epsilon) \setminus \{0\}$:

$$z(h,q,p) = \frac{X(p+hq) - X(p)}{h} \in C^1(\mathbb{T};\mathbb{R}^n) \subset C^{0,1}(\mathbb{T};\mathbb{R}^n).$$

We first prove that z is continuous in all arguments for h in $(-\epsilon, \epsilon)$ and a suitably small ϵ (in particular, that z has a limit for $h \to 0$). By definition of X, z satisfies the fixed point equation (dropping all arguments from z)

$$z = E_N q + Q_N L \frac{F(x+hz) - F(x)}{h} \tag{29}$$

for $h \neq 0$. Let us introduce

$$\tilde{A}_1(x,z,h) = \begin{cases} [F(x+hz) - F(x)]/h & \text{if } h \neq 0 \\ \partial_1 F(x) z & \text{if } h = 0, \end{cases}$$

which maps $C^1(\mathbb{T};\mathbb{R}^n) \times C^{0,1}(\mathbb{T};\mathbb{R}^n) \times \mathbb{R}$ into $C^0(\mathbb{T};\mathbb{R}^n)$. By construction (see (28)) $\tilde{A}_1$ is continuous in all arguments. Using $\tilde{A}_1$ we extend the fixed point problem (29) to $h = 0$:

$$z = E_N q + Q_N L \tilde{A}_1(x,z,h). \tag{30}$$

In the next step we prove that the fixed point problem (30) has a unique solution:

Proposition 6.2 (Fixed point problem for linearization)

> Let p be in $D(X)$, $q \in \mathbb{R}^{n(2N+1)}$. Then there exists a $\epsilon > 0$ such that the map
>
> $$g : z \mapsto E_N q + Q_N L \tilde{A}_1(X(p), z, h)$$
>
> has a fixed point z_* in $C^1(\mathbb{T};\mathbb{R}^n)$ for all $h \in (-\epsilon, \epsilon)$. The fixed point z_* depends continuously on h, p and q.

The bound ϵ depends on p and $|q|$.

Proof of Proposition 6.2 The map g maps $C^{0,1}(\mathbb{T};\mathbb{R}^n)$ continuously into $C^1(\mathbb{T};\mathbb{R}^n) \subset C^{0,1}(\mathbb{T};\mathbb{R}^n)$. Moreover, since $x' = X'(p)$ and $x = X(p)$ depend continuously on p (see Lemma 5.2 and expression (25)), the map also depends continuously on its parameters p and h. We choose the constants $C_0 > 0$ and $C_1 > 0$ such that

$$C_0 \geq 2\|E_N q\|_0 \tag{31}$$
$$C_1 \geq \|D_N E_N q\|_0 + \left(\|Q_0\|_0 + \|D_N P_N L\|_0\right) K C_0, \tag{32}$$

where K is the Lipschitz constant of F with respect to the $\|\cdot\|_0$-norm in $B_\delta^{0,1}$.

We choose ϵ such that for all z satisfying $\|z\|_{0,1} \leq C_1$ the function $x + hz$ is in $B_\delta^{0,1}(x_0)$ for all $h \in (-\epsilon, \epsilon)$. This implies that for any z_1 and z_2 satisfying $\|z_1\|_{0,1} \leq C_1$ and $\|z_2\|_{0,1} \leq C_1$ we have

$$\begin{aligned} \frac{1}{h}\|F(x+hz_1) - F(x+hz_2)\|_0 &\leq K\|z_1 - z_2\|_0, \\ \|\partial_1 F(x)\,[z_1 - z_2]\|_0 &\leq K\|z_1 - z_2\|_0, \end{aligned} \tag{33}$$

(where K was the Lipschitz constant for F in $B_\delta^{0,1}(x_0)$) and, thus,

$$\|g(z_1) - g(z_2)\|_0 \leq \frac{1}{2}\|z_1 - z_2\|_0 \tag{34}$$

for all $h \in (-\epsilon, \epsilon)$ by choice of N (N was such that $\|Q_N L\|_0 \leq (2K)^{-1}$). This estimate for g implies

$$\|g(z)\|_0 \leq \|E_N q\|_0 + \frac{1}{2}\|z\|_0 \qquad \text{if } \|z\|_{0,1} \leq C_1, \tag{35}$$

since $g(0) = E_N q$.

The time derivative of $g(z)$ exists and its $\|\cdot\|_0$-norm can be estimated by differentiating the definition of g with respect to time in the same manner as we obtained (26):

$$\left\|\frac{\mathrm{d}}{\mathrm{d}t} g(z)\right\|_0 \leq \|D_N E_N q\|_0 + (\|Q_0\|_0 + \|D_n P_N L\|_0) K \|z\|_0. \tag{36}$$

The combination of the bounds (35) and (36) and the definition of the constants C_0 and C_1 guarantee that $g(z)$ maps the set

$$B = \{z \in C^{0,1}(\mathbb{T}; \mathbb{R}^n) : \|z\|_0 \leq C_0 \text{ and } \|z\|_{0,1} \leq C_1\}$$

back into itself. The contraction estimate (34) for the $\|\cdot\|_0$-norm and the completeness of B with respect to the $\|\cdot\|_0$-norm make the contraction mapping principle applicable with a uniform contraction rate for all $p \in D(X)$, all q within a bounded ball of $\mathbb{R}^{n(2N+1)}$ and $h \in (-\epsilon, \epsilon)$. This ensures that the fixed point z_* depends continuously on p, $q \in \mathbb{R}^{n(2N+1)}$ and $h \in (-\epsilon, \epsilon)$ with respect to the $\|\cdot\|_0$-norm (since g is continuous in all arguments).

The time derivative $z'(h, q, p)$ of $z(h, q, p)$ also exists and is continuous in p, q and h: we differentiate the fixed point equation (30) for z with respect to time to get

$$z' = D_N E_N q + Q_0 \tilde{A}_1(X(p), z(h, q, p), h) - D_N P_N L \tilde{A}_1(X(p), z(h, q, p), h),$$

which is a continuous function in all arguments with respect to the $\|\cdot\|_0$-norm. □

Proof of Lemma 6.1 continued Proposition 6.2 showed that $z(h, q, p)$ depends continuously on p, q and h as an element of $C^1(T; \mathbb{R}^n)$. Moreover, $z(0, q, p)$ is linear in q because $z \mapsto \partial_1 F(x) z$ is linear. Consequently, the directional derivative of X, $z(0, q, p)$, is continuous in its base point p and in its direction q, and it is linear in q. Thus, $z(0, p, q)$ is the Frechét derivative: $z(0, q, p) = \partial X(p) q$ in the $\|\cdot\|_0$-norm for $q \in \mathbb{R}^{n\times(2N+1)}$ and $p \in D(X)$.

The third statement of Lemma 6.1 is a consequence of the second statement and the fact that the derivative of F exists in the $\|\cdot\|_0$-norm for elements of $C^1(\mathbb{T}; \mathbb{R}^n)$. We split the difference quotients into two parts:

$$\frac{F(X(p + hq)) - F(X(p))}{h} = \frac{F(X(p) + h\partial X(p) q) - F(X(p))}{h} + \tag{37}$$

$$+ \frac{F(X(p + hq)) - F(X(p) + h\partial X(p) q)}{h} \tag{38}$$

The right-hand side in (37) converges for $h \to 0$ to $\partial_1 F(X(p))[\partial X(p)q]$, since $X(p)$ and $\partial X(p)q = z(0,q,p)$ are in $C^1(\mathbb{T};\mathbb{R}^n)$, and the limit is in the $\|\cdot\|_0$-norm. For the term in (38) we can apply the local EC Lipschitz continuity (again, $\partial X(p)q = z(0,q,p) \in C^1(\mathbb{T};\mathbb{R}^n)$) such that we get

$$\left\| \frac{F(X(p+hq)) - F(X(p)+h\partial X(p)q)}{h} \right\|_0 \leq K \left\| \frac{X(p+hq)-X(p)}{h} - \partial X(p)q \right\|_0,$$

which converges to 0 for $h \to 0$ due to the second statement of the lemma. Consequently, we obtain from the limit of (37) for $h \to 0$ that the directional derivative of $F(X(p))$ in direction q is equal to $\partial_1 F(X(p))[\partial X(p)q]$, which is continuous with respect to p and q and linear in q. Thus,

$$\left[\frac{\partial}{\partial p} F(X(p))\right] q = \partial_1 F(X(p))[\partial X(p)q], \tag{39}$$

and $p \mapsto F(X(p))$ is continuously differentiable with respect to p in the $\|\cdot\|_0$-norm. The linear operators R_0 and $P_N L$ preserve the continuity (and the linearity in q) of (39). □

Corollary 6.3 (Spectral radius of linearized fixed point problem)

> For $x = X(p)$ (where $p \in D(X)$) consider the map
>
> $$M : z \mapsto Q_N L \partial_1 F(x) z.$$
>
> The spectral radius of M as a map from $C^0(\mathbb{T};\mathbb{R}^n)$ back into itself, or as a map from $C^1(\mathbb{T};\mathbb{R}^n)$ back into itself, is less or equal $1/2$.

Proof The upper bound for the spectral radius of M from $C^0(\mathbb{T};\mathbb{R}^n)$ into $C^0(\mathbb{T};\mathbb{R}^n)$ is a consequence of the local Lipschitz constant $1/2$ of $x \mapsto Q_N LF(x)$. Since M is compact as an element of $L(C^k(\mathbb{T};\mathbb{R}^n); C^k(\mathbb{T};\mathbb{R}^n))$ for $k = 0, 1$ the spectral radius is identical to the modulus of the maximal (in modulus) eigenvalue, which is of finite algebraic multiplicity. The eigenvector corresponding to this maximal eigenvalue is an element of $C^1(\mathbb{T};\mathbb{R}^n)$ such that the spectral radius of M is the same for $k = 0$ and $k = 1$. □

Thus, the derivative $z = \partial X(p)q$ of X in p is the unique solution of the contractive linear fixed point problem in $C^1(\mathbb{T};\mathbb{R}^n)$

$$z = E_N q + Q_N L \partial_1 F(X(p)) z. \tag{40}$$

7. Higher degrees of smoothness

We observe that the pair $(x, y) = (X(p), \partial X(p)q)$ satisfies the system of equations

$$\begin{aligned} x &= E_N p + Q_N LF(x) \\ y &= E_N q + Q_N L \partial_1 F(x) y. \end{aligned}$$

This has a similar structure to the original fixed point problem (14) but in dimension $n_2 = 2n$ with the variables (x, y) and parameters (p, q). Thus, we aim to apply a linear version of the arguments of Section 6 recursively, assuming that f is EC^k smooth as recursively defined in Definition 2.1. Throughout this section we assume that f is EC^k smooth for all degrees up to order $j_{\max}$.

Generalizing the definition of F and $\partial_1 F$ we introduce the spaces D_j and the operators $\partial_j F$ for $j \geq 0$:

$$\begin{aligned} D_0 &= C^0(\mathbb{T};\mathbb{R}^n) & D_j &= D_{j-1}^1 \times D_{j-1} \\ \partial_j F : D_j &\mapsto C^0(\mathbb{T};\mathbb{R}^n), & [\partial_j F(x)](t) &= \partial_j f(\Delta_t x). \end{aligned}$$

For the spaces D_j^k we denote the norms as $\|\cdot\|_{j,k}$. The maps $\partial_j F$ are all continuous and map indeed into $C^0(\mathbb{T};\mathbb{R}^n)$ due to the continuity of $\partial_j f$ and Δ_t. It is also clear from the definition that $\partial_{j+k} F = \partial_j[\partial_k F]$ if $j + k \leq j_{\max}$.

The following proposition is a direct consequence of the EC^k smoothness of f (see Lemma B.6 from Appendix B):

Proposition 7.1

> For $j + k \leq j_{\max}$ the operator $\partial_j F$ is a continuous map from D_j^k into $C^k(\mathbb{T};\mathbb{R}^n)$.

Proof of Proposition We start the induction for $k = 0$ (where the statement is a consequence of the continuity of $\partial_j f$ and Δ_t for all $j \leq j_{\max}$) and then increase k step by step. For the inductive step let us assume that for k we know that $\partial_j F : D_j^k \mapsto C^k(\mathbb{T};\mathbb{R}^n)$ is continuous for all $j \leq j_{\max} - k$. Let us fix a $j \leq j_{\max} - k - 1$. We have to show that $\partial_j F$ maps D_j^{k+1} continuously into $C^{k+1}(\mathbb{T};\mathbb{R}^n)$. We know (by inductive assumption) that $\partial_j F$ maps D_j^k continuously into $C^k(\mathbb{T};\mathbb{R}^n)$ and that $\partial_{j+1} F$ maps $D_{j+1}^k = D_j^{k+1} \times D_j^k$ continuously into $C^k(\mathbb{T};\mathbb{R}^n)$. Thus, we can apply Lemma B.6 to $\partial_j F$ (this takes the place of the operator F in Lemma B.6) and $D = D_j^k$, obtaining that $\partial_j F : D_j^{k+1} \mapsto C^{k+1}(\mathbb{T};\mathbb{R}^n)$ is continuous. □

An immediate consequence of this is that $X(p)$ and $\partial X(p)q$ as constructed in Section 6 are as smooth as the right-hand-side:

Corollary 7.2 (Smoothness of X and ∂X in time)

> For every $p \in D(X)$ and every $q \in R^{n(2N+1)}$ the functions $X(p)$ and $\partial X(p)q$ satisfy $X(p) \in C^{j_{\max}+1}(\mathbb{T};\mathbb{R}^n)$ and $\partial X(p)q \in C^{j_{\max}}(\mathbb{T};\mathbb{R}^n)$. Moreover, the maps
>
> $$p \mapsto X(p) \in C^{j_{\max}+1}(\mathbb{T};\mathbb{R}^n) \quad \text{and} \qquad [p,q] \mapsto \partial X(p)q \in C^{j_{\max}}(\mathbb{T};\mathbb{R}^n)$$
>
> are continuous.

Proof The function $x = X(p)$ satisfies $x = E_N p + Q_N L F(x)$. Since F maps $D_0^k = C^k(\mathbb{T};\mathbb{R}^n)$ back into itself continuously for all $k \leq j_{\max}$, $Q_N L$ maps D_0^k into D_0^{k+1} continuously for all k, and $E_N p \in C^\infty(\mathbb{T};\mathbb{R}^n)$, the fixed point equation implies the following: if $x \in D_0^k$

then $F(x) \in D_0^k$, thus, $x = E_N p + Q_N L F(x) \in D_0^{k+1}$(for all $k \leq j_{\max}$). Similarly, $z = E_N q + Q_N L \partial_1 F(x) z$, and $\partial_1 F$ maps D_1^k into D_0^k for all $k \leq j_{\max} - 1$. Thus, the fixed point equation implies: if $z \in D_0^k$ and $x \in D_0^{k+1}$ then $(x, z) \in D_1^k$, thus, $\partial_1 F(x) z \in D_0^k$, thus, $z = E_N q + Q_N L \partial_1 F(x) z \in D_0^{k+1}$ for all $k \leq j\max - 1$. All of the above dependencies are continuous such that the continuous dependence on p and q in the norms of $D_0^{j_{\max}+1}$ and $D_0^{j_{\max}}$, respectively, follows. □

We define inductively the operators F_j by

$$F_0(x) = F(x) \qquad \text{for } x \in D_0 \tag{41}$$

$$F_j \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} F_{j-1}(x) \\ \partial_1 F_{j-1}(x) y \end{bmatrix}, \qquad \text{for } \begin{bmatrix} x \\ y \end{bmatrix} \in D_j = D_{j-1}^1 \times D_j. \tag{42}$$

The operators F_j are combinations of derivatives of F. Thus, we can use Proposition 7.1 to determine which spaces the operators F_j map into:

Corollary 7.3 (Image of right-hand side)

> For $j + l + k \leq j_{\max}$ the operator $\partial_l F_j$ maps D_{j+l}^k continuously into D_j^k. In particular, F_j maps D_j continuously back into itself.

Proof We apply Proposition 7.1 to start our induction over j (for $j = 0$ the statement is identical to Proposition 7.1). For the inductive step let us assume that we know that $\partial_l F_{j-1}$ maps D_{j+l-1}^k continuously into D_{j-1}^k for all k and l satisfying $l + k \leq j_{\max} - j + 1$. By definition (42) of F_j the derivative $\partial_l F_j$ for $l \leq j_{\max} - j$ is

$$\partial_l F_j \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} \partial_l F_{j-1}(x) \\ \partial_{l+1} F_{j-1}(x) y \end{bmatrix} \qquad \text{for } \begin{bmatrix} x \\ y \end{bmatrix} \in D_{l+j} = D_{l+j-1}^1 \times D_{l+j}.$$

The first component, $\partial_l F_{j-1}$ maps D_{l+j-1}^{k+1} continuously into D_{j-1}^{k+1} for all k from 0 to $j_{\max} - l - j$ (shifting index k down by 1) due to the assumption of the inductive step. Similarly, $\partial_{l+1} F_{j-1}$ maps $D_{j+l-1}^{1+k} \times D_{j+l-1}^k = D_{j+l}^k$ continuously into D_{j-1}^k for all k from 0 to $j_{\max} - l - j$, again due to the assumption of the inductive step. Consequently, $\partial_l F_j$ maps $D_{j+l-1}^{k+1} \times D_{j+l-1}^k = D_{j+l}^k$ continuously into D_j^k for all k from 0 to $j_{\max} - l - j$, which is the statement we had to prove for the inductive step. □

Even though the map $x \in D_j^1 \mapsto \partial_1 F_j(x) \in L(D_j; D_j)$ is in general not continuous the following map is:

Proposition 7.4 (Continuity in operator norm)

> For $j < j_{\max}$ the map $x \in D_j^1 \mapsto Q_N L \partial_1 F_j(x) \in L(D_j^1; D_j^1)$ is continuous in $x \in D_j^1$.

Proof of Proposition 7.4 The EC^k smoothness of f (for $k \leq j_{\max}$) implies that F_j is continuously differentiable (in the classical sense) as a map from D_j^1 into D_j. Thus, the map $x \mapsto \partial_1 F_j(x)$ as a map from D_j^1 into $L(D_j^1; D_j)$ is continuous. Since $Q_N L$ maps D_j

continuously into D_j^1, the map $x \mapsto Q_N L\partial_1 F_j(x)$ is continuous as a map from D_j^1 into $L(D_j^1; D_j^1)$. □

The following Theorem provides continuous differentiability of order $j_{\max}$ for X and the $p \mapsto F(X(p))$ if the right-hand side is EC^k smoothness in the sense of Definition 2.1 for $k \le j_{\max}$:

Theorem 7.5 (Smoothness of algebraic system and X)

Define $n_0 = n(2N+1)$ and $n_j = 2^j n_0$, and the maps

$$X_0 : p \in D(X) \subseteq \mathbb{R}^{n_0} \mapsto X(p) \in D_0 \text{ and}$$
$$Y_0 : p \in D(X) \subseteq \mathbb{R}^{n_0} \mapsto F(X(p)) \in D_0,$$

and assume that f is $EC^{j_{\max}}$ smooth. Then the following maps exist and are continuous for all orders j up to $j_{\max}$:

$$\begin{aligned} X_j &: [p,q] \in D(X_j) := D(X_{j-1}) \times \mathbb{R}^{n_{j-1}} \subseteq \mathbb{R}^{n_j} && \mapsto [X_{j-1}(p), \partial X_{j-1}(p)q] \in D_j, \\ Y_j &: [p,q] \in D(X_j) && \mapsto [Y_{j-1}(p), \partial Y_{j-1}(p)q] \in D_j. \end{aligned}$$

The proof of Theorem 7.5 does not require the application of the contraction mapping principle for nonlinear maps. It uses only Corollary 6.3, Corollary 7.3 and Proposition 7.4 inductively.

Proof of Theorem 7.5 The main work is the proof of the existence of X_j, which we will do first.The assumption of the inductive step is comprised of the following two statements. We assume for j:

1. The map $(p_1, p_2) \in D(X_{j-1}) \times \mathbb{R}^{n_{j-1}} \mapsto X_j(p_1, p_2) \in D_j$ exists and is continuous. Moreover, the pair $(x_1, x_2) = X_j(p_1, p_2)$ satisfies
$$x_1 = E_N p_1 + Q_N L F_{j-1}(x_1) \tag{43}$$
$$x_2 = E_N p_2 + Q_N L\partial_1 F_{j-1}(x_1) x_2. \tag{44}$$
2. The linear map $z \mapsto Q_N L\partial_1 F_{j-1}(x_1) z$ maps D_{j-1}^1 back into itself and has spectral radius equal to $1/2$.

Both statements of the assumption of the inductive step have been proven for $j = 1$ in Lemma 6.1 and Corollary 6.3 . Let j be smaller than $j_{\max}$. Let us first establish that the map

$$p = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \mapsto x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = X_j \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{bmatrix} X_{j-1}(p_1) \\ \partial_1 X_{j-1}(p_1) p_2 \end{bmatrix}$$

does not only map continuously into D_j but into D_j^k for all $k \le j_{\max} - j + 1$. The argument is the same as in the proof of Corollary 7.2: the map F_j maps D_j^k continuously back into D_j^k for all $k \le j_{\max} - j$. If $x \in D_j^k$ then $F_j(x) \in D_j^k$, thus, $x = E_N p + Q_N LF(x) \in D_j^{k+1}$ for all $k \leq j_{\max} - j$ (and the dependence on p is continuous because all dependencies are continuous).

For $j+1$ we use the notation $p = (p_1, p_2)$ and $x = (x_1, x_2) = X_j(p)$. Let $q = (q_1, q_2) \in \mathbb{R}^{n_j}$ be arbitrary. We first show that $\partial_1 X_j(p)q$ exists. Consider for sufficiently small h the difference quotient

$$\frac{X_j(p+hq) - X_j(p)}{h} = \frac{1}{h}\begin{bmatrix} X_{j-1}(p_1 + hq_1) - X_{j-1}(p_1) \\ \partial_1 X_{j-1}(p_1 + hq_1)[p_2 + hq_2] - \partial_1 X_{j-1}(p_1)p_2 \end{bmatrix} =: \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

By assumption of the inductive step, X_{j-1} is continuously differentiable such that the first row of this difference quotient has the form

$$z_1(h, p_1, q_1) = \frac{1}{h}\left[X_{j-1}(p_1 + hq_1) - X_{j-1}(p_1)\right] = \int_0^1 \partial_1 X_{j-1}(p_1 + hsq_1)q_1 \,\mathrm{d}s \tag{45}$$

for $h \neq 0$. As established above $(p_1, q_1) \mapsto \partial_1 X_{j-1}(p_1)q_1 \in D^k_{j-1}$ is continuous for all $k \leq j_{\max} - j + 1$ such that

$$z_1(h, p_1, q_1) \in D^{j_{\max}-j+1}_{j-1}$$

($j_{\max} - j + 1 \geq 2$ since $j < j_{\max}$) and depends continuously on its arguments, including $h = 0$. For non-zero h the second row of the difference quotient, z_2, satisfies the fixed point equation (dropping the arguments from z_1, x_1 and x_2)

$$\begin{aligned} z_2 &= E_N q_2 + \tilde{z}(p_1, p_2, q_1, h) + Q_N L \partial_1 F_{j-1}(x_1 + hz_1)z_2 \quad \text{where} \\ \tilde{z}(p_1, p_2, q_1, h) &= Q_N L \frac{\partial_1 F_{j-1}(x_1 + hz_1)x_2 - \partial_1 F_{j-1}(x_1)x_2}{h} \end{aligned} \tag{46}$$

The regularity of x_1, x_2 and z_1 is:

$$\begin{aligned} x_1 &= X_{j-1}(p_1) &&\in D^{j_{\max}-j+2}_{j-1} \subset D^3_{j-1}, \\ x_2 &= \partial_1 X_{j-1}(p_1)p_2 &&\in D^{j_{\max}-j+1}_{j-1} \subset D^2_{j-1} \quad \text{and} \\ z_1 &= \partial_1 X_{j-1}(p_1)q_1 &&\in D^{j_{\max}-j+1}_{j-1} \subset D^2_{j-1}. \end{aligned}$$

We can apply the mean value theorem to the difference quotient appearing in $\tilde{z}$ since x_1 and z_1 are at least in D^2_{j-1} and x_2 is at least in D^1_{j-1} (see Corollary 7.1, and Lemma B.2 and Corollary B.5 in Appendix B):

$$\tilde{z}(p_1, p_2, q_1, h) = Q_N L \int_0^1 \partial_2 F_{j-1}(x_1 + shz_1, x_2)[z_1, 0]\,\mathrm{d}s.$$

The map $(x_1, x_2, z_1, h) \mapsto \partial_2 F_{j-1}(x_1 + shz_1, x_2)[z_1, 0]$ maps x_1, x_2, z_1 and h continuously into the space D^k_{j-1} for $k \leq j - j_{\max} - 1$ (see Corollary 7.1). Thus, the quantity $\tilde{z}(p_1, p_2, q_1, h)$ is in $D^{j_{\max}-j}_{j-1} \subset D^1_{j-1}$ (since $j \leq j_{\max} - 1$) and depends continuously on p_1, p_2, q_1 and h in this space.

Hence, (46) is a linear fixed point problem for z_2 where the inhomogeneity is in $D_{j-1}^{j_{\max}-j}$ and depends continuously on (p,q,h). The linear map $M(h): z_2 \mapsto Q_N L\partial_1 F_{j-1}(x_1+hz_1)z_2$ in front of z_2 on the right-hand side of (46) depends continuously on h as an element of $L(D_{j-1}^1; D_{j-1}^1)$ (see Corollary 7.4 and note that x_1, $z_1 \in D_{j-1}^1$). Since the spectral radius of the map $M(0)$ (for $h=0$) is less or equal than $1/2$ by assumption of our inductive step, the spectral radius of $M(h)$ is less than unity if we choose h sufficiently small. Thus, for all $p \in D(X_j)$ and $q \in \mathbb{R}^{n_j}$ and sufficiently small h, z_2 satisfies a contractive linear fixed point equation with an inhomogeneity in D_{j-1}^1 and a contractive linear map that maps into D_{j-1}^1 where all coefficients depend continuously on (h,p,q). Consequently, z_2 has a limit for $h \to 0$ that depends continuously on (p,q). For $h=0$ the fixed point equation for (z_1,z_2) simplifies to

$$\begin{aligned} z_1 &= E_N q_1 + Q_N L\partial_1 F_{j-1}(x_1)z_1 \\ z_2 &= E_N q_2 + Q_N L\left[\partial_2 F_{j-1}(x_1,x_2)[z_1,0] + \partial_1 F_{j-1}(x_1)z_2\right]. \end{aligned} \tag{47}$$

Both equations are linear in q and $z=(z_1,z_2)$. Consequently, $z(0,p,q)$, which is by definition the directional derivative of X_j in p in direction q, depends linearly on q and continuously on p and q. Consequently,

$$z(0,p,q) = \left[\frac{\partial}{\partial p} X_j(p)\right] q$$

is the Frechét derivative of X_j. By definition of x, z and F_j, the functions x and z satisfy

$$\begin{aligned} x =& E_N p + Q_N L F_j(x) \\ z =& E_N q + Q_N L\partial_1 F_j(x) z, \end{aligned}$$

which completes the proof of statement 1 of the inductive assumption for $j+1$.

Finally, consider the map

$$M_j(x): \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \mapsto \begin{bmatrix} Q_N L\partial_1 F_{j-1}(x_1)z_1 \\ Q_N L\left[\partial_2 F_{j-1}(x_1,x_2)[z_1,0] + \partial_1 F_{j-1}(x_1)z_2\right] \end{bmatrix}$$

We have established already that $x_1 \in D_{j-1}^3$ and $x_2 \in D_{j-1}^2$ (thus, $x=(x_1,x_2) \in D_j^2$). Thus, the first row of $M_j(x)$ (which depends on z_1 but not on z_2) maps D_{j-1}^2 back into itself. The second row maps D_j^1 into D_{j-1}^1 such that, overall, $M_j(x)$ maps D_j^1 back into itself. The map $M_j(x)$ is also lower triangular (the first row does not depend on z_2), and its diagonal terms are:

$$z_k \in D_{j-1}^k \mapsto Q_N L\partial_1 F_{j-1}(x) z_k \in D_{j-1}^k \quad \text{for } k=1,2.$$

The spectral radius of both diagonal maps less or equal than $1/2$ by statement 2 of the inductive assumption (both are eventually compact maps and the eigenvector corresponding to the spectral radius is in D_{j-1}^2). Thus, the overall spectral radius of $M_j(x)$ is also equal to $1/2$, which proves statement 2 of the inductive assumption for $j+1$.

Existence of Y_j We show inductively that $Y_j(p) = F_j(X_j(p))$. For $j = 1$ this statement was proven in Lemma 6.1. Let $j < j_{\max}$ and assume that $Y_j = F_j(X_j(p))$ for $p \in D(X_j)$. Since

$$X_j(p) = E_N p + Q_N L F_j(X_j(p))$$

and F_j maps D_j^1 into D_j^1, X_j is an element of D_j^1. Let $q \in \mathbb{R}^{n_j}$ be arbitrary, and let us denote $(x_1, x_2) = (X_j(p), \partial_1 X_j(p) q) = X_{j+1}(p, q)$. The component x_2 satisfies

$$x_2 = E_N q + Q_N L \partial_1 F_j(x_1) x_2$$

such that x_2 is in D_j^1, too. Consequently,

$$\begin{aligned}
\frac{Y_j(p+hq) - Y_j(p)}{h} &= \frac{F_j(X_j(p+hq)) - F_j(X_j(p))}{h} \\
&= \frac{F_j(x_1 + hx_2) - F_j(x_1)}{h} + \frac{F_j(X_j(p+hq)) - F_j(x_1 + hx_2)}{h}.
\end{aligned} \tag{48}$$

Since F_j is differentiable for $x_1 \in D_j^1$ and deviations $hx_2 \in D_j^1$ the first quotient in the expression (48) converges to $\partial_1 F_j(x_1) x_2$. Since F_j is locally Lipschitz continuous in the $\|\cdot\|_{j,0}$-norm for elements of D_j^1 the second term in (48) can be bounded by

$$\left\| \frac{F_j(X_j(p+hq)) - F_j(X_j(p) + h\partial_1 X_j(p) q)}{h} \right\|_0 \le K_1 \left\| \frac{X_j(p+hq) - X_j(p)}{h} - \partial_1 X_j(p) q \right\|_0,$$

which converges to zero for $h \to 0$ because X_j is differentiable. Consequently, the directional derivative of Y_j in p in direction q is $\partial_1 F_j(X_j(p))[\partial X_j(p) q]$, which is continuous in p and q and linear in q. Therefore, the Frechét derivative of Y_j exists and

$$\left[\frac{\partial}{\partial p} Y_j(p)\right] q = \partial_1 F_j(X_j(p))[\partial X_j(p) q],$$

which implies by definition of F_j and X_j that $Y_{j+1} = F_{j+1}(X_{j+1}(p, q))$. □

We can refine the statement of Theorem 7.5 slightly by noting that $X_j : D(X_j) \mapsto D_j^1$ is continuous for all $j \le j_{\max}$ (instead of $X_j : D(X_j) \mapsto D_j$). This follows from the continuity of $Y_j = F_j(X_j(p))$ as a map into D_j and the relation

$$X_j(p) = E_N p + Q_N L Y_j(p).$$

8. Periodic orbits of autonomous systems

If one wants to find periodic orbits of autonomous differential equations the boundary value problem (5) has to be extended (and, for example, rescaled) in order to make it well-posed (this is also true for ODEs) because the period T of the orbit is unknown and for every periodic solution $x(t)$ of (5) its time shift $x(t + \delta)$ is also a solution. Assume that we have a functional $f(\cdot, \mu)$ (where $\mu \in \mathbb{R}^\nu$ is a system parameter) that is defined for arguments in $C^0(\mathbb{R};\mathbb{R}^n)\times\mathbb{R}^\nu$. Treating $\mathbb{R}^\nu$ as a subspace of $C^0(\mathbb{R};\mathbb{R}^\nu)$ we assume that $f(x,\mu)$ is also well defined as a functional from $C^0(\mathbb{R};\mathbb{R}^n)\times C^0(\mathbb{R};\mathbb{R}^\nu)=C^0(\mathbb{R};\mathbb{R}^{n+\nu})$ into $\mathbb{R}^n$, and is EC^k smooth for all degrees k up to $j_{\max}$ in the sense of Definition 2.1.

Let x be a periodic function of period $T=2\pi/\omega$. Then the function $y(s)=x(s/\omega)$ is a function of period 2π ($s\in\mathbb{T}$). This makes it useful to define the map

$$S: C^0(\mathbb{R};\mathbb{R}^n)\times\mathbb{R}\mapsto C^0(\mathbb{R};\mathbb{R}^n) \qquad\qquad [S(x,\omega)](t)=x(\omega t),$$

such that $S(y,\omega)=x$. Then $x\in C^1(\mathbb{R};\mathbb{R}^n)$ satisfies the differential equation

$$\dot{x}(t)=f(\Delta_t x,\mu) \tag{49}$$

on the real line and has period $2\pi/\omega$ if and only if $y=S(x,1/\omega)\in C^1(\mathbb{T};\mathbb{R}^n)$ satisfies the differential equation

$$\dot{y}(s)=\frac{1}{\omega}f(S(\Delta_s y,\omega),\mu).$$

Let us define an extended differential equation

$$\dot{x}_{\mathrm{ext}}(s)=f_{\mathrm{ext}}(\Delta_s x_{\mathrm{ext}}) \tag{50}$$

where f_{ext} maps $C^0(\mathbb{T};\mathbb{R}^{n+1+\nu})$, into $\mathbb{R}^{n+1+\nu}$, and is defined by

$$f_{\mathrm{ext}}\begin{pmatrix}y\\ \omega\\ \mu\end{pmatrix}=\begin{bmatrix}f\left(S(y,\omega(0)),\mu(0)\right)/\operatorname{cut}(\omega(0))\\ 0\\ 0\end{bmatrix},\quad\text{where}$$

$$\operatorname{cut}(\omega)=\begin{cases}\omega & \text{if } \omega>\omega_{\mathrm{cutoff}}>0\\ \text{smooth, uniformly non-negative extension} & \text{for } \omega<\omega_{\mathrm{cutoff}}\end{cases}$$

for $y\in C^0(\mathbb{T};\mathbb{R}^n)$, $\omega\in C^0(\mathbb{T};\mathbb{R})$ and $\mu\in C^0(\mathbb{T};\mathbb{R}^\nu)$. We have used in our definition that any functional f defined for $x\in C^0(\mathbb{R};\mathbb{R}^n)$ is also a functional on $C^0(\mathbb{T};\mathbb{R}^n)$ (periodic functions have a natural extension $x(t)=x(t \bmod_{[-\pi,\pi)})$ if $t\in\mathbb{R}$ is arbitrary). The extended system has introduced the unknown ω and the system parameter μ as functions of time and the additional differential equations $\dot{\omega}=0$, $\dot{\mu}=0$ which force the new functions to be constant. We have also introduced a cut-off for ω close to zero to keep f_{ext} globally defined. The functional f_{ext} is also EC^k smooth for all degrees $k\le j_{\max}$. Thus, the differential equation (50) satisfies the assumptions of Theorem 5.1 and Theorem 7.5. Any solution (y,ω,μ) we find for (50) corresponds to a periodic solution $t\mapsto y(\omega t)$ of period $2\pi/\omega$ at parameter μ for (49) and vice versa as long as $\omega>\omega_{\mathrm{cutoff}}$.

The resulting system of algebraic equations has $(n+\nu+1)(2N+1)$ variables and equations. Let us denote as $F=(F_y,F_\omega,F_\mu)$ the components of the right-hand side (of which F_μ and F_ω are identically zero), $p=(p_y,p_\omega,p_\mu)$ the $2N+1$ leading Fourier coefficients of y, ω and μ, respectively (these are the variables of the system), and $X(p)=(X_y(p),X_\omega(p),X_\mu(p))$ the map from $R^{(n+\nu+1)(2N+1)}$ into $C^{j_{\max}}(\mathbb{T};\mathbb{R}^{n+\nu+1})$. Then several of the variables can be eliminated and the equations correspondingly simplified. Since F is identically zero in its last $\nu+1$ components we have

$$X_\omega(p)=E_Np_\omega, \qquad\qquad X_\mu(p)=E_Np_\mu.$$

Hence, the algebraic system (19) has $\nu+1$ equations of the type $0=0$ (since $P_0F(X(p))=0$ for the equations $\dot\omega=0$ and $\dot\mu=0$). Furthermore, system (19) contains the equations $Q_0E_Np_\omega=0$ and $Q_0E_Np_\mu=0$, which require that all Fourier coefficients (except the averages $R_0\omega$ and $R_0\mu$) of μ and ω are equal to zero. This means that the algebraic system (19) (unsurprisingly) forces ω and μ to be constant. Thus, we can eliminate $R_NQ_0E_Np_\omega$ and $R_NQ_oE_Np_\mu$ (which are $2N(\nu+1)$ variables), replacing them by zero, and drop the corresponding equations. This leaves the first $n(2N+1)$ algebraic equations

$$0=R_N\left(P_0F_y(X_y(p_y,\omega,\mu),\omega,\mu)+Q_0\left[E_Np_y-P_NLF_y(X_y(p_y,\omega,\mu),\omega,\mu)\right]\right), \tag{51}$$

which depends smoothly (with degree $j_{\max}$) on the $n(2N+1)$ variables p_y and the parameters $\omega\in\mathbb{R}$ and $\mu\in\mathbb{R}^\nu$. We have identified the vector $R_0\mu$ and the scalar $R_0\omega$ with the corresponding constant functions μ and ω in (51).

Rotational Invariance The original nonlinearity F, defined by $[F(x)](t)=f(\Delta_tx)$ is invariant with respect to time shift: $\Delta_tF(x)=F(\Delta_tx)$ for all $t\in\mathbb{T}$ and $x\in C^0(\mathbb{T};\mathbb{R}^n)$. Furthermore, Δ_t commutes with the following operations:

$$\Delta_tQ_NL=Q_NL\Delta_t \quad \text{(if } N\geq 0\text{) and} \qquad\qquad \Delta_tP_N=P_N\Delta_t.$$

This property gets passed on to the algebraic equation in the following sense: let us define the operation Δ_t for a vector p in $\mathbb{R}^{n(2N+1)}$, which we consider as a vector of Fourier coefficients of the function $E_Np\in C^0(\mathbb{T};\mathbb{R}^n)$, by

$$\Delta_tp=R_N\Delta_tE_Np.$$

With this definition Δ_t commutes with R_N and E_N. It is a group of rotation matrices: Δ_t is regular for all t, $\frac{\mathrm{d}}{\mathrm{d}t}\Delta_t|_{t=0}=M_N$ (M_N was differentiation: $(E_Np)'=E_NM_Np$) and $\Delta_{2k\pi}=I$ for all integers k. The definition of $X(p)$ as a fixed point of $x\mapsto E_Np+Q_NLF(x)$ implies that $\Delta_tX(p)=X(\Delta_tp)$. From this it follows that the algebraic system of equations is also invariant. If we denote the right-hand-side of the overall system (51) by $G(p_y,\omega,\mu)$ then G satisfies

$$\Delta_tG(p_y,\omega,\mu)=G(\Delta_tp_y,\omega,\mu) \quad \text{for all } t\in\mathbb{T},\ p_y\in\mathbb{R}^{n(2N+1)},\ \omega>0 \text{ and } \mu\in\mathbb{R}^\nu.$$

9. Illustrative example — Hopf bifurcation

One useful aspect of the reduction theorem is that it provides an alternative proof to the Hopf bifurcation theorem for equations with state-dependent delays. The first proof that the Hopf bifurcation occurs as expected is due to Eichmann [4]. The reduction of periodic boundary-value problems to smooth algebraic equations reduces the Hopf bifurcation problem to an equivariant algebraic pitchfork bifurcation.

Let us consider the equation

$$0 = f(x,\mu) \tag{52}$$

where $f : C^0([-\tau,0];\mathbb{R}^n)\times\mathbb{R}\mapsto\mathbb{R}^n$, $\mu\in\mathbb{R}$, and $x\in C^0([-\tau,0];\mathbb{R}^n)$ is a constant x_0, that is, x satisfies $x(t)=x_0$ for all $t\in[-\tau,0]$ (thus, $x\in C^k([-\tau,0];\mathbb{R}^n)$ for all $k\geq 0$). This means that (52) is a system of n algebraic equations for the $n+1$ variables (x_0,μ). If f is EC^k smooth in the sense of Definition 2.1 then the algebraic system $f(x,\mu)$ (where $x(t)=x_0$) is k times continuously differentiable. Let us assume that this algebraic system has a regular solution $x_0(\mu)\in\mathbb{R}^n$ for μ close to 0. Without loss of generality we can assume that $x_0(\mu)=0$, otherwise, we introduce the new state variable $x_{\mathrm{new}}=x_{\mathrm{old}}-x_0(\mu)\in C^0([-\tau,0];\mathbb{R}^n)$. Hence, $f(0,\mu)=0$ for all μ close to 0. Let us denote the EC^1 derivative of f (in the sense of Definition 2.1) in $(x,\mu)=(0,\mu)$ by $a(\mu)\in L(C^0([-\tau,0];\mathbb{R}^n);\mathbb{R}^n)$. The linear operator $a(\mu)$ can be easily complexified, defining $a(\mu)[x+iy]=a(\mu)[x]+ia(\mu)[y]$ for $x+iy\in C^0([-\tau,0];\mathbb{C}^n)$. If f is EC^2 smooth in the sense of Definition 2.1 then the $n\times n$-matrix $K(\lambda,\mu)$ (called the *characteristic matrix*), defined by

$$K(\lambda,\mu)\,v = \lambda v - a(\mu)[v\exp(\lambda t)] \tag{53}$$

is analytic in its complex argument λ and differentiable in its real argument μ. The Hopf bifurcation theorem states the following:

Theorem 9.1 (Hopf bifurcation)

Assume that f is EC^k smooth ($k\geq 2$) in the sense of Definition 2.1 and that the characteristic matrix $K(\lambda,\mu)$ satisfies the following conditions:

1. **(Imaginary eigenvalue)** there exists a ω_0 such that $\det K(i\omega_0,0)=0$ and $i\omega_0$ is an isolated root of $\lambda\mapsto\det K(\lambda,0)$. We denote the corresponding null vector by $v_1=v_r+iv_i\in\mathbb{C}^n$ (scaling it such that $|v_r|^2+|v_i|^2=1$).

2. **(Non-resonance)** $\det K(ik\omega,0)\neq 0$ for all integers $k\neq\pm1$.

3. **(Transversal crossing)** The root curve $\mu\mapsto\lambda(\mu)$ of $\det K(\lambda,\mu)$ which corresponds to the isolated root $i\omega_0$ at $\mu=0$ (that is, $\lambda(0)=i\omega_0$) has a non-vanishing derivative
$$0\neq\left.\frac{\partial}{\partial\mu}\operatorname{Re}\lambda(\mu)\right|_{\mu=0}.$$

Then there exists a $k-1$ times differentiable curve

$$\beta\in(-\epsilon,\epsilon)\mapsto(x,\omega,\mu)\in C^1(\mathbb{T};\mathbb{R}^n)\times\mathbb{R}\times\mathbb{R}$$

such that for sufficiently small $\epsilon>0$ the following holds:

1. x is a periodic orbit of $\dot{x} = f(\Delta_t x, \mu)$ of period $2\pi/\omega$, that is, $x \in C^1(\mathbb{T}; \mathbb{R}^n)$ and
$$\dot{x}(t) = \frac{1}{\omega} f(S(\Delta_t x, \omega), \mu), \tag{54}$$
2. the amplitude of the first Fourier coefficients of x is equal to β, that is,
$$\begin{aligned} 0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \operatorname{Re}\left[v_1 \exp(it)\right]^T x(t)\,\mathrm{d}t, \quad \text{and} \\ \beta &= \frac{1}{\pi} \int_{-\pi}^{\pi} \operatorname{Im}\left[v_1 \exp(it)\right]^T x(t)\,\mathrm{d}t, \end{aligned} \tag{55}$$
3. $x|_{\beta=0} = 0$, $\mu|_{\beta=0} = 0$ and $\omega|_{\beta=0} = \omega_0$, that is, the solution x, the system parameter μ and the frequency ω of x, which are differentiable functions of the amplitude β, are identical to $x = 0$, $\mu = 0$, $\omega = \omega_0$ for $\beta = 0$.

The statement is identical to the classical Hopf bifurcation theorem for ODEs in its assumptions and conclusions. Note that the existence of the one-parameter family (parametrized in β) automatically implies the existence of a two-parameter family for $\beta \neq 0$ due to the rotational invariance: if x is a solution of (54) then $\Delta_s x$ is also a solution of (54). Condition (55) fixes the time shift s of x such that x is orthogonal to $\operatorname{Re}[v_1 \exp(it)]$ using the $\mathbb{L}^2$ scalar product.

Proof The proof of the Hopf bifurcation theorem is a simple fact-checking exercise. We have to translate the assumptions on the derivative of $f : C^0([-\tau, 0]; \mathbb{R}^n) \mapsto \mathbb{R}^n$ into properties of the right-hand side of the nonlinear algebraic system (51) near $(x, \omega, \mu) = (0, \omega_0, 0)$, and then apply algebraic bifurcation theory to the algebraic system. The only element of the proof that is specific to functional differential equations comes in at the linear level: the fact that the eigenvalue $i\omega_0$ is simple implies that the right nullvector $v_1 \in \mathbb{C}^n$ and any non-trivial left nullvector w_1 satisfy
$$w_1^H \left[\left. \frac{\partial}{\partial \lambda} K(\lambda, 0) \right|_{\lambda = i\omega_0} \right] v_1 \neq 0.$$
This is the generalization of the orthogonality condition $w_1^H v_1 \neq 0$, known from ordinary matrix eigenvalue problems, to exponential matrix eigenvalue problems of the type $K(\lambda, \mu) v = 0$. The proof of Theorem (9.1) is entirely based on the standard calculus arguments for branching of solutions to algebraic systems as can be found in textbooks [1]. Thus, we have relegated the details to Appendix C. □

In applications one often requires differentiability of the nonlinear system of at least degree 3 (that is, f has to be at least EC^3 smooth) and imposes an additional non-degeneracy condition, which ensures that the graph $\mu(\beta)$ has a non-zero second derivative at $\beta = 0$: $\mu''(0) \neq 0$ ($\mu'(0)$ is zero due to rotational symmetry, see proof in Appendix C). The sign of $\mu''(0)$ determines if the Hopf bifurcation is *supercritical* ($\mu''(0) > 0$) or *subcritical* ($\mu''(0) < 0$). In the statement of the Theorem 9.1 only EC^2 smoothness is required for f such that μ could be continuously differentiable once but not twice, leaving criticality undetermined.

10. Conclusion

We have proved that periodic boundary-value problems for functional differential equations (FDEs) are equivalent to systems of smooth algebraic equations if the functional f defining the right-hand side of the boundary-value problem satisfies natural smoothness assumptions. The assumptions are identical to those imposed in the review [6] and do not exclude FDEs with strate-dependent delay. There are several immediate extensions of the results presented in this paper. The list below addresses some of them.

Globally valid algebraic system

The main result was formulated locally in the neighborhood of a given $x_0 \in C^{1,1}(\mathbb{T};\mathbb{R}^n)$ and required only local Lipschitz continuity. The proof makes obvious that the domain of definition for the map X, which maps between the function space and the finite-dimensional space is limited by the size of the neighborhood of x_0 for which one can find a uniform (EC) Lipschitz constant. In problems with delay the right-hand side is typically a combination of Nemytskii operators generated by smooth functions and the evaluation operator ev. These typically satisfy a (semi-)global Lipschitz condition (see also condition (Lb) in [6]): for all R there exists a constant K such that

$$|f(x) - f(y)| \leq K\|x - y\|_0$$

for all x and y satisfying $\|x\|_1 \leq R$ and $\|y\|_1 \leq R$. Under this condition one can for any bounded ball choose an algebraic system that is equivalent to the periodic boundary-value problem in this bounded ball. For periodic orbits of autonomous systems this means that one can find an algebraic system such that all periodic orbits of amplitude less than R and of period and frequency at most R are given by the roots of the algebraic system.

Implicitly given delays

In practical applications the delay is sometimes given implicitly, for example, the position control problem considered in [13] and the cutting problem in [8] contain a separate algebraic equation, which defines the delay implicitly. In simple cases these problems can be reduced to explicit differential equations using the standard reduction for index-1 differential algebraic equations. For example, in the cutting problem the delay τ depends on the current position x via the implicit linear equation

$$\tau(t) = a - bx(t) - bx(t - \tau(t)), \tag{56}$$

which can be transformed into a differential equation by differentiation with respect to time:

$$\dot{\tau}(t) = \frac{-bv(t) - [\tau(t) - a + bx(t) + bx(t - \tau(t))]}{1 + bv(t - \tau(t))} \tag{57}$$

($v(t) = \dot{x}(t)$ is explicitly present as a variable in the cutting model, which is a second order differential equation). The original model accompanied with the differential equation (57) instead of the algebraic equation (56) fits into the conditions of the reduction theorem 5.1. The regularity statement of Lemma 5.2 guarantees that the resulting periodic solutions have Lipschitz continuous derivatives with respect to time. This implies that the defect $d = \tau - (a - bx - bx(t - \tau))$ in the algebraic condition (56) satisfies $\dot{d}(t) = -d(t)$ along solutions. Since the solutions are periodic the defect d is periodic, too, and, hence, d is identically zero. The denominator appearing in Equation (57) becomes zero exactly in those points in which the implicit condition (56) cannot be solved for the delay τ with a regular derivative.

The same argument can be applied to the position control problem as long as the object, at position x, does not hit the base at position $-w$ (the model contains the term $|x + w|$ in the right-hand side).

Neutral equations

The index reduction works only if the delay τ, which is itself a function of time, is not evaluated at different time points. For example, changing $bx(t - \tau(t))$ to $bx(t - \tau(t - 1))$ on the right-hand-side of (56) would make the index reduction impossible. However, certain simple neutral equations permit a similar reduction directly on the function space level. Consider

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\Delta_t(x + g(x))\right] = f(\Delta_t x) \tag{58}$$

where the functional f satisfies the local EC Lipschitz condition 2.2 in a neighborhood U of a point $x_0 \in \mathbb{C}^{1,1}(\mathbb{T}; \mathbb{R}^n)$, and $g : C^0(\mathbb{T}; \mathbb{R}^n) \mapsto \mathbb{R}^n$ has a global (classical) Lipschitz constant less than unity (this excludes state-dependent delays in the essential part of the neutral equation). Then one can define the map $X_g(y)$ as the unique solution x of the fixed point problem

$$x(t) = y(t) - g(\Delta_t x) \quad \text{for } y \text{ near } y_0 = x_0 + g(x_0),$$

which reduces (58) to the equation

$$\dot{y}(t) = f(\Delta_t X_g(y)) = f(X_g(\Delta_t y)). \tag{59}$$

Equation (59) satisfies the conditions of Theorem 5.1. One implication of this reduction is that periodic solutions of (58) are k times continuously differentiable if the functional f is EC^k smooth in the sense of Definition 2.1 and g is k times continuously differentiable as a map from $C^0(\mathbb{T}; \mathbb{R}^n)$ into $\mathbb{R}^n$.

References

[1] E. Allgower and M. K. Georgi. *Introduction to Numerical Continuation Methods*. SIAM, 1979.

[2] O. Arino, K. P. Hadeler, and M. L. Hbid. Existence of periodic solutions for delay differential equations with state dependent delay. *Journal of Differential Equations*, 144(2):263 – 301, 1998.

[3] Odo Diekmann, Stephan A. van Gils, Sjoerd M. Verduyn Lunel, and Hans-Otto Walther. *Delay equations*, volume 110 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1995. Functional, complex, and nonlinear analysis.

[4] Markus Eichmann. *A local Hopf Bifurcation Theorem for differential equations with state-dependent delays*. PhD thesis, University of Giessen, 2006.

[5] Jack K. Hale and Sjoerd M. Verduyn Lunel. *Introduction to functional-differential equations*, volume 99 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1993.

[6] Ferenc Hartung, Tibor Krisztin, Hans-Otto Walther, and Jianhong Wu. Functional differential equations with state-dependent delays: Theory and applications. In P. Drábek A. Cañada and A. Fonda, editors, *Handbook of Differential Equations: Ordinary Differential Equations*, volume 3, chapter 5, pages 435 – 545. North-Holland, 2006.

[7] Qingwen Hu and Jianhong Wu. Global hopf bifurcation for differential equations with state-dependent delay. *Journal of Differential Equations*, 248(12):2801 – 2840, 2010.

[8] T. Insperger, G. Stépán, and J. Turi. State-dependent delay model for regenerative cutting processes. In *Proceedings of the Fifth EUROMECH Nonlinear Dynamics Conference, Eindhoven, Netherlands*, 2005.

[9] D. Jackson. *The theory of Approximation*, volume XI. AMS Colloquium Publication, New York, 1930.

[10] J. Sieber and R. Szalai. Characteristic matrices for linear periodic delay differential equations. Technical report, http://arxiv.org/abs/1005.4522, 2010.

[11] G. Stépán. *Retarded Dynamical Systems*: *Stability and Characteristic Functions*. Longman Scientific and Technical, Harlow, Essex, 1989.

[12] R. Szalai, G. Stépán, and S.J. Hogan. Continuation of bifurcations in periodic delay differential equations using characteristic matrices. *SIAM Journal on Scientific Computing*, 28(4):1301–1317, 2006.

[13] H.-O. Walther. Stable periodic motion of a system with state-dependent delay. *Differential and Integral Equations*, 15:923–944, 2002.

[14] H.-O. Walther. Smoothness properties of semiflows for differential equations with state-dependent delays. *Journal of Mathematical Sciences*, 124(4):5193–5207, 2004.

A. Proof of Lemma 3.1

We find the norm $\|Q_N L\|_0$ first, and start out with the well-known estimate for interpolating trigonometric polynomials for continuous functions on $\mathbb{T}$. Let x be a continuous function on $\mathbb{T}$ with modulus of continuity ω. Then (see [9])

$$\|Q_N x\|_0 \le C_0 \omega\left(\frac{2\pi}{N}\right) \log N$$

where C_0 is a constant that does not depend on x or N. A function $\omega : [0,\infty) \mapsto [0,\infty)$ is called a modulus of continuity of a continuous function x if

$$|x(t) - x(s)| \le \omega(|t-s|).$$

holds for all s and $t \in \mathbb{T}$. For a function $x \in C(\mathbb{T};\mathbb{R}^n)$ the anti-derivative

$$[Lx](t) = \int_0^t x(s) - R_0 x \, \mathrm{d}s$$

has the Lipschitz constant $\|x\|_0 = \max\{|x(t)| : t \in \mathbb{T}\}$ such that a modulus of continuity for Lx is $\omega(h) = \|x\|_0 h$. Consequently,

$$\|Q_N Lx\|_0 \le C_0 \frac{2\pi \|x\|_0}{N} \log N, \tag{60}$$

where C_0 does not depend on x or N. This proves the claim of the lemma for $j = 0$. For $x \in C^j(\mathbb{T};\mathbb{R}^n)$ we notice that all derivatives of x up to order j are continuous. Applying estimate (60) to each of the derivatives of x we get

$$\|Q_N L x^{(l)}\|_0 \le \frac{2\pi C_0}{N} \log N \, \|x^{(l)}\|_0 \quad \text{for } l = 0, \ldots j.$$

Consequently, the maximum of the left-hand sides over all $l \in \{0, \ldots, j\}$ must be less than the maximum of the right-hand sides:

$$\|Q_N L x\|_j = \max_{l=0,\ldots,j} \|Q_N L x^{(l)}\|_0 \le \frac{2\pi C_0}{N} \log N \max_{l=0,\ldots,j} \|x^{(l)}\|_0 = 2\pi C_0 \frac{\log N}{N} \|x\|_j,$$

which implies the desired estimate for $\|Q_N L\|_j$ using the constant $C = 2\pi C_0$.

The estimate of $Q_N L$ in the Lipschitz norm is a continuity argument. The operator $Q_N L$ is bounded (and, thus, continuous) on $C^{0,1}(\mathbb{T};\mathbb{R}^n)$. For every element y of $C^1(\mathbb{T};\mathbb{R}^n)$ (which is a dense subspace of $C^{0,1}(\mathbb{T};\mathbb{R}^n)$) the Lipschitz constant is identical to $\|y'\|_0 = \max_{t\in\mathbb{T}}|y'(t)|$, and, thus, $\|y\|_1 = \|y\|_{0,1}$. Let $x_n \in C^1(\mathbb{T};\mathbb{R}^n)$ be a sequence of continuously differentiable functions that converges to $x \in C^{0,1}(\mathbb{T};\mathbb{R}^n)$ in the $\|\cdot\|_{0,1}$-norm: $\|x_n - x\|_{0,1} \to 0$ for $n \to \infty$. Then

$$\|Q_N L x_n\|_{0,1} = \|Q_N L x_n\|_1 \leq C\frac{\log N}{N}\|x_n\|_1 = C\frac{\log N}{N}\|x_n\|_{0,1}.$$

On both sides of the inequality the limit for $n \to \infty$ exists, resulting in the desired estimate for $\|Q_N L\|_{0,1}$. □

B. Basic differentiability properties of the right-hand side

Let $(D, \|\cdot\|_D)$ be a Banach space of the form

$$D = C^{k_1}(\mathbb{T};\mathbb{R}^{m_1}) \times \ldots \times C^{k_l}(\mathbb{T};\mathbb{R}^{m_l})$$

where $l \geq 1$, the integers k_j are non-negative and the integers m_j are positive. We use the notation

$$\begin{aligned} D^k &= \{x \in D : x^{(k)} \in D\}, \quad \text{with the norm} & \|x\|_{D,k} &= \max_{j=0\ldots k}\{\|x^{(j)}\|_D\} \\ D^{0,1} &= \left\{x \in D : L(x) := \sup_{t\neq s}\frac{\|x(t)-x(s)\|_D}{|t-s|} < \infty\right\}, & \|x\|_{D,L} &= \max\left\{\|x\|_D, L(x)\right\}. \end{aligned}$$

Sets that are closed and bounded in $D^{0,1}$ are complete with respect to the norm of D. We aasume that f is EC^1 smooth in the sense of Definition 2.1 in this Appendix.

B.1. Basic properties of f

This section proves three properties that continuously differentiable functions f have: first that the derivative limit (6) exists also for Lipschitz continuous deviations, second a weakened form of the mean value theorem, and third a weakened form of Lipschitz continuity.

Lemma B.1 (Extension of derivative to deviations in $D^{0,1}$)

> Let f be EC^1 smooth in the sense of Definition 2.1. Then the limit required to exist in Definition(2.1) exists also in the $\|\cdot\|_{D,L}$-norm: for all $x \in D^1$
>
> $$\lim_{\substack{y\in D^{0,1} \\ \|y\|_{D,L}\to 0}} \frac{|f(x+y) - f(x) - \partial_1 f(x)y|}{\|y\|_{D,L}} = 0. \tag{61}$$

Note that in (61) the norm in which y goes to zero is $\|\cdot\|_{D,L}$ instead of $\|\cdot\|_{D,1}$.

Proof Let $\epsilon > 0$ be arbitrary. We pick $\delta > 0$ such that

$$|f(x+\tilde{y}) - f(x) - \partial_1 f(x)\,\tilde{y}| < \epsilon \|\tilde{y}\|_{D,1} \tag{62}$$

for all $\tilde{y} \in D^1$ satisfying $\|\tilde{y}\|_{D,1} < \delta$. Let $y \in D^{0,1}$ be such that $\|y\|_{D,L} < \delta$. We can choose a $\tilde{y} \in D^1$ that satisfies

$$\|\tilde{y}\|_{D,1} < \min\{\delta, 2\|y\|_{D,L}\} \tag{63}$$

$$|f(x+y) - f(x+\tilde{y})| < \epsilon \|y\|_{D,L} \tag{64}$$

$$|\partial_1 f(x)[y - \tilde{y}]| < \epsilon \|y\|_{D,L}. \tag{65}$$

Combining estimate (62) with (63)–(65) we obtain

$$|f(x+y) - f(x) - \partial_1 f(x)\,y| < 4\epsilon \|y\|_{D,L}.$$

□

Lemma B.2 (Existence of mean value)

> There exists a continuous function
>
> $$\tilde{a} : D^1 \times D^1 \times D \mapsto \mathbb{R}^n$$
>
> which is linear in its third argument and satisfies for all $x, y \in D^1$
>
> $$f(x+y) - f(x) = \tilde{a}(x,y)\,y. \tag{66}$$
>
> Moreover, $\tilde{a}(x,0)y = \partial_1 f(x)y$ for all $x \in D^1$ and $y \in D$.

Proof The argument for the existence of a mean value follows exactly the proof of the general mean value theorem [6]: the candidate for $\tilde{a}(u,v)\,w$ is

$$\tilde{a}(u,v)\,w = \int_0^1 \partial_1 f(u+sv)\,w\,\mathrm{d}s. \tag{67}$$

Since $\partial_1 f$ is assumed to be continuous in its arguments the integral is well defined and continuous in its arguments $u \in D^1$, $v \in D^1$, $w \in D$. All one has to show is that the $\tilde{a}$ defined in (67) satisfies (66): let $x, y \in D^1$ and $\epsilon > 0$ be arbitrary and choose m such that uniformly for all $s \in [0,1]$

$$\left| \int_0^1 \partial_1 f(x+sy)\,y\,\mathrm{d}s - \frac{1}{m}\sum_{k=0}^{m-1} \partial_1 f\left(x + \frac{k}{m}y\right) y \right| < \epsilon,$$

$$\left| f\left(x + sy + \frac{y}{m}\right) - f(x+sy) - \partial_1 f(x+sy)\frac{y}{m} \right| < \frac{\epsilon}{m}.$$

Then it follows that

$$\left| f(x+y) - f(x) - \int_0^1 \partial_1 f(x+sy)\,y\,\mathrm{d}s \right| < 2\epsilon.$$

Since $\epsilon > 0$ was arbitrary the left-hand side must be zero. □

Lemma B.3 (Local C^0 Lipschitz continuity)

> For all $x \in D^{0,1}$ there exists a neighborhood $U(x) \subseteq D^{0,1}$ and a constant $K_x > 0$ such that for all y_1 and $y_2 \in U(x)$ the following Lipschitz estimate holds:
>
> $$|f(y_1) - f(y_2)| \leq K_x \|y_1 - y_2\|_D.$$

Note that the upper bound depends only on the $\|\cdot\|_D$-norm, not on the $\|\cdot\|_{D,L}$-norm, which would be a weaker statement.

Proof We prove the Lipschitz continuity first for y_1 and y_2 from a sufficiently small neighborhood $U(x) \cap D^1 \subseteq D^1$ of $x \in D^1$.

Let x be an element of D^1. Since the mean value $\tilde{a}$ is continuous in $(x, 0, 0)$ we have a $\delta > 0$ such that for all $u, v \in D^1$ and $w \in D$ satisfying $\|u\|_{D,1} < \delta$, $\|v\|_{D,1} < \delta$ and $\|w\|_D < \delta$

$$|\tilde{a}(x+u, v)w| < \epsilon.$$

This implies that $|\tilde{a}(x+u, v)w| < [\epsilon/\delta]\|w\|_D$ for all u and v with $\max\{\|u\|_{D,1}, \|v\|_{D,1}\} < \delta$ and $w \in D$. Thus, $\|\tilde{a}(x+u, v)\|_D \leq \epsilon/\delta =: K_x$ for $\tilde{a}(x+u, v)$ as an element of $L(D; D)$ in the operator norm corresponding to D. Consequently, if $\|y_1 - x\|_{D,1} < \delta/2$ and $\|y_2 - x\|_{D,1} < \delta/2$

$$|f(y_1) - f(y_2)| = \left| \int_0^1 \tilde{a}(y_2, y_1 - y_2)[y_2 - y_1] \, \mathrm{d}s \right| \leq K_x \|y_1 - y_2\|_D.$$

The extension of the statement to $D^{0,1}$ follows from the continuity of f in D: $U(x_0) \cap D^1$ is dense in $U(x_0) \subset D^{0,1}$. Pick two sequences y_n and z_n in $U(x_0) \cap D^1$ that converge to y and z in $U(x_0)$ in the Lipschitz norm. Then $f(y_n) \to f(y)$ and $f(z_n) \to f(z)$ since f is continuous in D. Moreover, $\|y_n - z_n\|_D \to \|y - z\|_D$ for $n \to \infty$. Since

$$|f(y_n) - f(z_n)| \leq K_x \|y_n - z_n\|_D \tag{68}$$

for all n the inequality also holds for the limit for $n \to \infty$. □

B.2. Basic properties of F

Let $F : D \mapsto C^0(\mathbb{T}; \mathbb{R}^n)$ be defined as $F(x)(t) = f(\Delta_t x)$. As a map into $C^0(\mathbb{T}; \mathbb{R}^n)$ the map F is continuous because f and Δ_t are continuous. The first lemma lists properties that F has if f satisfies only local EC Lipschitz continuity in the sense of Definition 2.2. That is, we do *not* assume that f is EC^1 smooth in the sense of Definition 2.1. Since Lemma B.3 was proved using only the assumption of EC^1 smoothness of f this is a strictly weaker condition.

Lemma B.4 (C^0 Lipschitz continuity of F)

> Assume that f is locally EC Lipschitz continuous in the sense of Definition 2.2. Then F has the following properties:

1. for all $x \in D^{0,1}$ there exists a neighborhood $U(x) \subseteq D^{0,1}$ and a constant $K_x > 0$ such that for all y_1 and $y_2 \in U(x)$

$$\|F(y_1) - F(y_2)\|_0 \leq K_x \|y_1 - y_2\|_D.$$

2. F maps elements of $D^{0,1}$ into $C^{0,1}(\mathbb{T}; \mathbb{R}^n)$. Moreover, for every $x \in D^{0,1}$ there exists a bounded neighborhood of $U(x) \subseteq D^{0,1}$ such that the image of $U(x)$ under F is bounded: there exists a $R > 0$ such that $\|F(y)\|_{0,1} \leq R$ for all $y \in U(x)$ (R depends on $U(x)$).

Proof Statement 1 is a simple consequence of the local EC Lipschitz continuity of f and the compactness of $\mathbb{T}$ (which allows to choose a uniform Lipschitz bound K_x for all $t \in \mathbb{T}$).

Concerning statement 2: let $x \in D^{0,1}$ be arbitrary. We choose the neighborhood $U(x)$ bounded (say, $\|y - x\|_{D,L} \leq \delta$) and such that we can apply the local EC Lipschitz continuity of f in $U(x)$. Then there exists a K_x such that for all $y, z \in U(x)$ and $t, s \in \mathbb{T}$ the estimate

$$|f(\Delta_t y) - f(\Delta_s z)| \leq K_x \|\Delta_t y - \Delta_s z\|_D = K_x \|\Delta_{t-s} y - z\|_D$$

holds. Initially setting $z = x$ and $s = t$ we get a bound on $\|F(y)\|_0$: $\|F(y)\|_0 \leq \|F(x)\|_0 + K_x \delta =: R_0$ for all $y \in U(x)$. It remains to be shown that the Lipschitz constant of $F(y)$ is bounded for $y \in U(x)$:

$$|F(y)(t) - F(y)(s)| = |f(\Delta_t y) - f(\Delta_s y)| \leq K_x \|\Delta_t y - \Delta_s y\|_D \leq K_x \|y\|_{D,L} |t - s|.$$

Since $\|y - x\|_{D,L} \leq \delta$ for $y \in U(x)$, choosing

$$R = \max\left\{R_0, K_x \left(\|x\|_{D,L} + \delta\right)\right\}$$

ensures that $\|F(y)\|_{0,1} \leq R$. ∎

Define the maps

$$\begin{aligned}
[\partial_1 F(u) v](t) &= \partial_1 f(\Delta_t u) \Delta_t v && \text{for } u \in D^1,\ v \in D, \\
[\tilde{A}(u, v) w](t) &= \tilde{a}(\Delta_t u, \Delta_t v) \Delta_t w && \text{for } u \in D^1,\ v \in D^1,\ w \in D.
\end{aligned}$$

The following Corollary B.5 and Lemma B.6 assume that f is EC^1 smooth in D in the sense of Definition 2.1.

Corollary B.5 (Differentiability of F)

Let $f : D \mapsto \mathbb{R}^n$ be EC^1 smooth. Then F, $\partial_1 F$ and $\tilde{A}$ have the following properties.

1. The map $\partial_1 F(u) v)$ is continuous in both arguments (and linear in its second argument) as a map from $D^1 \times D$ into $C^0(\mathbb{T}; \mathbb{R}^n)$. It satisfies for all $x \in D^1$

$$\lim_{\substack{y \in D^{0,1} \\ \|y\|_{D,L} \to 0}} \frac{\|F(x + y) - F(x) - \partial_1 F(x) y\|_0}{\|y\|_{D,L}} = 0. \tag{69}$$

2. The map $\tilde{A}(u,v)w$ is continuous in all three arguments (and linear in its third argument) as a map from $D^1 \times D^1 \times D$ into $C^0(\mathbb{T};\mathbb{R}^n)$. It satisfies for all $x, y \in D^1$
$$F(x+y) - F(x) = \tilde{A}_1(x,y)\,y.$$
Moreover, $\tilde{A}(x,0)\,y = \partial_1 F(x)\,y$ for all $x \in D^1$ and $y \in D$.

Note that in the limit (69) we allow for deviations $y \in D^{0,1}$.

Proof The statements 1 and 2 are direct consequences of the EC^1 smoothness of f, the compactness of $\mathbb{T}$, Lemma B.1 and Lemma B.2, the corresponding lemmas about $D^{0,1}$ deviations and the mean value for f. □

Lemma B.6 (Differentiability of images of F)

Let $f : D \mapsto \mathbb{R}^n$ be EC^1 smooth. We also assume that $F : D \mapsto C^k(\mathbb{T};\mathbb{R}^n)$ (for some $k \geq 0$) and $\partial_1 F : D^1 \times D \mapsto C^k(\mathbb{T};\mathbb{R}^n)$ are continuous maps. Then F maps elements of D^1 into $C^{k+1}(\mathbb{T};\mathbb{R}^n)$ and F is continuous as a map from D^1 to $C^{k+1}(\mathbb{T};\mathbb{R}^n)$.

Proof Let x be in D^1, that is, $x' \in D$. If $\partial_1 F : D^1 \times D \mapsto C^k(\mathbb{T};\mathbb{R}^n)$ is continuous then $\tilde{A} : D^1 \times D^1 \times D \mapsto C^k(\mathbb{T};\mathbb{R}^n)$, which is given by $\tilde{A}(u,v)\,w = \int_0^1 \partial_1 F(u+sv)\,w\,\mathrm{d}s$, is continuous, too. Using statement 2 of Corollary B.5 we have
$$\frac{F(\Delta_h x) - F(x)}{h} = \tilde{A}(x, \Delta_h x - x)\left[\frac{\Delta_h x - x}{h}\right].$$
On the right side $\|\Delta_h x - x\|_{D,1}$ converges to 0 and $[\Delta_h x - x]/h$ converges to x' in the norm of D for $h \to 0$ because $x \in D^1$. Since $\tilde{A}$ is continuous in its arguments the right side converges to $\tilde{A}(x,0)x' = \partial_1 F(x)x' \in C^k(\mathbb{T};\mathbb{R}^n)$ for $h \to 0$. Thus, $F(x) \in C^{k+1}(\mathbb{T};\mathbb{R}^n)$. Since $(v,w) \in D^1 \times D \mapsto \partial_1 F(v)\,w$ is continuous in (u,v), the time derivative of F, $F(x)' = \partial_1 F(x)x'$ is also continuous in x if we use the norm $\|\cdot\|_{D,1}$ for the argument and $\|\cdot\|_k$ for the image. □

C. Proof of Hopf bifurcation theorem

First, we note that $x \mapsto S(x,\omega)^{-1} = x(\omega^{-1}\cdot)$ maps $C^k(\mathbb{T};\mathbb{R}^n)$ into a closed subspace of $C^k([-\tau,0];\mathbb{R}^n)$ if we choose $\tau > 2\pi/\omega_0$ (which we can do without loss of generality), and if we extend functions x on $\mathbb{T}$ to the whole real line by setting $x(t) = x(t \mod [-\pi,\pi))$. This implies that, if the functional $f : C^0([-\tau,0];\mathbb{R}^n) \times \mathbb{R} \mapsto \mathbb{R}^n$ is EC^k smooth then the functional
$$(x,\mu,\omega) \in C^0(\mathbb{T};\mathbb{R}^n) \times \mathbb{R}^2 \mapsto \frac{1}{\omega} f(S(x,\omega),\mu)$$
is EC^k smooth, too, such that we can reduce the problem of finding periodic orbits of frequency ω to the algebraic system (51). Let us choose the periodic orbit $x_0 = (x,\omega,\mu)$

with $x=0$, $\omega=\omega_0$, $\mu=0$ as the solution in the neighborhood of which we construct the equivalent algebraic system. We choose the number of Fourier modes, $N\geq 1$, and the size of the neighborhood B in $C^{0,1}(\mathbb{T};\mathbb{R}^{n+2})$ such that the conditions of Theorem 5.1 are satisfied in B. The nonlinear algebraic system then reads (after multiplication with ω and mapping it onto the space $\operatorname{rg} P_N$ from $\mathbb{R}^{n(2N+1)}$ by applying R_N^{-1})

$$\begin{aligned}0=&P_0 f(S(\Delta_t X_p(p,\omega,\mu),\omega),\mu)+\\&+\omega Q_0 P_N E_N p - Q_0 P_N L f(S(\Delta_t X_p(p,\omega,\mu),\omega),\mu)\end{aligned}\tag{70}$$

The variables are $p\in\mathbb{R}^{n(2N+1)}$, μ and ω. We know from Theorem 7.5 that

$$\begin{aligned}Y_p:&(p,\omega,\mu)\in\mathbb{R}^{n(2N+1)}\times\mathbb{R}\times\mathbb{R}\mapsto f(S(\Delta_t X_p(p,\omega,\mu),\omega),\mu)\in C^0(\mathbb{T};\mathbb{R}^n),\\X_p:&(p,\mu,\omega)\in\mathbb{R}^{n(2N+1)}\times\mathbb{R}\times\mathbb{R}\mapsto X_p(p,\omega,\mu)\in C^1(\mathbb{T};\mathbb{R}^n)\end{aligned}$$

are k times differentiable and note that

$$f(S(\Delta_t X_p(0,\omega,\mu),\omega),\mu)=0\tag{71}$$

for all $\omega\approx\omega_0$ and $\mu\approx 0$. The derivative of

$$F:(x,\omega,\mu)\mapsto f(S(\Delta_t x,\omega),\mu)$$

in $x=0$, $\omega\approx\omega_0$ and $\mu\approx 0$ with respect to x is

$$a(\mu)S(\Delta_t x,\omega),\quad\text{which we write as } A(\omega,\mu)\,x$$

where $a(\mu)$ is the same linear functional as used in the definition of the characteristic matrix $K(\lambda,\mu)$ in (53) (the derivatives of F with respect to ω and μ are zero due to (71)). We observe that $A(\omega,\mu)$ commutes with P_j and Q_j for all $j\geq 0$.

Let us now determine the linearization of $X_p(p,\omega,\mu)$ in $(p,\omega,\mu)=(0,\omega,\mu)$: due to (71) $X_p(0,\omega,\mu)$ is equal to zero for all $\omega\approx\omega_0$ and $\mu\approx 0$. Thus, we have

$$0=\left.\frac{\partial}{\partial\omega}X_p(p,\omega,\mu)\right|_{p=0}\qquad\text{and}\qquad 0=\left.\frac{\partial}{\partial\mu}X_p(p,\omega,\mu)\right|_{p=0}.$$

Moreover, the fixed point equation (40) defining $z=[\partial X_p/\partial p](p,\omega,\mu)q$, evaluated in $(p,\omega,\mu)=(0,\omega,\mu)$ reads

$$z=E_N q+Q_N LA(\mu,\omega)z=E_N q+Q_N LA(\mu,\omega)Q_N z,$$

exploiting that $Q_N L=Q_N L Q_N$ and $Q_N A(\omega,\mu)=A(\omega,\mu)Q_N$. We can choose the neighborhood of $(p,\omega,\mu)=(0,\omega_0,0)$ such that the spectral radius of $Q_N LA(\mu,\omega)$ is less than unity (see Corollary 6.3), which implies that $Q_N z=0$, and, thus

$$\left[\left.\frac{\partial}{\partial p}X_p(p,\omega,\mu)\right|_{p=0}\right]q=E_N q.$$

Consequently, the linearization of the nonlinear algebraic system (70) in $(p, \omega, \mu) = (0, \omega, \mu)$ is

$$0 = P_0 A(\omega, \mu) E_N p + \omega Q_0 P_N E_N p - Q_0 P_N L A(\omega, \mu) E_N p. \quad (72)$$

We observe that the linear system (72) decouples into equations for

$$\begin{aligned} y_0 &= P_o E_N p = E_0 p = p_0 && \text{(the average of } E_N p\text{)}, \\ y_1 &= Q_0 E_1 p = p_{-1} \sin t + p_1 \cos t && \text{(the first Fourier component of } E_N p\text{)}, \\ y_j &= Q_{j-1} E_j p = p_{-j} \sin(jt) + p_j \cos(jt) && \text{(the } j\text{-th Fourier component of } E_N p, \\ & && 2 \le j \le N), \end{aligned}$$

where we denote the components of p by $p_j \in \mathbb{R}^n$ ($j = -N \dots N$). This decoupling is achieved by pre-multiplication of (72) with P_0 and $Q_{j-1}P_j$ for $j = 1 \dots N$:

$$\begin{aligned} P_0 \cdot (72): &\quad 0 = A(\omega, \mu) y_0 = a(\mu)\, p_0 && (73) \\ Q_0 P_1 \cdot (72): &\quad 0 = \omega y_1 - Q_0 L A(\omega, \mu)\, y_1 && (74) \\ Q_{j-1} P_j \cdot (72): &\quad 0 = \omega y_j - Q_0 L A(\omega, \mu)\, y_j \qquad \text{for } j = 2 \dots N. && (75) \end{aligned}$$

Equations (74) and (75) are equivalent (in complex notation) to

$$\begin{aligned} i\omega u_1 - a(\mu) \left[u_1 \exp(i\omega s) \right] &= K(i\omega, \mu)\, u_1 = 0, && (76) \\ ij\omega u_j - a(\mu) \left[u_j \exp(ij\omega s) \right] &= K(ij\omega, \mu)\, u_j = 0 \quad (2 \le j \le N), && (77) \end{aligned}$$

that is, $u_j \in \mathbb{C}^n$ is a solution of (76) (or (77), respectively) if and only if $y_j = u_j \exp(ijt)$ is a solution of (74) (or (75), respectively).

The non-resonance assumption of the theorem guarantees that equation (73) is a regular linear system for p_0, and that (75) is a regular linear algebraic system for p_{-j} and p_j ($j \ge 2$) at $\mu = 0$ and $\omega = \omega_0$ (and, hence, for all ω and μ near-by). The condition on the simplicity of the eigenvalue $i\omega_0$ of K ensures that equation (76) (and, thus, (74)) has a one-dimensional (in complex notation) subspace of solutions for $\omega = \omega_0$ and $\mu = 0$, spanned by the nullvector v_1 of $K(i\omega, 0)$. Let us denote the adjoint nullvector of $K(i\omega_0, 0)$ by $w_1 \in \mathbb{C}^n$ (again, using complex notation, $w_1^H K(i\omega_0, 0) = 0$). Since $i\omega_0$ is simple

$$w_1^H \frac{\partial K}{\partial \lambda}(i\omega, 0)\, v_1 \neq 0,$$

which implies that we can choose $w_1 \in \mathbb{C}^n$ without loss of generality such that

$$w_1^H \frac{\partial K}{\partial \lambda}(i\omega, 0)\, v_1 = 1.$$

With this convention we note that

$$w_1^H \frac{\partial K}{\partial \mu}(i\omega, 0)\, v_1 = - \left. \frac{\partial \lambda}{\partial \mu} \right|_{\mu = 0} =: c_\mu \in \mathbb{C} \qquad w_1^H \frac{\partial}{\partial \omega} K(i\omega, 0)\, v_1 = i \in \mathbb{C} \quad (78)$$

where $\operatorname{Re} c_\mu \neq 0$ by the transversal crossing assumption of the theorem. In complex notation any scalar multiple of the nullvector $v_1 = v_r + i v_i$ is also a nullvector. Thus, the complex scalar factor $\alpha + i\beta$ in front of v_1 makes up two components of the variable p (in real notation): in short, p solves the linearized algebraic system (72) if and only if all p_j with $|j| \neq 1$ are zero and $p_{-1} \sin t + p_1 \cos t = \operatorname{Re}\left[(\alpha + i\beta) v_1 \exp(it)\right]$ for some $\alpha, \beta \in \mathbb{R}$, that is,

$$\begin{bmatrix} p_{-1} \\ p_1 \end{bmatrix} = \alpha \begin{bmatrix} -v_i \\ v_r \end{bmatrix} + \beta \begin{bmatrix} -v_r \\ -v_i \end{bmatrix} =: \alpha b_r + \beta b_i. \tag{79}$$

Let us collect the statements so far and introduce coordinates. Collecting all components p_j with $|j| \neq 1$ and the orthogonal complement in $\mathbb{R}^{2n}$ of the space spanned by $\{b_1, b_2\}$ into a single variable q (of real dimension $n_q = n(2N-1) + 2(n-1)$), the coordinates for the variables are

$$(\alpha, \beta) =: r \in \mathbb{R}^2, \quad \text{and} \quad q \in \mathbb{R}^{n_q}.$$

We split up the nonlinear system of equations (70) in the same way as we did for the linearized problem, by pre-multiplication with P_0 and $Q_{j-1}P_j$ for $j = 1 \ldots N$:

$$\begin{aligned} P_0 \cdot (70): \quad & 0 = P_0 f(S(\Delta_t X_p(p,\omega,\mu),\omega),\mu) && (80) \\ Q_0 P_1 \cdot (70): \quad & 0 = \omega Q_0 E_1 p - Q_0 P_1 L f(S(\Delta_t X_p(p,\omega,\mu),\omega),\mu) && (81) \\ Q_{j-1} P_j \cdot (70): \quad & 0 = \omega Q_{j-1} E_j p - Q_{j-1} P_j L f(S(\Delta_t X_p(p,\omega,\mu),\omega),\mu). && (82) \end{aligned}$$

We split equation (81) further using w_1^H and its orthogonal complement $w_1^\perp = I - w_1 w_1^H / (w_1^H w_1)$. This gives rise to a splitting into two real equations ($w_1^H \cdot$(81)) and $2(n-1)$ real equations ($w_1^\perp \cdot$(81)). Collecting $w_1^\perp \cdot$(81) and the equations (80) and (82) into a subsystem of $n(2N-1) + 2(n-1) = n_q$ equations the full nonlinear algebraic system (70) in the coordinates (r, q) has the form

$$0 = \begin{bmatrix} M_{rr}(r,q,\omega,\mu) & M_{rq}(r,q,\omega,\mu) \\ M_{qr}(r,q,\omega,\mu) & M_{qq}(r,q,\omega,\mu) \end{bmatrix} \begin{bmatrix} r \\ q \end{bmatrix}. \tag{83}$$

By our choice of coordinates the matrices $M_{rr} \in \mathbb{R}^{2\times 2}$, $M_{rq} \in \mathbb{R}^{2\times n_q}$ and $M_{qr} \in \mathbb{R}^{n_q \times 2}$ are identically zero in $r = 0$, $q = 0$, $\mu = 0$, $\omega = i\omega_0$ such that the system matrix has the form

$$\begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & \ldots & 0 \\ 0 & \ldots & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} & \begin{matrix} M_{qq}(0,0,i\omega,0) \\ \text{(regular)} \end{matrix} \end{bmatrix}$$

in $r = 0$, $q = 0$, $\mu = 0$ and $\omega = i\omega_0$. Thus, we can eliminate q by solving the n_q lower equations, obtaining a graph $q(r,\omega,\mu) r$. This graph respects rotational invariance: $q(\Delta_s r, \omega, \mu) \Delta_s r = \Delta_s [q(r,\omega,\mu) r]$. Note that the application of Δ_s to $r = (\alpha, \beta)$

corresponds to the rotation of r by angle s (or the multiplication $\exp(is)(\alpha+i\beta)$). The Lyapunov-Schmidt reduction of (83), replacing q by the graph $q(r,\omega,\mu)\,r$, then reads

$$0 = M_{rr}(r, q(r,\omega,\mu)\,r,\omega,\mu)\,r =: M_r(r,\omega,\mu)\,r, \tag{84}$$

where M_r is still rotationally symmetric in r: $M_r(\Delta_s r,\omega,\mu)\,\Delta_s r = \Delta_s M_r(r,\omega,\mu)\,r$. Equation (78) in real notation implies that

$$\begin{aligned}\frac{\partial M_r}{\partial\omega}(0,\omega_0,0) &= \frac{\partial M_{rr}}{\partial\omega}(0,0,\omega_0,0) = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix},\\ \frac{\partial M_r}{\partial\mu}(0,\omega_0,0) &= \frac{\partial M_{rr}}{\partial\mu}(0,0,\omega_0,0) = \begin{bmatrix}\operatorname{Re} c_\mu & -\operatorname{Im} c_\mu\\ \operatorname{Im} c_\mu & \operatorname{Re} c_\mu\end{bmatrix}.\end{aligned}$$

Equation (84) is a system of two equations with four unknowns ($r=(\alpha,\beta)$, ω and μ). We now fix one of the unknowns setting

$$\alpha = 0$$

such that we can expect one-parametric families of solutions (β,ω,μ). Introducing M_β as the second column of M_r and dropping the dependence on α (which is zero), the first derivatives of $M_\beta(\beta,\omega,\mu)$ in $(0,\omega_0,0)$ with respect to the pair ω and μ is:

$$\begin{bmatrix}\dfrac{\partial M_\beta}{\partial\omega} & \dfrac{\partial M_\beta}{\partial\mu}\end{bmatrix}(0,\omega_0,0) = \begin{bmatrix}-1 & -\operatorname{Im} c_\mu\\ 0 & \operatorname{Re} c_\mu\end{bmatrix},$$

which is regular (as established in (78), $\operatorname{Re} c_\mu \neq 0$ due to the assumption that the eigenvalue crosses the imaginary axis transversally). Note that M_β itself is a projection of the first derivative of the original right-hand side of the full nonlinear algebraic system (70). Thus, M_β is $k-1$ times continuously differentiable, and we end up with a nonlinear system of (two) equations for three scalar variables (β,ω,μ):

$$0 = M_\beta(\beta,\omega,\mu)\,\beta.$$

Hence, either $\beta=0$, which corresponds to the trivial solution or (after division by β)

$$0 = M_\beta(\beta,\omega,\mu), \tag{85}$$

where $M_\beta(0,\omega_0,0)=(0,0)$ and the derivative with respect to the pair (ω,μ) is regular. Thus, we can apply the implicit function theorem to (85) to obtain a unique graph $(\omega(\beta),\mu(\beta))$ solving (85). The graph satisfies $\omega(0),\mu(0))=(\omega_0,0)$ and thus, branches off from the trivial solution (which has $\beta=0$ and ω and μ arbitrary). The rotational symmetry of M_r implies reflection symmetry of M_β in β such that $M_\beta(-\beta,\omega,\mu)=M_\beta(\beta,\omega,\mu)$ for all β, ω and μ. Hence, the solution graph is reflection symmetric, too: $\omega(-\beta)=\omega(\beta)$ and $\mu(-\beta)=\mu(\beta)$. Thus, for small β there is a unique non-trivial solution of the full nonlinear algebraic system of the form $r=(0,\beta)$, $q=q(r,\omega(\beta),\mu(\beta))\,r$. As Equation (79) shows, we can extract the coordinates α (which is zero) and β from the full solution $x \in C^k(\mathbb{T};\mathbb{R}^n)$ by applying the projections

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \cos(t)v_r^T x(t) - \sin(t)v_i^T x(t)\,\mathrm{d}t = \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{Re}\left[v_1 \exp(it)\right]^T x(t)\,\mathrm{d}t = \alpha,$$
$$\frac{1}{\pi}\int_{-\pi}^{\pi} \sin(t)v_r^T x(t) + \cos(t)v_i^T x(t)\,\mathrm{d}t = \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{Im}\left[v_1 \exp(it)\right]^T x(t)\,\mathrm{d}t = -\beta.$$

(Note that the vector $v_1 = v_r + v_i$ was scaled to have unit length and that the decomposition was orthogonal.) Thus, the coordinate β in the statement of the theorem is the negative of the coordinate β chosen in (79). $\square$