

## Computation of Limit Cycles and Their Isochrons: Fast Algorithms and Their Convergence\*

Gemma Huguet<sup>†</sup> and Rafael de la Llave<sup>‡</sup>

**Abstract.** We present efficient algorithms to compute limit cycles and their isochrons (i.e., the sets of points with the same asymptotic phase) for planar vector fields. We formulate a functional equation for the parameterization of the invariant cycle and its isochrons, and we show that it can be solved by means of a Newton method. Using the right transformations, we can solve the equation of the Newton step efficiently. The algorithms are efficient in the sense that if we discretize the functions using  $N$  points, a Newton step requires  $O(N)$  storage and  $O(N \log N)$  operations in Fourier discretization or  $O(N)$  operations in other discretizations. We prove convergence of the algorithms and present a validation theorem in an a posteriori format. That is, we show that if there is an approximate solution of the invariance equation that satisfies some mild nondegeneracy conditions, then there is a true solution nearby. Thus, our main theorem can be used to validate numerically computed solutions. The theorem also shows that the isochrons are analytic and depend analytically on the base point. Moreover, it establishes smooth dependence of the solutions on parameters and provides efficient algorithms to compute perturbative expansions with respect to external parameters. We include a discussion on the numerical implementation of the algorithms as well as numerical results for representative examples.

**Key words.** parameterization method, isochrons, invariant manifolds, numerical computation of invariant objects, phase resetting curves, biological oscillators

**AMS subject classifications.** 37M99, 65P40, 92-08, 37C10, 37C27, 34C45

**DOI.** 10.1137/120901210

**1. Introduction.** Oscillators appear in many areas of biology as well as in physics and chemistry. Mathematically, we regard them as systems that have a stable limit cycle, whose dynamics on the limit cycle can be easily described by their phase. The phase can be extended to a neighborhood of the limit cycle via the asymptotic phase. The sets of points with a constant asymptotic phase, called isochrons [47], approach the same orbit on the limit cycle. Isochrons play a key role in understanding changes in phase due to brief perturbations. A brief stimulus occurring at phase  $\theta$  will send the trajectory away from the limit cycle and put it on the isochron for another phase, say  $\theta'$ . Changes in phase due to a brief stimulus at different times of the cycle are typically described by the so-called phase response curve (PRC). Phase

\*Received by the editors December 4, 2012; accepted for publication (in revised form) by D. Barkley July 10, 2013; published electronically October 22, 2013. A preliminary version of this paper was prepared at IMA, during the New Directions Summer Program 2011.

<http://www.siam.org/journals/siads/12-4/90121.html>

<sup>†</sup>Courant Institute of Mathematical Sciences, New York University, New York, NY 10012-1185. Current address: Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, 08028 Barcelona, Spain ([gemma.huguet@upc.edu](mailto:gemma.huguet@upc.edu)). The work of this author was supported by the Swartz Foundation, Spanish MCyT/FEDER grants MTM2009-06973 and MTM2012-31714, and Catalan 2009SGR-859.

<sup>‡</sup>School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332 ([rafael.delallave@math.gatech.edu](mailto:rafael.delallave@math.gatech.edu)). The work of this author was supported by NSF grant DMS 1162544.

reduction methods and PRCs have been extensively used to study weakly coupled oscillators and synchronization properties [21, 11, 45, 12, 17]. However, when these perturbations are not weak or occur very frequently, the trajectory is displaced away from the limit cycle and the phase reduction and/or the linear PRC are not enough to accurately describe the dynamics [19, 46, 27].

Thus, in recent times, there has been a large effort to compute isochrons and PRCs accurately and up to high order. Recent papers close to our goal are [28], which computes isochrons for dimensions higher than two; [43], which presents algorithms for PRCs to all orders; [41], which describes and implements algorithms for isochrons of bursting oscillations, and [33], which describes and implements algorithms for the computation of isochrons which are effective for systems with multiple time scales. We especially refer the reader to [19], where a parameterization method was used to compute local isochrons up to high order and PRCs for points which are not on the limit cycle. In that paper, the invariance equation was solved using an order by order method, which involved solving a large linear system.

In this paper we describe efficient algorithms (which do not require solving a full linear system) to compute a parameterization of the limit cycle and the isochrons that provides extremely accurate representations of the isochrons in a relatively large neighborhood of the limit cycle. The core mathematical result of this paper is Theorem 3.2, which provides estimates for the convergence of the algorithm. In addition to this, Theorem 3.2 shows that the isochrons are analytic and depend analytically on the base point, establishes smooth dependence of the solutions on parameters, and provides efficient algorithms to compute perturbative expansions with respect to external parameters.

The starting point of the mathematical formulation of the problem is the work by Guckenheimer in [18], which pointed out that the isochrons are just the stable manifold of a point in the limit cycle in the sense of the theory of normally hyperbolic manifolds. Hence, following [3, 4, 5], we formulate the problem as solving a functional equation (see (2.4)) for the parameterization of the invariant circle and its isochrons. This method was also used in [19], where the invariance equation was solved using an order by order method. The main novelty here is that we apply a Newton method to solve this functional equation. The key point for obtaining both an efficient algorithm and good estimates that lead to convergence is to use several identities and algebraic manipulations to transform the equation for the Newton step into a simpler equation.

These identities, which have a geometric interpretation, are obtained by taking derivatives of the invariance equation. As it turns out, the Newton step involves some “loss of derivatives”; i.e., the norm of the remainder after a Newton step is controlled by the square of the norm of derivatives of the remainder before the Newton step. It is well known from [24, 31, 30] that the quadratic estimates in one step lead to convergence even when there is loss of derivatives. Of course, numerical implementations exhibit quadratic convergence irrespective of the fact that to establish it one uses a sophisticated method.

The algorithms we present here can be implemented rather straightforwardly using a package manipulating Fourier–Taylor series of the type that is commonly used in celestial mechanics [2, 20, 16]. The algorithm is highly efficient in the sense that if we discretize the function using  $N$  terms of the Fourier–Taylor series, then a Newton step requires  $O(N)$  storage and  $O(N \log N)$  operations but has the quadratic convergence of the Newton algorithm (i.e.,

after an application of the algorithm, the error is roughly the square of the original error). If we discretize using splines or using collocation methods, using  $N$  points, a Newton step requires  $O(N)$  storage and  $O(N)$  operations (with a larger constant).

We have implemented these algorithms using Fourier–Taylor series, and we show the results for some representative examples in section 8 (the Rayleigh oscillator, the Morris–Lecar model, and a reduced Hodgkin–Huxley-type model). The algorithm is highly accurate (in the examples considered, one can get the isochrons up to ten thousand times the roundoff error), and it also gives information on the derivatives of the isochrons. The computations run in seconds on a standard laptop.

This paper is organized as follows: In section 2 we formulate the problem as a functional equation. In section 3 we state the main result: Theorem 3.2. In section 4 we describe the iterative step of the Newton method. In particular, in section 4.4 we present the iterative step in an algorithmic form (Algorithm 4.4). This iterative step will be used in section 5 to prove that the method converges in some appropriate norms defined in section 3.1. In section 6 we present some consequences of the formalism, such as smooth dependence on parameters and local uniqueness. In section 7 we discuss some implementation details, and in section 8 we present numerical results for some representative examples. We end with a discussion in section 9.

Note that the tools and standards of the sections of this paper are very different and could appeal to different communities. Hence, we have strived to make the specific sections readable independently. For instance, in sections 4.4, 7, and 8 we use the language of algorithm theory, and discuss storage, operation counts, etc. In contrast, in sections 5 and 6 we use rigorous mathematical estimates.

**2. Setup of the problem.** We consider a differential equation in the plane

$$(2.1) \quad \dot{x} = X(x), \quad x \in \mathbb{R}^2,$$

and denote by  $X^t$  the flow associated with (2.1). That is,  $X^t(x_0)$  solves  $\frac{d}{dt}X^t(x_0) = X(X^t(x_0))$ ,  $X^0(x_0) = x_0$ . We assume that (2.1) admits a hyperbolic limit cycle and that  $X$  is analytic.

More precisely, we assume, for some map  $K_0 : \mathbb{T} \rightarrow \mathbb{R}^2$ , that  $x(t) = K_0(\omega t)$  is a solution of (2.1) and, furthermore, that  $\mathcal{K} = K_0(\mathbb{T}^1)$  is an exponentially attracting set. That is, if  $y$  is close enough to  $\mathcal{K}$ , then

$$(2.2) \quad d(X^t(y), \mathcal{K}) \leq C e^{-\lambda t}$$

for some  $C, \lambda > 0$ .

Given (2.2), as pointed out in [18], it follows that we can find a unique  $\Phi(y)$  in such a way that

$$(2.3) \quad |X^t(y) - K_0(t + \Phi(y))| \leq C e^{-\lambda t}.$$

The goal is to show (see Theorem 3.2 for a more precise statement) that, in these circumstances, we can find  $K : \mathbb{T} \times [-1, 1] \rightarrow \mathbb{R}^2$ , an analytic local diffeomorphism, such that

$$(2.4) \quad X \circ K(\theta, s) = DK(\theta, s) \begin{bmatrix} \omega \\ \lambda s \end{bmatrix}.$$

Using the more concise notation  $A_{\omega,\lambda} = [\begin{smallmatrix} \omega \\ \lambda s \end{smallmatrix}]$ , (2.4) can be written as

$$X \circ K(\theta, s) = DK(\theta, s) A_{\omega,\lambda}.$$

Equation (2.4) will be the centerpiece of our approach. Note that if (2.4) holds, (2.2) also holds with the same value of  $\lambda$ . Of course, (2.2) holds for a range of  $\lambda$ 's, while (2.4) determines  $\lambda$  in a unique way. From now on, we will use  $\lambda$  to denote the optimal  $\lambda$  in (2.2).<sup>1</sup>

We should think of (2.4) as a functional equation for the mapping  $K$  and for the real numbers  $\omega$  and  $\lambda$ . The vector field  $X$  is, of course, known.

Note that we are taking the convention that the functions are 1-periodic, not  $2\pi$ -periodic. Hence,  $\omega$  is the inverse of the period.

Note also that with our conventions,  $\lambda$  is positive when the limit cycle is repulsive and negative when the limit cycle is attractive. The method works in both cases. In the stable case, the isochrons are obtained by fixing the asymptotic phase in the future. In the unstable case, the isochrons are obtained by fixing the asymptotic phase in the past. Of course, one can pass from the stable case to the unstable case just by changing the direction of time (equivalently, the sign of the vector field  $X$ ).

**2.1. Geometric interpretation of invariance equation (2.4).** We can think of (2.4) as a change of variables that turns the vector field  $X$  into the straight vector field on  $\mathbb{T} \times [-1, 1]$ :

$$(2.5) \quad A_{\omega,\lambda} \equiv \begin{bmatrix} \omega \\ \lambda s \end{bmatrix}.$$

We can also write (2.4) as

$$X \circ K(\theta, s) = [\omega \partial_\theta + \lambda s \partial_s] K(\theta, s),$$

but the formulation as in (2.4) is geometrically more natural.

It is straightforward to show that if we have (2.4), the evolution in the coordinates  $(\theta, s)$  becomes

$$(2.6) \quad X^t(K(\theta, s)) = K(\theta + t\omega, s e^{\lambda t}).$$

Indeed, note that, using (2.4), we have

$$\begin{aligned} \frac{d}{dt} K(\theta + t\omega, s e^{\lambda t}) &= (\omega \partial_\theta + s \lambda e^{\lambda t} \partial_s) K(\theta + t\omega, s e^{\lambda t}) \\ &= X(K(\theta + t\omega, s e^{\lambda t})). \end{aligned}$$

We can describe (2.6) as saying that if we perform the change of variables given by  $K$ , the coordinates  $(\theta, s)$  evolve by the linear flow

$$(2.7) \quad \Lambda^t(\theta, s) = (\theta + \omega t, s e^{\lambda t});$$

---

<sup>1</sup>One of the consequences of the theory developed here is that the optimal  $\lambda$  in (2.2) exists. That is, there are no subexponential corrections. See section 2.3 for more details.

that is, we have  $X^t \circ K = K \circ \Lambda^t$ .

In particular, the isochrons are just the sets obtained by fixing  $\theta$  and letting  $s$  vary:

$$(2.8) \quad \mathcal{S}_\theta = \{K(\theta, s) \mid s \in [-1, 1]\}.$$

In Theorem 3.2, we will show that  $K$  is analytic and, as a corollary, that  $\mathcal{S}_\theta$  are analytic manifolds that depend analytically on  $\theta$ .

Note that if we have (2.4), then

$$K_0(\theta) = K(\theta, 0)$$

is a limit cycle, so that the isochrons  $\mathcal{S}_\theta$  are curves transversal to the limit cycle. Indeed,

$$K_1(\theta) = \partial_s K(\theta, s)|_{s=0}$$

is tangent to the isochron at  $K_0(\theta)$  and is transversal to  $\partial_\theta K(\theta, s)|_{s=0}$ , which is tangent to the circle  $K(\theta)$  ( $(\partial_\theta K, \partial_s K)$  has full rank because  $K$  is a diffeomorphism).

Note also that even if the foliation by isochrons is invariant by the flow  $X^t$  of (2.1), the individual leaves are not invariant. Indeed, we have

$$(2.9) \quad \mathcal{S}_{\theta+\omega t} = X^t(\mathcal{S}_\theta).$$

However, we can use this property to numerically extend the computation of the isochrons to a larger neighborhood of the limit cycle. See section 7.

**2.2. Lack of uniqueness.** It is important to note that the solutions of (2.4) are never unique. Indeed, for any  $\theta_0 \in \mathbb{T}$  and  $b \in \mathbb{R}$ , if  $(K, \omega, \lambda)$  is a solution of (2.4) and  $\tilde{K}$  is defined by  $\tilde{K}(\theta, s) = K(\theta + \theta_0, bs)$ , then  $(\tilde{K}, \omega, \lambda)$  is also a solution of (2.4). We will show in Theorem 6.1 that this is the only source of nonuniqueness. In particular,  $\omega$  and  $\lambda$  are uniquely determined.

A practical consequence of this lack of uniqueness is that we can assume (by choosing  $b$ ) that the domain of the parameter  $s$  is  $[-1, 1]$ . This choice of parameter  $b$  is convenient for theoretical calculations, but not essential. On the other hand, it is very important for numerical calculations. We will discuss this issue in detail in sections 7 and 8.

**2.3. Topological characterization of isochrons.** One of the consequences of (2.4) is that the isochrons admit a topological characterization. This is somewhat different from the characterization given by the theory of normally hyperbolic manifolds, which involves not only convergence but convergence at a certain exponential rate. For simplicity, we present only the formula for the stable case, when  $\lambda < 0$  and the asymptotic phase is the phase in the future.

Using (2.4) and the fact that  $K$  is uniformly differentiable (so that it preserves rates of convergence up to a constant), we have for  $0 < \eta \ll 1$ ,

$$(2.10) \quad \begin{aligned} W_{K(\theta, 0)}^s &= \{P \in \mathbb{R}^2 \mid |X_t(P) - X_t(K(\theta, 0))| \leq C_{\eta, P} e^{-(|\lambda|-\eta)t} \text{ for all } t \geq 0\} \\ &= \{K(\theta, s), s \in \mathbb{R}\} \\ &= \{P \in \mathbb{R}^2 \mid |X_t(P) - X_t(K(\theta, 0))| \rightarrow 0 \text{ as } t \rightarrow \infty\} \\ &= \{P \in \mathbb{R}^2 \mid |X_t(P) - X_t(K(\theta, 0))| \leq C_P e^{-|\lambda|t} \text{ for all } t \geq 0\}. \end{aligned}$$

The first characterization of  $W_{K(\theta)}^s$  in (2.10) is the standard definition in the theory of normally hyperbolic manifolds, which involves rates of convergence.

The third line of (2.10) is a purely topological characterization; it involves just convergence irrespective of the rate. The link between all the characterizations in (2.10) is the characterization given in the second line of (2.10). The difference between the first and fourth lines of (2.10) is that in the first line we just assume that there is an open interval of rates of convergence. In the fourth line, we obtain that the endpoint of the rates of convergence allowed in the first line is also allowed.

The proof of the equivalence of all the characterizations follows from the observation that it is valid when the dynamics are described by the linear flow  $\Lambda^t$  (see (2.7)) and that the change of variables provided by  $K$  does not change the rates of convergence (or, of course, true convergence) of the orbits.

The characterizations above show that, for the problem at hand, the orbits that converge do so exponentially fast and the exponential rate is always the optimal one (i.e., there are no polynomial corrections). This result strongly relies on the low dimensionality of the problem and is false in more general normally hyperbolic manifolds.

**3. Statement of the main analytical result: Theorem 3.2.** To formulate Theorem 3.2 precisely, we need some definitions of norms in which to measure functions.

**3.1. Definition of norms.** We use the standard supremum norms in KAM theory. They seem to give the sharpest results in loss of differentiability. As for the numerical work with them, see section 7.1.4.

We denote

$$(3.1) \quad \begin{aligned} \mathbb{T}_\rho &= \{\theta \in \mathbb{C}/\mathbb{Z} \mid |\text{Im}(\theta)| \leq \rho\}, \\ B_\beta &= \{s \in \mathbb{C} \mid |s| \leq \beta\}, \\ \mathcal{D}_{\beta,\rho} &= \mathbb{T}_\rho \times B_\beta. \end{aligned}$$

**Definition 3.1.** Given a periodic function  $f(\theta) = \sum_{k \in \mathbb{Z}} \hat{f}_k e^{2\pi i k \theta}$  and a number  $\rho > 0$ , we define

$$(3.2) \quad \|f\|_\rho = \sup_{\theta \in \mathbb{T}_\rho} |f(\theta)|.$$

Given a family of periodic functions  $K(\theta, s) = \sum_{n \in \mathbb{N}} K_n(\theta) s^n$  with  $K_n(\theta)$  periodic and two numbers  $\rho, \beta > 0$ , we define

$$(3.3) \quad \|K\|_{\beta,\rho} = \sup_{|s| \leq \beta, \theta \in \mathbb{T}_\rho} |K(\theta, s)|.$$

Note that the definitions above are valid in the cases that the functions are real valued or vector valued.

We consider the spaces  $\mathcal{A}_\rho$  and  $\mathcal{A}_{\beta,\rho}$  consisting of the functions for which the norms  $\|\cdot\|_\rho$  and  $\|\cdot\|_{\beta,\rho}$ , respectively, are finite. We consider them equipped with the corresponding norms, which makes  $\mathcal{A}_\rho$  and  $\mathcal{A}_{\beta,\rho}$  Banach spaces.

Some elementary properties of these norms are presented in section 5.1.

### 3.2. Statement of the main analytical result.

**Theorem 3.2.** *Let  $X$  be an analytic vector field in a domain of  $\mathbb{R}^2$  which extends to a domain  $\mathcal{C} \subset \mathbb{C}^2$ .*

*Assume that we can find an analytic parameterization  $K : \mathbb{T} \times [-1, 1] \rightarrow \mathbb{R}^2$  and numbers  $\omega, \lambda$  in such a way that the following hold:*

- $K(\mathbb{T}_\rho \times B_\beta) \subset \mathcal{C}$ ,  $\text{dist}(K(\mathbb{T}_\rho \times B_\beta), \mathbb{C}^2 - \mathcal{C}) = \zeta > 0$ .
- $\|X \circ K - DKA_{\omega, \lambda}\|_{\beta, \rho} < \varepsilon$ .
- For some  $0 < \delta < \rho/2$ , we have

$$(3.4) \quad \varepsilon \delta^{-1} C \leq 1,$$

where  $C$  is an explicit expression developed in the proof depending on  $\sup_{x \in \mathcal{C}} |D^2 X(x)|$ ,  $\sup_{x \in \mathbb{T}_\rho \times B_\beta} |DK(x)|, |DK^{-1}(x)|, |D^2 K(x)|$ .

Then, there exists  $K^*$  an analytic local diffeomorphism and  $\omega^*, \lambda^* \in \mathbb{R}$  such that

$$(3.5) \quad X \circ K^*(\theta, s) = DK^*(\theta, s) A_{\omega^*, \lambda^*}.$$

Furthermore,

$$(3.6) \quad \|K - K^*\|_{\beta-\delta, \rho-\delta}, |\omega - \omega^*|, |\lambda - \lambda^*| \leq C\varepsilon.$$

Here, following the standard practice in KAM arguments, we denote by the letter  $C$  some quantities that depend on the quantities indicated after (3.4) even if the meaning could be different from line to line. In particular, the constant  $C$  that appears in the conclusions (3.6) is different from the constant  $C$  that appears in the hypothesis (3.4). Indeed, there are several pairs of constants that work.

The proof of Theorem 3.2 is given in section 5. It is based on a rapidly convergent iteration. This iteration, which takes advantage of some geometric calculations, is not only the method of proof but also yields efficient algorithms. For this reason, we discuss the iterative step in a separate section (see section 4).

Note that Theorem 3.2 has the format of a posteriori results of numerical analysis. That is, it states that if the initial approximation solves the equation with sufficient accuracy depending on explicit “condition numbers,” then there is a true solution nearby. Moreover, one can bound the difference between the initial approximation and the true solution by the residual of (2.4) evaluated on the initial approximation.

The solution produced by Theorem 3.2 is also unique in a neighborhood except for the reparameterizations discussed in section 2.2. Moreover, solutions depend smoothly on parameters. We discuss these results in section 6.

**4. The iterative step for the computation of  $(K, \omega, \lambda)$ .** In this section we present the procedure, based on the Newton method, for obtaining a more approximate solution  $(K, \omega, \lambda)$  of (2.4) out of a sufficiently approximate one. More precisely, we state explicitly the functional equation that needs to be solved, we describe the algebraic manipulations to simplify this equation, and finally we provide two methods for solving the simplified equation: one using Fourier series and another one using integral representations.

This procedure will be the basis of both the convergence proof of Theorem 3.2 (see section 5) and the numerical algorithm (see section 4.4).

**4.1. The Newton method.** Given an approximate solution  $(K, \omega, \lambda)$  of (2.4) such that

$$(4.1) \quad X \circ K - DKA_{\omega, \lambda} = E,$$

the Newton method seeks an improved solution  $(K + \Delta, \omega + \sigma, \lambda + \eta)$ , in such a way that  $(\Delta, \sigma, \eta)$  eliminates  $E$  “in the linear approximation.”

Since  $X \circ (K + \Delta) \approx X \circ K + DX \circ K\Delta$  and  $D(K + \Delta) = DK + D\Delta$ , we have that the equation for the Newton method is

$$(4.2) \quad (DX \circ K)\Delta - (D\Delta)A_{\omega, \lambda} - DK A_{\sigma, \eta} = -E.$$

One should think of (4.2) as an equation for  $\Delta$ ,  $\eta$ , and  $\sigma$  when all the other quantities are known. Indeed,  $X$  is given by the problem and  $(K, \omega, \lambda)$  is the known approximation we are trying to improve.

Note that at this stage it is not clear that (4.2) has a solution because we have periodicity requirements on  $K$ . The fact that (4.2) has solutions will be established in section 4.2.

If we discretize our functions in some appropriate basis of functions satisfying the periodicity conditions, (4.2) yields a linear equation that can be solved using a linear solver. This is a reasonably practical approach in many circumstances [19]. The main drawback is that one needs to solve a linear system which requires one to invert a full matrix of the dimension of the discretization.

In this paper, however, we take a different approach. We use several identities to obtain a change of variables which reduces (4.2) to a much simpler equation (up to a certain error which is smaller than the original error and does not change the quadratic character of the Newton method). This is what we call a quasi-Newton method, and we discuss it in detail in the next section.

*Remark 4.1.* The use of these identities in linearization problems was pointed out in [31]. A more systematic study based on “group structure” of the equations is in [48]. Some extensions of this approach were used in [26] for Hamiltonian systems, taking advantage of the geometric properties of the system. In our case, the geometric property we take advantage of is, mainly, the lower dimensionality of the system.

**4.2. The quasi-Newton method.** Consider an approximate solution  $(K, \omega, \lambda)$  of (2.4) satisfying (4.1). Notice that taking derivatives of (4.1) we obtain that our approximate solution  $(K, \omega, \lambda)$  also satisfies

$$(4.3) \quad (DX \circ K)DK - D^2K A_{\omega, \lambda} - DK DA_{\omega, \lambda} = DE,$$

where

$$DA_{\omega, \lambda} = \begin{pmatrix} 0 & 0 \\ 0 & \lambda \end{pmatrix}.$$

We emphasize that both (4.1) and, hence, (4.3) give information that is on hand at the beginning of the iterative step.

Next, we show how to use (4.3) to simplify the equation for the Newton method (4.2). The crucial idea is that rather than looking for  $(\Delta, \sigma, \eta)$  in (4.2), we look for  $(W, \sigma, \eta)$ , where

$$(4.4) \quad \Delta = DK W.$$

Note that if  $DK$  is invertible, both  $\Delta$  and  $W$  are equivalent unknowns in the sense that if we know one, we can find the other one. Then, if we substitute (4.4) into (4.2), we obtain that (4.2) is equivalent to

$$(4.5) \quad (DX \circ K)DK W - D^2 K W A_{\omega,\lambda} - DK DW A_{\omega,\lambda} - DK A_{\sigma,\eta} = -E.$$

Using (4.3), we obtain that (4.5) is equivalent to (recall that  $D^2 K W A_{\omega,\lambda} = D^2 K A_{\omega,\lambda} W$  because  $D^2 K$  is a symmetric quadratic form)

$$(4.6) \quad DK DA_{\omega,\lambda} W + DEW - DK DW A_{\omega,\lambda} - DK A_{\sigma,\eta} = -E.$$

The quasi-Newton method consists just in dropping the term  $DEW$  from (4.6), which we argue, heuristically at the moment, is “quadratically small” because it is the product of two terms which are small (think of  $W$  as of the same order of smallness as  $E$ ). This heuristic idea that  $DEW$  is small will be made rigorous when we perform estimates in section 5.2.

Hence, we will consider the equation for  $(W, \sigma, \eta)$ ,

$$(4.7) \quad DK DA_{\omega,\lambda} W - DK DW A_{\omega,\lambda} - DK A_{\sigma,\eta} = -E,$$

and then consider the improved solution  $(K + DK W, \omega + \sigma, \lambda + \eta)$ . This will be referred to as the quasi-Newton step.

If we premultiply (4.7) by  $DK^{-1}$ , we obtain

$$(4.8) \quad \begin{pmatrix} 0 & 0 \\ 0 & \lambda \end{pmatrix} W - DW A_{\omega,\lambda} - A_{\sigma,\eta} = -DK^{-1} E.$$

If we express (4.8) in components, using the shorthand  $\tilde{E} = DK^{-1} E$  and denoting the components of  $\tilde{E}$  and  $W$  by subindices,  $\tilde{E} = (\tilde{E}_1, \tilde{E}_2)$  and  $W = (W_1, W_2)$ , we obtain

$$(4.9) \quad \begin{aligned} -(\omega \partial_\theta + \lambda s \partial_s) W_1 - \sigma &= -\tilde{E}_1, \\ \lambda W_2 - (\omega \partial_\theta + \lambda s \partial_s) W_2 - \eta s &= -\tilde{E}_2. \end{aligned}$$

The remarkable feature of (4.9) is that it involves only a linear operator with constant coefficients. As we will see next, these equations can be solved very efficiently either in Fourier coefficients (see Lemma 4.2) or using explicit (and fast converging) integral formulas (see Lemma 4.3).

**4.3. Solutions of the constant coefficient linearized equations.** In this section, we study the solvability of equations (4.9) using both (formal) Fourier series and improper (but rapidly convergent) integrals. The latter are convenient when the functions are discretized using splines or collocation methods, like when the vector field  $X$  is known only at some points obtained experimentally. Detailed estimates will be established in section 5. Nevertheless, for the purpose of implementing algorithms, only the existence and the form of the solutions are needed.

#### 4.3.1. Solutions of linearized equations by Fourier methods.

**Lemma 4.2.** Consider a formal series  $\tilde{E} = \sum \tilde{E}_{j,k} s^j e^{2\pi i k \theta}$ . Then, if  $\tilde{E}_{00} = 0$ , the equation for  $u$ ,

$$(4.10) \quad (\omega \partial_\theta + \lambda s \partial_s) u = \tilde{E},$$

has the one-dimensional family of formal series solutions

$$\sum_{j \in \mathbb{N}, k \in \mathbb{Z}} u_{j,k} s^j e^{2\pi i k \theta}$$

with

$$(4.11) \quad \begin{aligned} u_{j,k} &= \frac{\tilde{E}_{j,k}}{2\pi i \omega k + \lambda j} \quad \text{if } (j, k) \neq (0, 0), \\ u_{0,0} &= \alpha \end{aligned}$$

for any  $\alpha \in \mathbb{R}$ . The solutions given in (4.11) are the only formal series solutions of (4.10). Furthermore if  $\tilde{E}_{00} \neq 0$ , there are no formal series solutions of (4.10).

If  $\tilde{E}_{10} = 0$ , then the equation for  $u$ ,

$$(4.12) \quad -\lambda u + (\omega \partial_\theta + \lambda s \partial_s) u = \tilde{E},$$

has the one-parameter family of formal series solutions

$$\sum_{j \in \mathbb{N}, k \in \mathbb{Z}} u_{jk} s^j e^{2\pi i k \theta}$$

with

$$(4.13) \quad \begin{aligned} u_{jk} &= \frac{\tilde{E}_{jk}}{2\pi i \omega k + \lambda(j-1)} \quad \text{if } (j, k) \neq (1, 0), \\ u_{1,0} &= \alpha \end{aligned}$$

for any  $\alpha \in \mathbb{R}$ . The solutions given by (4.13) are the only formal series solutions of (4.12). Furthermore, if  $\tilde{E}_{1,0} \neq 0$ , (4.12) has no solutions.

*Proof.* Taking Fourier series on both sides of (4.10) we have

$$u_{jk}(\lambda j + 2\pi i \omega k) = \tilde{E}_{jk}.$$

It is easy to see that  $\lambda j + 2\pi i \omega k = 0$  if and only if  $j = 0, k = 0$ . Therefore, the solution  $u$  is obtained by setting  $u_{jk} = \tilde{E}_{jk}/(\lambda j + 2\pi i \omega k)$  when  $(j, k) \neq (0, 0)$  and  $u_{00}$  is arbitrary.

Similarly, we observe that (4.12) is equivalent to

$$(\lambda j - \lambda + 2\pi i \omega k) u_{jk} = \tilde{E}_{jk}.$$

Again, we note that  $\lambda(j-1) + 2\pi i \omega k = 0$  if and only if  $j = 1, k = 0$ , and then the same argument as before applies. ■

Consider the expressions given by (4.11) and (4.13). If  $\tilde{E}$  is not just a formal power series, but rather a smooth function, the solutions above will also have several regularity properties. The reason is that regularity of the error implies fast decay properties for the Fourier–Taylor coefficients, which in turn imply fast decay of the Fourier–Taylor coefficients of the solutions and, hence, regularity properties of the solutions. Detailed estimates will be presented in section 5.

### 4.3.2. Solutions of linearized equations by improper integrals.

**Lemma 4.3.** Consider  $\tilde{E} : \mathbb{T} \times [-1, 1] \rightarrow \mathbb{R}^2$ , which is a  $C^r$  function  $r \in \mathbb{N} \cup \{\infty, \omega\}$ ,  $r \geq 1$ . Assume that we are in the stable case ( $\lambda < 0$ ).

If  $\int_0^1 \tilde{E}(\theta, 0) d\theta = 0$ , (4.10) has the solutions

$$(4.14) \quad u(\theta, s) \equiv \alpha + \frac{1}{\omega} \int_0^\theta \tilde{E}(\sigma, 0) d\sigma - \int_0^\infty [\tilde{E}(\theta + \omega t, se^{\lambda t}) - \tilde{E}(\theta + \omega t, 0)] dt.$$

Furthermore, if  $\int \tilde{E}(\theta, 0) \neq 0$ , there is no  $C^0$  solution of (4.10). The only solutions of (4.10) in  $C^0(\mathbb{T} \times [-1, 1])$  are (4.14).

If  $r \geq 2$ ,  $\int_0^1 \partial_s \tilde{E}(\theta, 0) d\theta = 0$ , (4.12) has the solutions

$$(4.15) \quad \begin{aligned} u(\theta, s) = & A(\theta) + sB(\theta) + \int_0^\infty e^{-\lambda t} [\tilde{E}(\theta + \omega t, se^{\lambda t}) \\ & - \tilde{E}(\theta + \omega t, 0) - se^{\lambda t} \partial_s \tilde{E}(\theta + \omega t, 0)] dt, \end{aligned}$$

where

$$\begin{aligned} A(\theta) &= \int_0^\infty e^{\lambda t} \tilde{E}(\theta - \omega t, 0) dt, \\ B(\theta) &= \alpha + \frac{1}{\omega} \int_0^\theta \partial_s \tilde{E}(\sigma, 0) d\sigma. \end{aligned}$$

Furthermore, if  $\int_0^1 \partial_s \tilde{E}(\theta, 0) d\theta \neq 0$ , there is no  $C^1$  solution of (4.12). The only solutions of (4.12) in  $C^2(\mathbb{T} \times [-1, 1])$  are those given by (4.15).

*Proof.* We observe that if we particularize (4.10) to  $s = 0$ , we obtain

$$\omega \partial_\theta u(\theta, 0) = \tilde{E}(\theta, 0).$$

Hence, by the fundamental theorem of calculus, the only continuous solutions of (4.10) should satisfy

$$(4.16) \quad u(\theta + \omega T, 0) - u(\theta, 0) = \int_0^T \tilde{E}(\theta + \omega t, 0) dt.$$

We see that the expression (4.16) is periodic in  $T$  if and only if  $\int_0^1 \tilde{E}(\theta, 0) d\theta = 0$ . Hence, if  $\int_0^1 \tilde{E}(\theta, 0) d\theta \neq 0$ , there is no solution  $u$ . In the rest of the discussion we will assume  $\int_0^1 \tilde{E}(\theta, 0) d\theta = 0$ . We also note that the expression (4.16) is a solution of (4.10) on the set  $s = 0$ .

Also, by the fundamental theorem of calculus and adding and subtracting terms we obtain that any solution of (4.10) should satisfy

$$(4.17) \quad \begin{aligned} & u(\theta + \omega T, se^{\lambda T}) - u(\theta, s) \\ &= \int_0^T [\tilde{E}(\theta + \omega t, se^{\lambda t}) - \tilde{E}(\theta + \omega t, 0)] dt + \int_0^T \tilde{E}(\theta + \omega t, 0) dt \\ &= \int_0^T [\tilde{E}(\theta + \omega t, se^{\lambda t}) - \tilde{E}(\theta + \omega t, 0)] dt + u(\theta + \omega T, 0) - u(\theta, 0), \end{aligned}$$

because  $u(\theta + \omega T, se^{\lambda T}) - u(\theta + \omega T, 0)$  converges to 0 as  $T \rightarrow \infty$ , and

$$|\tilde{E}(\theta + \omega t, se^{\lambda t}) - \tilde{E}(\theta + \omega t, 0)| \leq ce^{\lambda t}.$$

We obtain that the integral (4.17) is uniformly convergent, and we obtain that the function  $u$  is the only candidate for a solution.

Taking  $T \rightarrow \infty$  in the expression (4.17) we have

$$u(\theta, s) = u(\theta, 0) - \int_0^\infty [\tilde{E}(\theta + \omega t, se^{\lambda t}) - \tilde{E}(\theta + \omega t, 0)] dt,$$

and using again the fundamental theorem of calculus we have

$$u(\theta, 0) = u(0, 0) + \int_0^\theta \partial_\theta u(\sigma, 0) d\sigma = \alpha + \frac{1}{\omega} \int_0^\theta \tilde{E}(\sigma, 0) d\sigma,$$

which gives (4.14).

Since the integrand in (4.17) converges fast enough, we can compute the derivatives of the integral by computing the derivatives of the integrand and, therefore, conclude that  $u$  is indeed a solution of (4.10).

To prove the second claim of Lemma 4.3, we proceed as before. We start by computing candidates for  $u(\theta, 0)$  and  $\partial_s u(\theta, 0)$ , and then we show that the integrand converges fast enough that we can justify that they are indeed solutions. This strategy is very common in linearization problems and in invariant manifold theorems.

We observe that using the integrating factor  $e^{-\lambda\theta/\omega}$  in (4.12) we have

$$-\lambda e^{-\lambda\theta/\omega} u(\theta, s) + e^{-\lambda\theta/\omega} (\omega \partial_\theta + \lambda s \partial_s) u(\theta, s) = e^{-\lambda\theta/\omega} \tilde{E}(\theta, s)$$

and therefore

$$(\omega \partial_\theta + \lambda s \partial_s) [e^{-\lambda\theta/\omega} u(\theta, s)] = e^{-\lambda\theta/\omega} \tilde{E}(\theta, s).$$

Using the fundamental theorem of calculus

$$e^{-\lambda(\theta+\omega T)/\omega} u(\theta + \omega T, se^{\lambda T}) - e^{-\lambda\theta/\omega} u(\theta, s) = \int_0^T e^{-\lambda(\theta+\omega t)/\omega} \tilde{E}(\theta + \omega t, se^{\lambda t}) dt$$

and multiplying by  $e^{\lambda\theta/\omega}$ , we have the variation of parameters formula

$$(4.18) \quad e^{-\lambda T} u(\theta + \omega T, se^{\lambda T}) - u(\theta, s) = \int_0^T e^{-\lambda t} \tilde{E}(\theta + \omega t, se^{\lambda t}) dt.$$

We observe that using the variation of parameters formula (4.18) for  $s = 0$ , we obtain

$$e^{-\lambda T} u(\theta + \omega T, 0) - u(\theta, 0) = \int_0^T e^{-\lambda t} \tilde{E}(\theta + \omega t, 0) dt,$$

which after multiplying by  $e^{\lambda T}$  and performing the change of variables  $\tilde{\theta} = \theta + \omega T$  becomes

$$\begin{aligned} u(\tilde{\theta}, 0) &= u(\tilde{\theta} - \omega T, 0) e^{\lambda T} + \int_0^T e^{\lambda(T-t)} \tilde{E}(\tilde{\theta} - \omega(T-t), 0) dt \\ &= u(\tilde{\theta} - \omega T, 0) e^{\lambda T} + \int_0^T e^{\lambda t} \tilde{E}(\tilde{\theta} - \omega t, 0) dt. \end{aligned}$$

Taking limits as  $T \rightarrow \infty$ , we obtain the expression for  $A$  in (4.15).

Now, we observe that if we take derivatives with respect to  $s$  of (4.12) and evaluate at  $s = 0$ , we obtain

$$-\lambda \partial_s u(\theta, 0) + \omega \partial_\theta \partial_s u(\theta, 0) + \lambda \partial_s u(\theta, 0) = \partial_s \tilde{E}(\theta, 0),$$

which we rewrite as

$$\omega \partial_\theta [\partial_s u(\theta, 0)] = \partial_s \tilde{E}(\theta, 0).$$

Therefore, we have that if  $\int_0^1 \partial_s \tilde{E}(\theta, 0) d\theta = 0$ , there is no periodic solution. Otherwise, we obtain  $B$  in (4.15).

Equation (4.12) is obviously linear, and we have found solutions  $A(\theta)$ ,  $sB(\theta)$  corresponding to the right-hand side (RHS)  $\tilde{E}(\theta, 0)$ ,  $s\partial_s \tilde{E}(\theta, 0)$ . Therefore it suffices to find solutions for an RHS of (4.12) of the form

$$\tilde{\tilde{E}}(\theta, s) = \tilde{E}(\theta, s) - \tilde{E}(\theta, 0) - s\partial_s \tilde{E}(\theta, 0).$$

Again, we will find a candidate  $\tilde{u}(\theta, s)$  and verify that indeed it is a solution.

The variation of parameters formula (4.18) gives

$$e^{-\lambda T} \tilde{u}(\theta + \omega T, se^{\lambda T}) - \tilde{u}(\theta, s) = \int_0^T e^{-\lambda t} \tilde{\tilde{E}}(\theta + \omega t, se^{\lambda t}) dt.$$

We note that, because  $\tilde{\tilde{E}}$  is  $C^2$  and  $\tilde{\tilde{E}}(\theta, 0) = 0$ ,  $\partial_s \tilde{\tilde{E}}(\theta, 0) = 0$ ,

$$|\tilde{\tilde{E}}(\theta + \omega t, se^{\lambda t})| \leq Ce^{2\lambda t}.$$

Hence, the integral in the RHS is convergent if we take the limit  $T \rightarrow \infty$ .

Since we have found the linear parts for  $u(\theta, s)$ , it is natural to guess that  $\tilde{u}(\theta, s) \leq Cs^2$ . Thus, we guess that the only solution for (4.12) is

$$(4.19) \quad \tilde{u}(\theta, s) = \int_0^\infty e^{-\lambda t} \tilde{\tilde{E}}(\theta + \omega t, se^{\lambda t}) dt.$$

To prove that (4.19) is indeed a solution of (4.12), we note that taking derivatives under the integral sign (which is justified by the rapid convergence to zero of the integrand and its derivatives), we obtain that (4.19) satisfies (4.12).

Now we observe that, using the definition of  $\tilde{u}$ , if there was a solution  $u$  of (4.12), it would satisfy

$$-\lambda(u - \tilde{u}) + (\omega \partial_\theta + \lambda s \partial_s)(u - \tilde{u}) = \tilde{E}(\theta, 0) + s\partial_s \tilde{E}(\theta, 0),$$

but we have already established the uniqueness of this previous approximation. ■

Of course, from the form of the solutions (4.14) and (4.15) we can also obtain regularity properties, but this will be done in section 5.

Note that, since  $\lambda < 0$ , the integrals defining  $A$  are uniformly convergent. The integrals in (4.15) are also uniformly convergent because, by Taylor's theorem,

$$|\tilde{E}(\theta + \omega t, se^{\lambda t}) - \tilde{E}(\theta + \omega t, 0) - se^{\lambda t} \partial_s \tilde{E}(\theta + \omega t, 0)| \leq Ce^{2\lambda t}.$$

**4.4. Algorithm for the quasi-Newton step.** In this section we specify step by step the implementation of the quasi-Newton step.

The theory of solvability of linearized equations developed in Lemma 4.2 indicates how to approach the solution of (4.9). We determine the unknowns  $\sigma, \eta$  so that the equations are all solvable, namely,

$$(4.20) \quad \begin{aligned} \sigma &= \int_0^1 \tilde{E}_1(\theta, 0) d\theta, \\ \eta &= \int_0^1 \partial_s \tilde{E}_2(\theta, s)|_{s=0} d\theta. \end{aligned}$$

Therefore, we are led to the following algorithm.

**Algorithm 4.4.** Consider a vector field  $X : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Given  $K : \mathbb{T} \times [-1, 1] \rightarrow \mathbb{R}^2$ ,  $\omega \in \mathbb{R}$ ,  $\lambda \in \mathbb{R}$ , compute the following:

1.  $\alpha \leftarrow X \circ K$ .
2.  $\beta \leftarrow DK$ .
3.  $E \leftarrow \alpha - \beta A_{\omega\lambda}$ .
4.  $\tilde{E} = DKE$ . Denote  $\tilde{E} = (\tilde{E}_1, \tilde{E}_2)$ .
5.  $\sigma = \int_0^1 \tilde{E}_1(\theta, 0) d\theta$ ,
- $\eta = \int_0^1 \partial_s \tilde{E}_2(\theta, s)|_{s=0} d\theta$ .
6. Find  $W_1$  solving

$$(\omega \partial_\theta + \lambda s \partial_s) W_1 = \tilde{E}_1 - \sigma.$$

We also impose the normalization  $\int W_1|_{s=0} = 0$ , so that the solution is unique.

7. Find  $W_2$  solving

$$(\omega \partial_\theta + \lambda s \partial_s) W_2 - \lambda W_2 = \tilde{E}_2 - \eta s.$$

We also impose the normalization  $\int \partial_s W_2|_{s=0} = 0$ , so that the solution is unique.

8. Denote  $W = (W_1, W_2)$ . The improved solution is

$$\tilde{K} = K + DKW, \quad \omega = \omega + \sigma, \quad \lambda = \lambda + \eta.$$

A remarkable feature of Algorithm 4.4 is that even if it is a quadratically convergent algorithm, at no stage of the algorithm it required to compute (much less to invert) a matrix of the dimension of the discretization. We need only perform algebraic operations among functions, compute derivatives, and solve the cohomology equations.

All the above operations can be implemented either in Taylor–Fourier series or using a discretization in a grid of points and interpolating using, e.g., splines. We discuss the numerical implementation details in section 7.

**Remark 4.5.** Notice that the solution  $W_1$  obtained in step 6 is unique up to the addition of a constant and  $W_2$  computed in step 7 is unique up to the addition of a multiple of the first order coefficient. The solutions obtained for  $\eta, \sigma$  are unique. The indeterminacy in the solutions of 6 and 7 can be used to achieve other normalizations (see section 6.1). These are the only indeterminacies of the modified Newton equation (see sections 2.2 and 6.1).

**5. Convergence of the iterative step and proof of Theorem 3.2.** In this section we prove Theorem 3.2, which establishes the convergence of Algorithm 4.4 provided that we start with a sufficiently approximate solution of (2.4).

To obtain convergence proofs we need to supplement the discussion in section 4 with considerations of function space and estimates on the solution and the error that quantify the assertion that the result of the iterative step satisfies the invariance equation more accurately. Finally, we need to show that the procedure for producing more accurate solutions can be iterated, hence obtaining a sequence of functions which converge to a solution.

The a posteriori results such as Theorem 3.2 are very typical of results on convergence of methods based in Newton–Kantorovich algorithms, which have an iterative method leading to a fixed point. Nevertheless, since our iterative step involves taking derivatives, we will have to use Nash–Moser estimates rather than the more elementary Kantorovich ones. This will require introducing norms to measure the distance between functions. Although Nash–Moser methods are very robust and can work with several norms, we will discuss only the most customary supremum norms introduced in Definition 3.1.

**5.1. Some elementary properties of the norms in Definition 3.1.** In this section we review some elementary properties of the norms introduced in section 3.1.

**Proposition 5.1.** *Consider the notation introduced in Definition 3.1. For any  $f, g \in \mathcal{A}_\rho$  and  $K, L \in \mathcal{A}_{\beta, \rho}$ , we have*

$$(5.1) \quad \begin{aligned} \|f \cdot g\|_\rho &\leq \|f\|_\rho \|g\|_\rho, \\ \|K \cdot L\|_{\beta, \rho} &\leq \|K\|_{\beta, \rho} \|L\|_{\beta, \rho}. \end{aligned}$$

The proof of Proposition 5.1 is immediate using that the supremum of the product is less than the product of the supremums.

**Proposition 5.2.** *For any  $\delta > 0$ , we have for any  $f \in \mathcal{A}_\rho$ ,  $K \in \mathcal{A}_{\beta, \rho}$ ,*

$$(5.2) \quad \begin{aligned} \|\partial_\theta f\|_{\rho-\delta} &\leq C\delta^{-1}\|f\|_\rho, \\ \|\partial_s K\|_{\beta-\delta, \rho} &\leq C\delta^{-1}\|K\|_{\beta, \rho}, \\ \|\partial_\theta K\|_{\beta, \rho-\delta} &\leq C\delta^{-1}\|K\|_{\beta, \rho}. \end{aligned}$$

This is a very standard result in complex analysis that follows from the Cauchy formula for the derivative as a contour integral (see [1, 39]).

**Proposition 5.3.** *Let  $X$  be an analytic vector field in a domain  $\mathcal{C} \subset \mathbb{C}^2$ . Let  $K : \mathbb{T}_\rho \times B_\beta \rightarrow \mathbb{C}^2$  be such that*

$$\text{dist}(K(\mathbb{T}_\rho \times B_\beta), \mathbb{C}^2 - \mathcal{C}) \geq \zeta > 0.$$

*Then, the following hold:*

- $X \circ K \in \mathcal{A}_{\beta, \rho}$ . In particular,  $X \circ K$  is analytic on  $\mathbb{T}_\rho \times B_\beta$ .
- For all  $\gamma : \mathbb{T}_\rho \times B_\beta \rightarrow \mathbb{C}^2$  with  $\|\gamma\|_{\beta, \rho}$  sufficiently small, we have

$$(5.3) \quad \|X \circ (K + \gamma) - X \circ K - (DX \circ K)\gamma\|_{\beta, \rho} \leq C\|\gamma\|_{\beta, \rho}^2.$$

The proof of Proposition 5.3 follows from the observation that, for each  $x \in \mathbb{T}_\rho$ , we can use Taylor's theorem and then take the supremum. This gives that the constant  $C$  appearing in (5.3) can be taken to be

$$C = \frac{1}{2} \sup_{x \in \mathcal{C}} \|D^2 X\|.$$

Another useful property of the norms is that they are log convex in  $\rho$ . This is just *Hadamard's three circle theorem* [1, 39]. For all  $\rho_1, \rho_2 > 0$ ,  $0 \leq \alpha \leq 1$ , we have

$$(5.4) \quad \|\phi\|_{\alpha\rho_1+(1-\alpha)\rho_2} \leq \|\phi\|_{\rho_1}^\alpha \|\phi\|_{\rho_2}^{1-\alpha}.$$

**5.2. Estimates for the iterative step.** In this section we quantify the argument presented heuristically showing that the error in (2.4) after the iterative step is bounded by the square of the error before the step. There are some subtleties (standard in KAM theory) that need to be taken into account: (a) The bounds after the step are in a slightly smaller domain; (b) the bounds have constants that blow up (like a power) on the loss of analyticity; and (c) the bounds have constants that depend on some nondegeneracy conditions which can be written explicitly and consist of algebraic expressions involving derivatives of  $K$ .

**Lemma 5.4.** *Assume that  $X$  is analytic in some domain  $U \subset \mathbb{C}^2$ . Let  $K : \mathcal{D}_{\beta,\rho} = \mathbb{T}_\rho \times B_\beta \rightarrow U$  belong to  $\mathcal{A}_{\beta,\rho}$ .*

*Assume that*

$$d(\text{Range}(K(\mathcal{D}_{\beta,\rho})), \mathbb{C}^2 - U) \geq \zeta > 0.$$

*Assume, furthermore, that for some  $m \geq 0$ ,*

$$(5.5) \quad \begin{aligned} \|K\|_{\beta,\rho} &\leq m, \\ \|DK\|_{\beta,\rho}, \|D^2K\|_{\beta,\rho} &\leq m_+, \\ \|DK^{-1}\|_{\beta,\rho} &\leq m_-, \\ \omega &\geq \tilde{m}, \text{ and } \lambda \leq -\tilde{m}. \end{aligned}$$

*Let  $E$  be the error function defined as*

$$E = X \circ K - DKA_{\omega,\lambda},$$

*and let  $\delta > 0$  be such that*

$$(5.6) \quad \delta^{-1}m\|E\|_{\beta,\rho} \leq \zeta/100.$$

*Then, there is a constant  $C$  depending only on  $\zeta, m_+, m_-, \tilde{m}$  such that the improved solution  $(K + \Delta, \omega + \sigma, \lambda + \eta)$  obtained after the quasi-Newton step specified in Algorithm 4.4 satisfies*

$$(5.7) \quad \|X \circ (K + \Delta) - D(K + \Delta)A_{\omega+\sigma,\lambda+\eta}\|_{\beta-\delta,\rho-\delta} \leq C\delta^{-1}\|E\|_{\beta,\rho}^2.$$

**Proof.** We want to estimate the error of the improved approximation  $(X + \Delta, \omega + \sigma, \lambda + \eta)$ , where  $\Delta = DKW$ , and  $(W, \sigma, \eta)$  are obtained through Algorithm 4.4 (see steps 4–7). The proof follows just walking through the argument presented in section 4.4 and adding and subtracting appropriate terms in the linear expansion.

First we note that the formula for computing the Fourier coefficients of the function  $W$  is specified in (4.11) and (4.13). Then, using that  $\tilde{E} = DKE$  and Proposition 5.1, we have that

$$(5.8) \quad \begin{aligned} |\sigma|, |\eta| &\leq \|\tilde{E}\|_{\beta,\rho} \\ &\leq \|DK\|_{\beta,\rho}\|E\|_{\beta,\rho} \\ &\leq C\|E\|_{\beta,\rho}, \\ \|W\|_{\beta,\rho} &\leq C\left(\frac{1}{\omega} + \frac{1}{\lambda}\right)\|E\|_{\beta,\rho}, \end{aligned}$$

where  $C$  is a constant that depends on  $m_+$ .

The following identity is obtained just adding and subtracting some terms (those we declared as the leading coefficients and those that we cancel) and grouping

$$\begin{aligned}
 & X \circ (K + \Delta) - D(K + \Delta)A_{\omega+\sigma, \lambda+\eta} \\
 &= X \circ (K + DKW) - D(K + DKW)(A_{\omega, \lambda} + A_{\sigma, \eta}) \\
 &= X \circ (K + DKW) - DKA_{\omega, \lambda} - DKA_{\sigma, \eta} - D^2KWA_{\omega, \lambda} \\
 &\quad - D^2KWA_{\sigma, \eta} - DKDW A_{\omega, \lambda} - DKDW A_{\sigma, \eta} \\
 &= X \circ (K + DKW) - X \circ K - DX \circ KDKW \\
 (5.9) \quad &\quad + [DX \circ KDK - D^2KA_{\omega, \lambda} - DKDA_{\omega, \lambda}]W \\
 &\quad + X \circ K - DKA_{\omega, \lambda} \\
 &\quad + DKDA_{\omega, \lambda}W - DKDW A_{\omega, \lambda} - DKA_{\sigma, \eta} \\
 &\quad - DKDW A_{\sigma, \eta} \\
 &\quad - D^2KWA_{\sigma, \eta}.
 \end{aligned}$$

The different lines (which we denote by  $\ell_1 - \ell_6$ ) in the last expression of (5.9) can be estimated as follows.

The third line ( $\ell_3$ ) in (5.9) is just  $E$ , and  $(W, \sigma, \eta)$  are chosen so that the third and fourth lines of (5.9) cancel exactly (see (4.7)).

We recall that we denote by  $C$  numbers that are controlled by some function of the condition numbers  $m$  in (5.5) as well as  $\omega, \lambda$ , and  $\sup_{x \in C} |D^2X|$ .

The first line of (5.9) can be estimated using the Taylor remainder and the bound for  $\|W\|_{\beta, \rho}$  obtained in (5.8). We obtain

$$\begin{aligned}
 (5.10) \quad \|\ell_1\|_{\beta, \rho} &\leq \frac{1}{2}\|D^2X\|_C (\|DK\|_{\beta, \rho}\|W\|_{\beta, \rho})^2 \leq \frac{1}{2}C\|W\|_{\beta, \rho}^2 \\
 &\leq C\|E\|_{\beta, \rho}^2.
 \end{aligned}$$

The second line of (5.9) can be estimated observing that the expression in brackets is the derivative of  $E$  (see (4.3)). Then, using Cauchy bounds (Proposition 5.2) and the bound for  $\|W\|_{\beta, \rho}$  obtained in (5.8) as well as the Banach algebra properties (Proposition 5.1), we obtain

$$\begin{aligned}
 \|\ell_2\|_{\beta-\delta, \rho-\delta} &\leq C\delta^{-1}\|E\|_{\beta, \rho}\|W\|_{\beta, \rho} \\
 &\leq C\delta^{-1}\|E\|_{\beta, \rho}^2.
 \end{aligned}$$

The fifth and sixth lines of (5.9) can be estimated straightforwardly using the estimates for  $\sigma, \eta$ , and  $W$  in (5.8) and Cauchy bounds (Proposition 5.2) by

$$\begin{aligned}
 (5.11) \quad \|\ell_5\|_{\beta-\delta, \rho-\delta} &\leq C\delta^{-1}\|W\|_{\beta, \rho}|(\sigma, \eta)| \\
 &\leq C\delta^{-1}\|E\|_{\beta, \rho}^2, \\
 \|\ell_6\|_{\beta, \rho} &\leq C\|W\|_{\beta, \rho}|(\sigma, \eta)| \\
 &\leq C\|E\|_{\beta, \rho}^2.
 \end{aligned}$$

Taking the minimum of these estimates we obtain the bound in (5.7). ■

*Remark 5.5.* In contrast with the usual KAM problems, we do not lose derivatives in the solutions of the linearized problem. In that case, the loss of derivatives in the Newton step is because the functional equation (2.4) involves computation of derivatives to straighten the vector field and composition of functions (which is not a differentiable operator unless one loses some domain).

**5.3. Repeating the iteration and end of the proof of Theorem 3.2.** The proof is very standard in KAM theory. We follow very closely the presentation in [25] for KAM problems.

We assume by induction that the iterative step can be carried out  $n$  times (i.e., that hypothesis (5.6) is verified for the first  $n$  steps). We denote by superindices  $(n)$  the parameterization  $K$  after  $n$  steps of the iterative process. For all the other objects we use subindices. We will show that, under certain assumptions on the size of  $\delta_0$  and the error in the initial approximation  $\varepsilon_0$ , which will be independent of  $n$ , hypothesis (5.6) will be verified for  $n+1$ . Moreover, we will show that the error for successive approximations decreases very fast (superexponentially).

We start by fixing for  $n \geq 1$

$$(5.12) \quad \delta_n = \frac{1}{4} \delta_0 2^{-n},$$

where  $\delta_0$  is the global analyticity loss denoted by  $\delta$  in Theorem 3.2. We will show that this choice of  $\delta_n$  is acceptable when the error in the initial approximation  $\varepsilon_0$  is small enough.

The condition numbers  $m_-$ ,  $m_+$ , and  $\tilde{m}$  will be changing in the iteration. We will assume inductively that they are twice as bad as the initial value. We will show that this induction assumption is maintained if  $\varepsilon_0$  is small enough. We will denote by  $C$  the constant that corresponds to the values of  $m$ ,  $m_+$ , and  $\tilde{m}$ , which are twice the original values.

Denoting by  $\varepsilon_n$  the value of the error at step  $n$ , the estimates for the iterative step can be written as

$$(5.13) \quad \varepsilon_n \leq C(\delta_0 2^{-n-1})^{-1} \varepsilon_{n-1}^2.$$

Repeating (5.13), we obtain

$$\begin{aligned} \varepsilon_n &\leq C\delta_0^{-1} 2^{n+1} \varepsilon_{n-1}^2 \\ &\leq (C\delta_0^{-1}) 2^{n+1} (C\delta_0^{-1})^2 2^{2n} \varepsilon_{n-2}^{2 \cdot 2} \\ (5.14) \quad &\leq (C\delta_0^{-1})^{1+2+2^2+\dots+2^{n-1}} 2^{(n+1)+2(n)+2^2(n-1)+\dots+2^n} \varepsilon_0^{2^n} \\ &\leq (C\delta_0^{-1})^{2^n} 2^{2^{n+1} \sum_{k=1}^{n+1} k 2^{-k}} \varepsilon_0^{2^n} \\ &\leq (C\delta_0^{-1})^{2^n} 2^{2^{n+1} \sum_{k=0}^{\infty} k 2^{-k}} \varepsilon_0^{2^n} \\ &\leq (C\delta_0^{-1})^{2^n} 2^{2^{n+2}} \varepsilon_0^{2^n} \\ &\leq (C\delta_0^{-1} 2^2 \varepsilon_0)^{2^n}. \end{aligned}$$

Notice that if  $C\delta_0^{-1} 2^2 \varepsilon_0 < 1$ , then  $\varepsilon_n$  is superexponentially small. The fact that  $\varepsilon_n$  decreases superexponentially while  $\delta_n$  decreases only exponentially (5.12) has the consequence

that the inductive assumption (5.6) will be satisfied for all the iterative steps if the initial error  $\varepsilon_0$  is small enough.

Furthermore, denoting  $\mathcal{D}_{\beta_n, \rho_n}$  the domain of definition of  $K^{(n)}$  with  $(\beta_n, \rho_n) = (\beta_{n-1} - \delta_{n-1}, \rho_{n-1} - \delta_{n-1})$ , we have, using the estimates for each step and Cauchy estimates,

$$\begin{aligned}\|DK^{(n)} - DK^{(0)}\|_{\beta_n, \rho_n} &\leq \sum_{j=1}^n \|DK^{(j)} - DK^{(j-1)}\|_{\beta_j, \rho_j} \leq \sum_{j=1}^n C\varepsilon_j \delta_j^{-1} \\ &\leq \sum_{j=1}^n (C\delta_0^{-1}\varepsilon_0)^{2^j} \delta_0^{-1} 2^j.\end{aligned}$$

Similarly, we obtain

$$\|D^2K^{(n)} - D^2K^{(0)}\|_{\eta_n, \rho_n} \leq \sum_{j=1}^n (C\delta_0^{-1}\varepsilon_0)^{2^j} \delta_0^{-2} 2^{2j}.$$

Hence, we see that if  $\varepsilon_0$  is small enough, we obtain that the assumption that the change of  $m_+$  is small enough is satisfied.

Similarly, we see that the other smallness assumptions of the change are satisfied if  $\varepsilon_0$  is sufficiently small. We get, therefore, that the inductive assumptions amount to a finite number of smallness assumptions on  $\varepsilon_0$ .

Note also that, adding and subtracting terms and using (5.8), we have

$$\begin{aligned}(5.15) \quad \|K^{(0)} - K^{(\infty)}\|_{\beta_\infty, \rho_\infty} &\leq \sum_{n=0}^{\infty} \|K^{(n)} - K^{(n+1)}\|_{\beta_\infty, \rho_\infty} \\ &\leq \sum_{n=0}^{\infty} \|K^{(n)} - K^{(n+1)}\|_{\beta_{n+1}, \rho_{n+1}} \\ &\leq \sum_{n=0}^{\infty} C\varepsilon_n \leq C\varepsilon_0.\end{aligned}$$

The last inequality, of course, depends on  $\varepsilon_0\delta^{-1}$  being sufficiently small so that the superexponential convergence implies that the dominant term in the infinite sum above is the first one. Indeed, observe that the recurrence for the error (5.13) can be rewritten more transparently as

$$\varepsilon_n \leq (C(\delta_0 2^{-n-1})^{-1}\varepsilon_{n-1})\varepsilon_{n-1}.$$

Using (5.14) we obtain that for  $\varepsilon_0$  sufficiently small, we have  $C(\delta_0 2^{-n-1})^{-1}\varepsilon_{n-1} \leq 1/2$ . Hence, we can estimate the sums by a geometric series with an initial term  $\varepsilon_0$  and ratio  $1/2$ .

Analogous consideration leads to estimates of  $|\omega - \omega^*|$ ,  $|\lambda - \lambda^*|$ .

**Remark 5.6.** For the experts in KAM theory, we note that in our case, the size of the correction is bounded by the error without any factor from the loss of analyticity. The factors  $\delta^{-1}$  come only from the fact that the functional we are studying is not differentiable. In the regular KAM theory, the corrections at each step require a factor of the loss of differentiability. In both cases, we obtain that the total change in the function is bounded by a multiple of the first step. In the standard KAM case, this first step is the error times a power of the analyticity loss. In our case, the step is bounded by a step of the error.

**6. Some remarks and extensions of the analytic proof.** In this section we will show that the proof of Theorem 3.2 also leads to the conclusion that  $K^*$ ,  $\omega^*$ , and  $\lambda^*$  are locally unique (up to obvious choices of origins of coordinates and scaling factors in  $K$ ). Moreover, we will show that the solutions depend smoothly on parameters and that there are versions of the theorem for finitely differentiable vector fields.

### 6.1. Local uniqueness.

**Theorem 6.1.** *Let  $(K, \omega, \lambda)$  and  $(\tilde{K}, \tilde{\omega}, \tilde{\lambda})$  be solutions of (2.4) for the same vector field  $X$ . If*

$$(6.1) \quad \|K - \tilde{K}\|_{\beta, \rho}, |\omega - \tilde{\omega}|, |\lambda - \tilde{\lambda}| \leq C$$

for a constant  $C$  that depends on the condition numbers of the solution  $(K, \omega, \lambda)$ ,  $\beta$ , and  $\rho$ , then

$$(6.2) \quad \omega = \tilde{\omega}, \quad \lambda = \tilde{\lambda},$$

and there are  $\theta_0, b \in \mathbb{R}$ , such that

$$\tilde{K}(\theta, s) = K(\theta + \theta_0, bs).$$

**6.1.1. Proof of Theorem 6.1.** The proof is very similar to those of the local uniqueness results in other papers [26, 6] which also use automatic reducibility methods. The key observation is that the linearized equation admits a unique solution if we impose a normalization. In the language of abstract implicit function theorems, this is expressed as saying that the linearized equation admits a “left inverse.” For a discussion of this from an abstract point of view, we refer the reader to [6, Appendix A].

To overcome the ambiguity pointed out in section 2.2, we need to introduce a definition of normalized solutions.

**Definition 6.2.** *Given a solution  $(K, \omega, \lambda)$  of (2.4), we say that another embedding  $\tilde{K}$  is  $K$ -normalized when*

$$(6.3) \quad \begin{aligned} \int d\theta \Pi_1 \left[ (\tilde{K} - K) D K^{-1} \right]_{s=0} &= 0, \\ \int d\theta \Pi_2 \left[ D \tilde{K} D K^{-1} \right]_{s=0} &= 1, \end{aligned}$$

where  $\Pi_1, \Pi_2$  denote the projections over the first and second components, respectively.

The interpretation of (6.3) is that, when we express the difference between the solutions in the natural frame of reference of  $K$ , the first coordinate has average zero. Furthermore, the vector field representing the stable directions of the solutions has integral 1. These normalizations are natural since they eliminate the indeterminations we already identified in the solutions of (2.4), namely the change of the origin in the angle variable and the change of scale in the linear variable.

**Lemma 6.3.** *Assume there exists a constant  $C$  such that  $\|K - \tilde{K}\|_{\beta, \rho} \leq C$ . Then there exist small  $\theta_0$  and  $b$  close to 1 such that  $\tilde{K} \circ B_{\theta_0, b}$  is  $K$ -normalized, where  $B_{\theta_0, b}(\theta, s) = (\theta + \theta_0, bs)$ .*

*Proof.* The proof is similar to that of Lemma 14 in [26]. It is based on the application of the implicit function theorem to the function  $F^K(\theta_0, b)$  defined by evaluating (6.3) at  $K \circ B_{\theta_0, b}$ . That is,

$$(6.4) \quad \begin{aligned} F_1^K(\theta_0, b) &= \int d\theta \Pi_1 [(K \circ B_{\theta_0, b} - K)DK^{-1}]_{s=0}, \\ F_2^K(\theta_0, b) &= \int d\theta \Pi_2 [D(K \circ B_{\theta_0, b})DK^{-1}]_{s=0}. \end{aligned} \quad \blacksquare$$

Hence, to prove Theorem 6.1, it suffices to show that if we have two solutions of (2.4), namely  $(K, \omega, \lambda)$  and  $(\tilde{K}, \tilde{\omega}, \tilde{\lambda})$ , and that  $\tilde{K}$  is  $K$ -normalized, then they are equal. To do so, it will be useful to introduce a more abstract point of view similar to that in [6]. We introduce the notation

$$(6.5) \quad \mathcal{T}(K, \omega, \lambda) \equiv X \circ K - DKA_{\omega, \lambda}$$

so that (2.4) can be written as

$$(6.6) \quad \mathcal{T}(K, \omega, \lambda) = 0.$$

If we are given two solutions  $(K, \omega, \lambda)$  and  $(\tilde{K}, \tilde{\omega}, \tilde{\lambda})$  of (6.6), we can write, using Taylor's theorem from one to the other,

$$(6.7) \quad \begin{aligned} 0 &= \mathcal{T}(\tilde{K}, \tilde{\omega}, \tilde{\lambda}) \\ &= \mathcal{T}(K, \omega, \lambda) + D\mathcal{T}(K, \omega, \lambda)[\tilde{K} - K, \tilde{\omega} - \omega, \tilde{\lambda} - \lambda] + R \\ &= D\mathcal{T}(K, \omega, \lambda)[\tilde{K} - K, \tilde{\omega} - \omega, \tilde{\lambda} - \lambda] + R, \end{aligned}$$

where  $R$  is the remainder of the Taylor expansion of the functional  $\mathcal{T}$ . The fact that  $\mathcal{T}$  is differentiable and the form of the derivative have been established in Lemma 5.4.

Note that  $\mathcal{T}$  involves only composing  $X$  on the RHS with  $K$ , taking derivatives of  $K$  and performing some algebraic operations. Hence, we have that

$$(6.8) \quad \|R\|_{\rho-\delta} \leq C\delta^{-1}(\|\tilde{K} - K\|_\rho^2 + |\tilde{\omega} - \omega|^2 + |\tilde{\lambda} - \lambda|^2).$$

The identity (6.7) relates the increments in the unknown to the Taylor estimates in exactly the same way that the corrections of the Newton method were related to the error. We can regard (6.7) as an equation for  $(\tilde{K} - K, \tilde{\omega} - \omega, \tilde{\lambda} - \lambda)$ . Using that these equations have unique solutions (because  $\tilde{K}$  is  $K$ -normalized), we have

$$(6.9) \quad \begin{aligned} \|\tilde{K} - K\|_{\rho-2\delta}, |\tilde{\omega} - \omega|, |\tilde{\lambda} - \lambda| &\leq C\delta^{-1}\|R\|_{\rho-\delta} \\ &\leq C\delta^{-2}(\|\tilde{K} - K\|_\rho^2 + |\tilde{\omega} - \omega|^2 + |\tilde{\lambda} - \lambda|^2). \end{aligned}$$

Using Hadamard's three circle theorem (5.4) with  $\alpha = 1/2$ ,  $\rho_1 = \rho + 2\delta$ , and  $\rho_2 = \rho - 2\delta$ , we obtain

$$\|\tilde{K} - K\|_\rho^2 \leq \|\tilde{K} - K\|_{\rho+2\delta}\|\tilde{K} - K\|_{\rho-2\delta}.$$

Hence,

$$(6.10) \quad \begin{aligned} & \|\tilde{K} - K\|_{\rho-2\delta} + |\tilde{\omega} - \omega| + |\tilde{\lambda} - \lambda| \\ & \leq C\delta^{-2}(\|\tilde{K} - K\|_{\rho+2\delta} + |\tilde{\omega} - \omega| + |\tilde{\lambda} - \lambda|) \\ & \cdot (\|\tilde{K} - K\|_{\rho-2\delta} + |\tilde{\omega} - \omega| + |\tilde{\lambda} - \lambda|). \end{aligned}$$

Therefore, if  $C\delta^{-2}(\|\tilde{K} - K\|_{\rho+2\delta} + |\tilde{\omega} - \omega| + |\tilde{\lambda} - \lambda|) < 1$ , we conclude that  $K = \tilde{K}$ ,  $\omega = \tilde{\omega}$ , and  $\lambda = \tilde{\lambda}$ . The statement of Theorem 6.1 is obtained just by redefining  $\rho$ .

**6.2. Dependence on parameters.** In many applications, the models depend on extra parameters. We will show how the automatic reducibility methods used in the proof of Theorem 3.2 lead to very efficient computations of the perturbative expansions with respect to these parameters. We will also show, following [32], that these perturbative expansions converge.

**6.2.1. Lipschitz dependence on parameters.** If we consider a family of vector fields  $X_\mu$  and we assume Lipschitz dependence of the vector field with respect to the parameter  $\mu$ , we can use Theorem 3.2 to obtain Lipschitz dependence of the solution with respect to the parameter  $\mu$ .

Following the notation introduced in (6.5), denote

$$\mathcal{T}_\mu(K, \omega, \lambda) \equiv X_\mu \circ K - DKA_{\omega, \lambda},$$

so that a solution  $(K_\mu, \omega_\mu, \lambda_\mu)$  of (2.4) for the vector field  $X_\mu$  can be written as

$$(6.11) \quad \mathcal{T}_\mu(K_\mu, \omega_\mu, \lambda_\mu) = 0.$$

Consider  $(K_\mu, \omega_\mu, \lambda_\mu)$  satisfying (6.11). Then, using that  $X_\mu$  is Lipschitz dependent with respect to parameter  $\mu$ , we clearly have

$$\begin{aligned} \|\mathcal{T}_{\tilde{\mu}}(K_\mu, \omega_\mu, \lambda_\mu)\|_{\beta-\delta, \rho-\delta} &= \|\mathcal{T}_{\tilde{\mu}}(K_\mu, \omega_\mu, \lambda_\mu) - \mathcal{T}_\mu(K_\mu, \omega_\mu, \lambda_\mu)\|_{\beta-\delta, \rho-\delta} \\ &= \|X_{\tilde{\mu}} \circ K_\mu - X_\mu \circ K_\mu\|_{\beta-\delta, \rho-\delta} \leq C\delta^{-1}|\mu - \tilde{\mu}|. \end{aligned}$$

So, we have that  $(K_\mu, \omega_\mu, \lambda_\mu)$  is as an approximate solution of (6.11) for values of the parameter  $\mu$  close to the original one.

Hence, we can apply Theorem 3.2 and obtain that for  $|\mu - \tilde{\mu}|$  sufficiently small, there exists  $K_{\tilde{\mu}}$  an analytic local diffeomorphism and  $\omega_{\tilde{\mu}}, \lambda_{\tilde{\mu}} \in \mathbb{R}$  such that

$$\|K_\mu - K_{\tilde{\mu}}\|_{\beta-2\delta, \rho-2\delta}, |\omega_\mu - \omega_{\tilde{\mu}}|, |\lambda_\mu - \lambda_{\tilde{\mu}}| \leq C\delta^{-1}|\mu - \tilde{\mu}|.$$

**6.2.2. Computation of perturbative expansions on parameters.** Consider a parametric family of vector fields  $X_\mu$  as well as a solution  $(K_0, \omega_0, \lambda_0)$  of (6.11) for  $\mu = 0$ . We want to compute a formal solution of (6.11) for  $\mu \neq 0$  by considering asymptotic expansions on the parameter  $\mu$ ,

$$(6.12) \quad \begin{aligned} K_\mu &= \sum_n \mu^n K_n, \\ \omega_\mu &= \sum_n \mu^n \omega_n, \quad \lambda_\mu = \sum_n \mu^n \lambda_n. \end{aligned}$$

We discuss two different methods for computing the asymptotic expansions efficiently. The convergence will be discussed in the next subsection.

**An order-by-order method.** We assume inductively that we have computed the expansion (6.12) up to order  $n - 1$ , and we want to show that it is possible to compute the expansion up to order  $n$ .

We substitute (6.12) into (2.4) and by matching coefficients of order  $n$ , we obtain

$$(6.13) \quad (DX_\mu \circ K_0)K_n - DK_n A_{\omega_0, \lambda_0} - DK_0 A_{\omega_n, \lambda_n} = S_n(K_0, \dots, K_{n-1}),$$

where  $S_n$  is an explicit polynomial expression in  $K_0, \dots, K_{n-1}$  whose coefficients are derivatives of  $X_\mu$  evaluated at  $K_0$ . These coefficients can be calculated efficiently using the methods of automatic differentiation when  $X_\mu$  is formed using elementary functions (polynomials, exponentials, trigonometric functions, etc.).

We observe that (6.13) are identical to (4.2), the equation we studied in section 4.2. Hence, we can use the same method used there with some minor differences that we discuss next.

Note that because  $K_0$  satisfies exactly the invariance equation (2.4), the factorization of (6.13) into elementary steps achieved in Algorithm 4.4 holds exactly. Furthermore, all the auxiliary quantities involved in the factorization need to be computed only once because for all steps we consider only the linearization around  $K_0$ , which does not change during the iteration.

**A quadratically convergent method.** A faster method for computing the perturbative expansions (6.12) consists in considering  $K(\theta, s, \mu)$ ; that is,  $K$  is a function of the parameter  $\mu$ . It is easy to see that Algorithm 4.4 lifts to functions of three variables and that one can also obtain quadratic convergence in the space of functions in these three variables using the argument in section 5.

**6.2.3. Convergence of perturbative expansions.** Convergence of perturbative expansions (6.12) is guaranteed by Theorem 3.2. It suffices to take the solution of (6.11) for  $\mu = 0$  as an approximate solution of (6.11), and by Theorem 3.2 we have that there exists a solution  $(K_\mu, \omega_\mu, \lambda_\mu)$  of (6.11) for  $\mu$  small and complex. Then, using Lemma 6.3 we can assume that the solutions are  $K_0$ -normalized in the sense of Definition 6.2. We also know that functions  $(K_\mu, \omega_\mu, \lambda_\mu)$  are differentiable for  $\mu$  small and complex. Hence, they are analytic in  $\mu$ .

**6.3. Finite differentiability.** There is a standard procedure in [31, 30], systematized and extended in [48], that shows that one can deduce results for finite differentiable problems from quantitative results such as Theorem 3.2 for analytic problems.

The key point is the characterization of finitely differentiable functions by the speed of approximation by analytic functions summarized in the following lemma.

**Lemma 6.4.** *A function  $f : \mathbb{T}^d \times B^l$  is  $r$  times continuously differentiable,  $r \in \mathbb{N}$ , and the  $r$  derivative is Hölder continuous with exponent  $\alpha$ ,  $0 < \alpha < 1$ , if and only if we can find a sequence of functions  $f_n$ , each of them analytic in a complex extension of size  $\rho_n = 2^{-n}$ , such that*

- $\|f_n - f_{n-1}\|_{\rho_n} \leq C 2^{-n(r+\alpha)}$ ,
- $\|f_n - f\|_{C^0} \rightarrow 0$ .

A streamlined proof of Lemma 6.4 can be found in [48]. It is also well known that the characterization given by Lemma 6.4 is false for  $\alpha = 0, 1$ .

Notice that (2.4) is linear in  $X$ . If  $X$  is  $C^{r+\alpha}$ , we can construct a sequence  $X_n$  which is analytic in decreasing domains. If  $(K_0, \omega_0, \lambda_0)$  is analytic and solves (2.4) for  $X_0$  with a good enough approximation, we can apply Theorem 3.2 and construct a true solution  $(K_1, \omega_1, \lambda_1)$  of (2.4) for  $X_0$ . This will be an approximate solution of (2.4) for  $X_1$ ; then, applying Theorem 3.2, we can construct an exact solution  $(K_2, \omega_2, \lambda_2)$  of the problem for  $X_1$ , which will be an approximate solution for the problem for  $X_2$ , etc.

In general, under appropriate inductive assumptions in the domain, we have that

$$\|X_n \circ K_n - DK_n A_{\omega_n, \lambda_n}\|_{\rho_n} \leq C \|X_n - X_{n-1}\|_{\rho_n} \leq C 2^{-(r+\alpha)n}.$$

Applying Theorem 3.2, we obtain that

$$\|K_{n+1} - K_n\|_{\rho_{n+1}}, |\omega_{n+1} - \omega_n|, |\lambda_{n+1} - \lambda_n| \leq C 2^{-(r+\alpha-1)n}.$$

Hence, we conclude that  $K$  is  $C^{r-1+\alpha}$ .

**7. Numerical implementation.** In this section we discuss some aspects of the numerical implementation.

**7.1. Implementation of Algorithm 4.4.** In this section we discuss several possible ways to implement Algorithm 4.4. The implementations require choices on the discretization of functions and on the ways to perform the elementary operations (algebraic operations, composition, derivatives, integrals, etc.). Many practical properties of the algorithm depend on these choices, among them

- (a) storage requirements,
- (b) speed,
- (c) accuracy,
- (d) stability, and
- (e) parallelizability.

Next we discuss several possible discretizations and the numerical properties of each of them.

**7.1.1. Fourier–Taylor series.** This is a very well established method in celestial mechanics (see [2, 9, 40] for classical implementations and [16, 20] for more modern implementations).

In this representation, one stores the Fourier–Taylor coefficients of a function

$$(7.1) \quad f(\theta, s) = \sum_{j \in \mathbb{N}, k \in \mathbb{Z}} f_{jk} s^j e^{2\pi i k \theta}.$$

For the discretization (7.1), steps 6 and 7 of Algorithm 4.4 are diagonal (see expressions (4.11) and (4.13) in Lemma 4.2). Therefore,  $N$  Fourier–Taylor coefficients require only  $O(N)$  operations and  $O(N)$  storage. Similarly, the computation of the derivative in step 2, as well as addition and multiplication of functions by numbers, also require only  $O(N)$  operations.

The most difficult calculation is the computation of  $X \circ K$  in step 1. When  $X$  can be obtained applying elementary operations (addition, multiplication, trigonometric functions,

exponentials, etc.) there is a well-defined toolkit that goes under the name of *automatic differentiation*.

It consists of a set of techniques for computing the derivatives of arbitrary order of a function evaluated at a given point, accurate to working precision, avoiding in this way the numerical problems inherent in symbolic and numerical differentiation. They are based on writing the function as a sequence of algebraic operations (sum, product, etc.) and elementary transcendental functions ( $\exp$ ,  $\sin$ ,  $\cos$ ,  $\log$ , power, etc.) and then systematically applying the chain rule to these operations (see [23] and also the Web page of the automatic differentiation community <http://www.autodiff.org/>).

Take, for instance, the case of the exponential function. Consider the Taylor expansion of a function  $f$  in the variable  $s$ ,

$$f(\theta, s) = \sum_{j \in \mathbb{N}} f_j(\theta) s^j,$$

and suppose that we want to find the Taylor expansion of the function  $\exp(f)$ :

$$\exp(f)(\theta, s) = \sum_{j \in \mathbb{N}} [\exp(f)]_j(\theta) s^j.$$

If we apply the chain rule to  $\exp(f)$ , we obtain the relation

$$\partial_s \exp(f) = \exp(f) \partial_s f.$$

Substituting  $f$  by its Taylor expansion and equating terms of order  $n$  we obtain the recursive expression

$$(7.2) \quad (n+1)[\exp(f)]_{n+1}(\theta) = \sum_{\ell=0}^n [\exp(f)]_{n-\ell}(\theta) (\ell+1) f_{\ell+1}(\theta).$$

Note that (7.2) allows us to compute  $[\exp(f)]_{n+1}(\theta)$  provided that we know  $[\exp(f)]_0(\theta), \dots, [\exp(f)]_{n-1}(\theta)$ . The recursion can be initialized because  $[\exp(f)]_0$  is just the constant  $\exp(f_0)$ .

Similar algorithms can be obtained for  $\sin f$ ,  $\cos f$ ,  $\log(1+f)$ , and  $f^\alpha$ , or indeed for any function of  $f$  that satisfies a differential equation or some recurrence on the coefficients.

Finally, one can apply similar algorithms to compute the Fourier series of the Taylor coefficients  $[\exp(f)]_j(\theta)$  (think of them as the sum of two polynomials in  $e^{2\pi i\theta}$  and  $e^{-2\pi i\theta}$ ). Hence, one can use the previous algorithm to compute

$$\begin{aligned} & \exp \left( \sum_{j \geq 0} f_k (e^{2\pi i\theta})^k \right), \\ & \exp \left( \sum_{j \geq 0} f_k (e^{-2\pi i\theta})^k \right) \end{aligned}$$

and then use the addition formula for the exponentials.

**7.1.2. Fourier-real mixed representation.** A variant of the Fourier–Taylor representation which has proved very useful is to keep at the same time both a representation based on the Fourier–Taylor coefficients and a discrete representation in real space

$$f(j/N, e^{2\pi i \ell/N})$$

for  $j = 0, \dots, N - 1$  and  $\ell = 0, \dots, N - 1$ .

Since we keep both the Fourier–Taylor representation and the discrete representation, we can use the representation that makes the computation faster in each step of the algorithm. Thus, steps 2—differentiation—and 6 and 7—solving the cohomology equations—of Algorithm 4.4 are diagonal in Fourier–Taylor representation, while step 1—evaluation of the vector field—is diagonal in the discrete representation. Of course, the FFT algorithm allows us to switch from real space to Fourier space in  $\mathcal{O}(N \log N)$  computations.

Note that Taylor–Fourier series can be considered also as Fourier series in two variables. Given a function  $f(\theta, s) = \sum f_n(\theta) s^n$ , when we consider  $f(\theta, e^{2\pi i s})$ , it becomes just a Fourier series. Hence, it is possible to compute the composition  $X \circ K$  in step 1 by evaluating  $X$  on the discretization  $K(\theta_j, e^{2\pi i s_k})$  and using the Fast Fourier transform (FFT) to obtain the Fourier series of  $X \circ K$  in the variables  $\theta$  and  $s$ .

An alternative which requires a higher operation count but which is customary in celestial mechanics [2, 40] is to deal with operations among periodic functions using FFT methods but to use automatic differentiation or the customary methods of polynomial manipulation (e.g., the Cauchy product formula to deal with the products).

Going back to the example discussed in the previous section, we can compute the Fourier expansions of Taylor coefficients  $[\exp(f)]_j(\theta)$  just evaluating the exponential function on the points  $\theta_j = j/N$  for  $j = 0, \dots, N - 1$  and then applying the FFT to obtain the Fourier coefficients.

In this way, if we discretize in  $M$  Fourier modes and  $L$  degree polynomials, we have  $N = M L$  and the operation count of one step is  $\mathcal{O}(M \log M L^2)$  (slightly higher than the  $\mathcal{O}(N \log N)$  obtained by Fourier methods).

In practice, for most computers, one can find highly optimized implementations of the FFT, for example in [14], so that the algorithm is  $\mathcal{O}(aN \log N + bN)$  with  $a \ll b$ .

**7.1.3. Splines/Chebyshev polynomials.** When the vector field is not analytic or is given by empirical measurements, a method of choice for discretizing the vector field and the function  $K$  is to use splines [8]. By now, splines, including multidimensional splines, are well supported in many packages [15, 10].

The discretization in terms of splines takes  $\mathcal{O}(N)$  operations to evaluate the vector field, compute derivatives, etc. However, the evaluation of splines makes nontrivial the computation of solutions of cohomology equations (4.10) and (4.12). In this case, it is more efficient to use solutions (4.14) and (4.15), which require only quadratures. We note that if the spline representation (by polynomials of low order) is known, the quadratures can be computed in closed form and evaluated rather fast (again only  $\mathcal{O}(N)$  operations).

Thus, using the spline discretization, a step requires only  $O(N)$  operations. Furthermore, given that the operations required for splines are local, they can be easily parallelized, especially in machines with multiple cores. Hence, splines seem to be extremely fast, even for a

large number of data points.

An alternative to the traditional splines that is of recent interest is the use of Chebyshev functions, which are in some way close to the Fourier series. See [34].

**7.1.4. Numerical norms.** Depending on the discretizations used, the norms that are appropriate may be different.

If we use splines, the error can be measured easily in  $C^r$  spaces. Analytic norms are not appropriate for splines since the functions involved are not analytic.

If we use Fourier series, the norms that are easy to compute are those that can be expressed in terms of Fourier series. Amongst the most effective norms in Fourier analysis are the weighted  $\ell^1$  norms because for them it is easy to compute the norm of operators in terms of matrix elements. For example,  $\|f\|_{w\ell^1} = \sum_k |k|^n |\hat{f}_k| e^{\rho|k|}$  is a norm that has many advantages: it is easy to compute reliably, it is easy to compute for operators, and it satisfies the Banach algebra property for  $n$  large enough.

For the experts we remark that one could have developed the theoretical results such as Theorem 3.2 in terms of weighted  $\ell^1$  norms of the Fourier coefficients, but it turns out that estimates of the composition operator are not so easy. Also, the characterization of finite differentiable functions by approximation is only true in the supremum norms considered here.

In practice, one can get useful upper bounds of  $\|f\|_\rho$  by noting that  $\|f\|_\rho \leq \|f\|_{w\ell^1}$  or, more generally,

$$(7.3) \quad \|f\|_\rho \leq \left( \|f^k\|_{w\ell^1} \right)^{1/k}.$$

The bounds (7.3) are very easy to implement, and they are very sharp in practice. Indeed, it is a consequence of the theory of Banach algebras [38, Thm. 18.9] that, for any norm which is a Banach algebra under multiplication, one has

$$\|f\|_\rho = \lim_{k \rightarrow \infty} \left( \|f^k\|_{w\ell^1} \right)^{1/k}.$$

**7.2. Other aspects of the numerical implementation.** Algorithm 4.4 has been incorporated in a more general program that computes the globalization of isochrons. In this section we discuss other details of the program.

**Initial guess.** To apply the Newton method we need to start with an initial approximation for the function  $K$  and the parameters  $\lambda$  and  $\omega$ . To do so, we will use a Poincaré section and reduce the problem to finding a zero of the Poincaré map. This will provide  $K_0$  (the periodic orbit) and  $\omega = 1/T$ , where  $T$  is the period of the orbit. A reasonable approximation for  $K_1$  and  $\lambda$  can be obtained by observing that

$$DX \circ K_0(\theta) K_1(\theta) = \omega \frac{d}{d\theta} K_1(\theta) + \lambda K_1(\theta).$$

Hence, we can solve for  $U(\theta)$  the variational equation

$$(7.4) \quad \begin{aligned} DX \circ K_0(\theta) U(\theta) &= \omega \frac{d}{d\theta} U(\theta), \\ U(0) &= \text{Id}_2, \end{aligned}$$

and we will have that  $e^{\lambda/\omega}$  will be an eigenvalue of  $U(1)$  and  $K_1(0)$  will be the corresponding eigenvector.

Note that once we obtain the result for a point, say  $\theta = 0$ , it is easy to propagate using that

$$(7.5) \quad K_1(\theta) = U(\theta)K_1(0)e^{-\lambda\theta/\omega}.$$

Our initial approximation for  $K$  will be  $K(\theta, \sigma) = K_0(\theta) + K_1(\theta)\sigma$ . We store  $K_0(\theta)$  and  $K_1(\theta)$  for equidistant values of  $\theta$ ; that is,  $\theta_j = j/N$  for  $j = 0, \dots, N - 1$ . Notice that this is equivalent to storing the coefficients of the Fourier series up to degree  $N$  by means of the FFT algorithm.

**Newton step.** We use a Fourier–Taylor and mixed representation to implement Algorithm 4.4. See section 7.1 for a detailed description. Since  $X$  can be obtained by applying elementary operations, we use automatic differentiation methods to perform the Taylor expansions as described in section 7.1.1. In order to pass from a grid representation to Fourier series and vice versa we use the FFT. In our program we have used the fftw3 library [14].

At each step of the Newton method we double the order of the Taylor series, so that after  $n$  Newton steps we have computed the Fourier–Taylor series up to order  $L = 2^n$ . The Newton method stops when the solution has been computed up to an error of order  $10^{-11}$  up to the desired order  $L$ . The norm we use to estimate the error is the  $\ell_1$  norm.

**Local approximation.** Up to this point we assume that we have converged to an approximate solution  $K$  of the invariance equation. We need to determine the domain  $\Omega_{loc}$  where the solution  $K$  is accurate; that is, the function  $K$  satisfies the invariance equation up to a certain tolerance  $E$  that we established at  $10^{-10}$ .

Given a fixed tolerance  $E$  we compute

$$(7.6) \quad \Omega_{loc} := \{(\theta, \sigma) \in \mathbb{T} \times \mathbb{R} \mid \|X(K(\theta, \sigma)) - DK(\theta, \sigma)A_{\omega, \lambda}\| < E\},$$

where  $\|\cdot\|$  is a norm in  $\mathbb{R}^2$ . We remark that  $\Omega_{loc}$  contains the limit cycle  $\gamma$ . Notice that the higher the order  $L$  the larger the  $\Omega_{loc}$ . We compute the local isochron of  $K_0(\theta)$  by fixing  $\theta = \theta_0$  and evaluating the function  $K$  on  $(\theta_0, s) \in \Omega_{loc}$ .

**Globalization of the isochrons.** Since the flow of the vector field  $X$  takes isochrons to isochrons, we can obtain several points on the isochron of phase  $\theta$ ,  $\mathcal{S}_\theta$ , by integrating backwards for a time  $t$  points on the isochron of phase  $\theta + \omega t$  obtained from the local approximation:

$$\mathcal{S}_\theta = X^{-t}(\mathcal{S}_{\theta + \omega t}).$$

Note that if  $\mathcal{S}_{\theta + \omega t}$  is known as a small curve, for  $t > 0$ , the isochron  $\mathcal{S}_\theta$  is much longer. Hence, this procedure extends the isochron to a larger domain  $\Omega$ . We refer to this procedure as the globalization of isochrons.

We use the procedure described in [19], which follows the numerical method proposed in [42]. The main idea is to select a nonuniform mesh of points on the isochron so that the globalization procedure provides dense points on the isochron.

The integration method used is a Taylor method (we have used the routines provided by Jorba and Zou; see [22]). We used an adaptive step size, degree, and tolerance (absolute and relative) of order  $10^{-16}$ .

**8. Numerical examples.** We present here the application of the numerical method described in the previous section to some relevant examples in the literature.

**Rayleigh oscillator.** It is a Van der Pol-type oscillator [44], and it was introduced by Rayleigh to show the appearance of sustained vibrations in acoustics [36]. The equations have the form

$$\begin{aligned}\dot{x} &= -y + \mu(x - x^3), \\ \dot{y} &= x.\end{aligned}$$

The system is analytic, has an unstable fixed point (focus) at  $(0, 0)$ , and for  $\mu = 1$  has a stable limit cycle of period  $T = 6.663$  and characteristic exponent  $\lambda = -1.059$ . We show the isochrons in Figures 1 and 2.

**Morris–Lecar model.** It was initially conceived as a model for an electrically excitable barnacle muscle [29], but well studied in the neuroscience literature, after [37], as a paradigm for different types of neuronal excitability. The model takes the form

$$(8.1) \quad \begin{aligned}C\dot{V} &= I_{app} - g_L(V - V_L) - g_K\omega(V - V_K) - g_{Ca}m_\infty(V)(V - V_{Ca}), \\ \dot{w} &= \phi \frac{w_\infty(V) - w}{\tau_w(V)},\end{aligned}$$

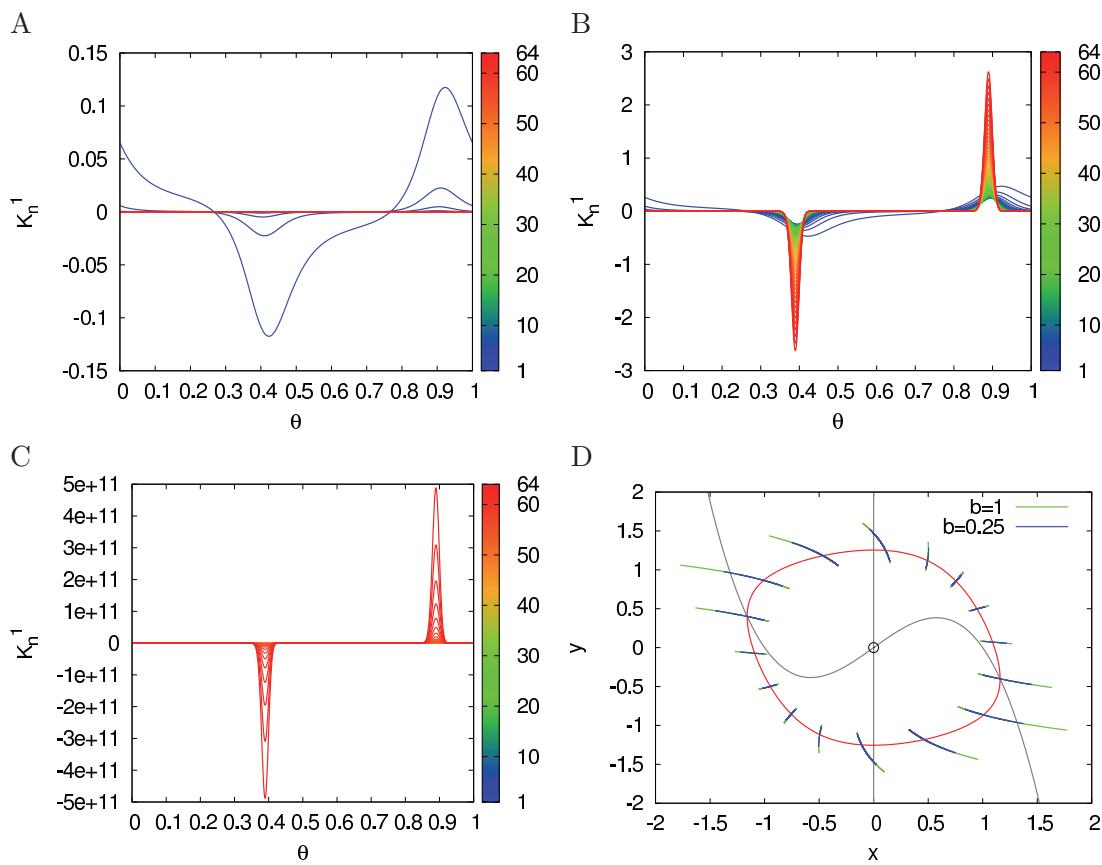
where

$$\begin{aligned}m_\infty(V) &= \frac{1}{2}(1 + \tanh((V - V_1)/V_2)), \\ w_\infty(V) &= \frac{1}{2}(1 + \tanh((V - V_3)/V_4)), \\ \tau_w(V) &= (\cosh((V - V_3)/(2V_4)))^{-1}.\end{aligned}$$

We consider three sets of parameters that correspond to three different types of bifurcations from steady state to periodic behavior as  $I_{app}$  varies: Hopf, saddle-node on an invariant circle (SNIC), and homoclinic (see [13]). We will refer to them as ML-Hopf, ML-SNIC, and ML-Hom, respectively. Some parameters are common for the three types of bifurcations. Namely,  $C = 20$ ,  $V_L = -60$ ,  $V_K = -84$ ,  $V_{Ca} = 120$ ,  $V_1 = -1.2$ ,  $V_2 = 18$ ,  $g_L = 2$ ,  $g_K = 8$ . The rest of the parameters are listed in Table 1. For each bifurcation type we consider a value of the parameter  $I_{app}$  for which the system is in the oscillatory regime but “close” to the critical value. The value of  $I_{app}$  that we considered for each set of parameters, as well as the period and the characteristic exponent of the limit cycle, are indicated in Table 1. We show the isochrons in Figures 3, 4, and 5.

**Reduced ( $I_{Na,p} + I_K$ )-model.** This model is somehow equivalent to the Morris–Lecar model. It was introduced in [21] and describes a fast persistent sodium current and a slower persistent potassium current:

$$(8.2) \quad \begin{aligned}C\dot{V} &= I_{app} - g_{Na}m_\infty(V)(V - V_{Na}) - g_Kn(V - V_K) - g_L(V - V_L), \\ \dot{n} &= n_\infty(V) - n,\end{aligned}$$



**Figure 1.** The Rayleigh oscillator. First component  $x$  of the functions  $K_n(\theta)$  for  $n = 1 \dots 64$  colored according to the corresponding order  $n$  for (A)  $b = 0.25$ , (B)  $b = 1$ , and (C)  $b = 1.5$ . (D) 16 local isochrons uniformly distributed in time along the limit cycle (red) computed using the expansions up to  $n = 64$  shown in panels (A) and (B) corresponding to  $b = 0.25$  (blue curves) and  $b = 1$  (green curves), respectively. The  $x$ - and  $y$ -nullclines (grey) and the fixed point (open circle) at  $(0, 0)$  (unstable focus) are included for reference.

where the open-state probability functions are

$$m_\infty(V) = \frac{1}{1 + \exp(-(V - V_{max,m})/k_m)},$$

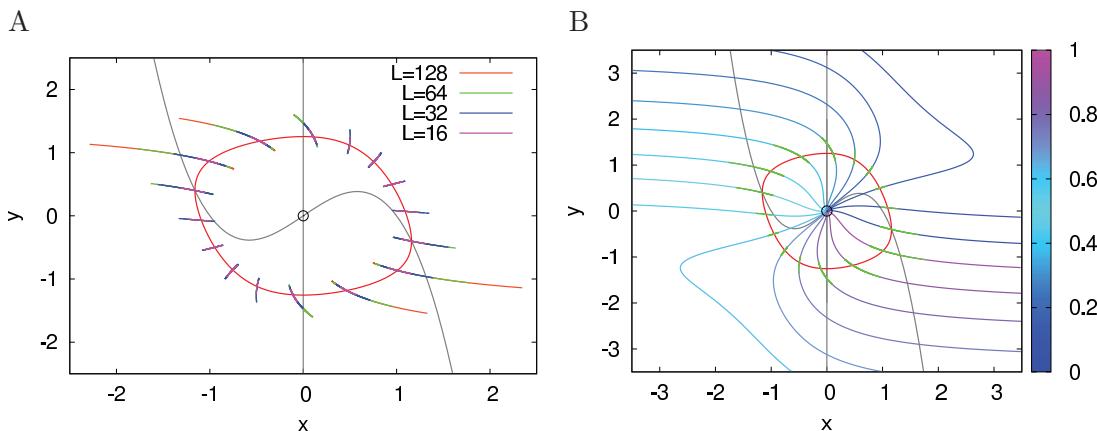
$$n_\infty(V) = \frac{1}{1 + \exp(-(V - V_{max,n})/k_n)},$$

and the parameters are  $C_m = 1$ ,  $g_{Na} = 20$ ,  $V_{Na} = 60$ ,  $g_K = 10$ ,  $V_K = -90$ ,  $g_L = 8$ ,  $v_L = -80$ ,  $V_{max,m} = -20$ ,  $k_m = 15$ ,  $V_{max,n} = -25$ ,  $k_n = 5$ .

For  $I_{app} = 10$ , the system has a stable limit cycle of period  $T = 7.0735$  and characteristic exponent  $\lambda = -3.91$  that has a strong slow-fast dynamics. The limit cycle encircles an unstable fixed point (focus) at  $(-26.83, 0.41)$ . We show the isochrons in Figure 6.

### 8.1. Numerical examples; Some considerations.

We have chosen these examples to em-



**Figure 2.** The Rayleigh oscillator. (A) 16 local isochrons uniformly distributed in time along the limit cycle (red) computed using the functions  $K_n$  up to order  $L = 16, 32, 64$ , and 128 with  $b = 1$ . (B) Globalization of the local isochrons (green curves) corresponding to  $L = 64$  in (A). Global isochrons are colored according to their phases. The  $x$ - and  $y$ -nullclines are shown in gray, and the fixed point at  $(0, 0)$  corresponding to an unstable focus is represented by an open circle.

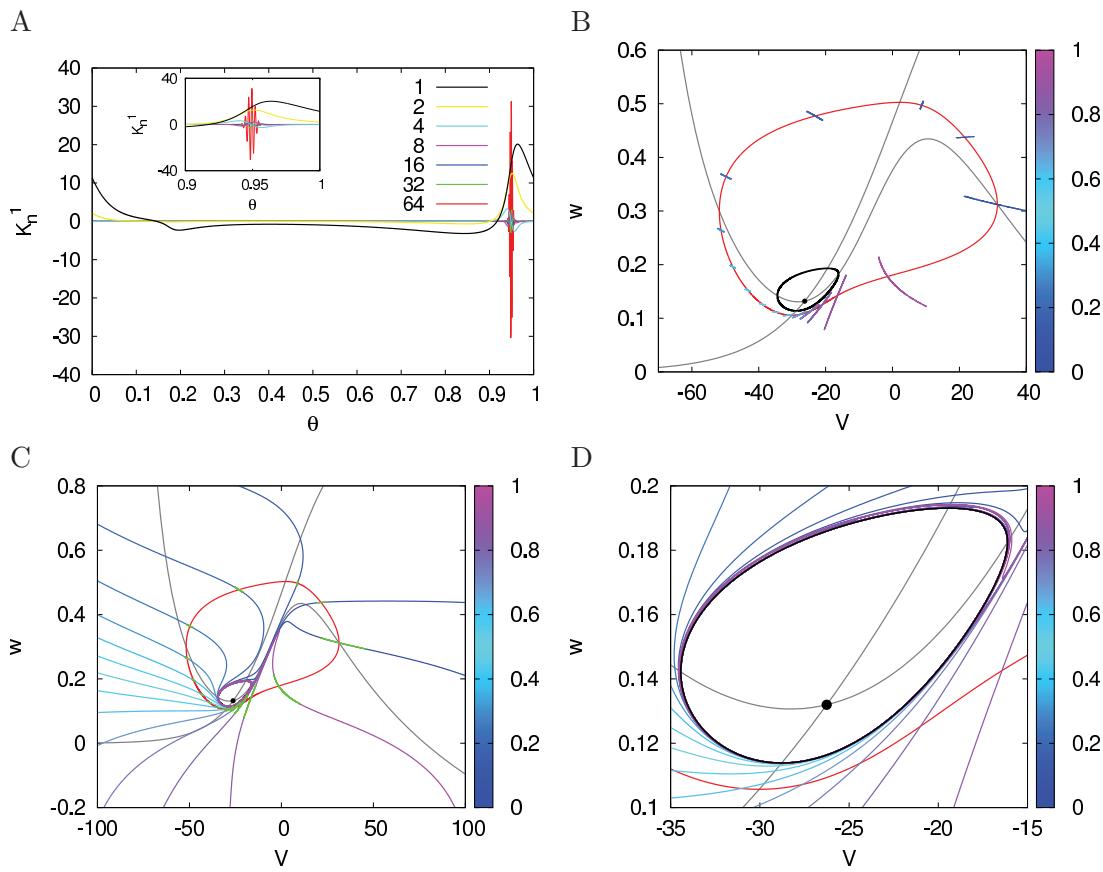
**Table 1**

Parameter values for the Morris–Lecar model for the corresponding type of bifurcation. For the values of  $I_{app}$  indicated the system has a stable periodic orbit of period  $T$  and characteristic exponent  $\lambda$ .

| Parameter | Hopf    | SNIC    | Homoclinic |
|-----------|---------|---------|------------|
| $\phi$    | 0.04    | 0.067   | 0.23       |
| $g_{Ca}$  | 4.4     | 4       | 4          |
| $V_3$     | 2       | 12      | 12         |
| $V_4$     | 30      | 17.4    | 17.4       |
| $I_{app}$ | 91      | 45      | 37         |
| $T$       | 99.27   | 99.192  | 33.94      |
| $\lambda$ | -0.0919 | -0.1198 | -0.06      |

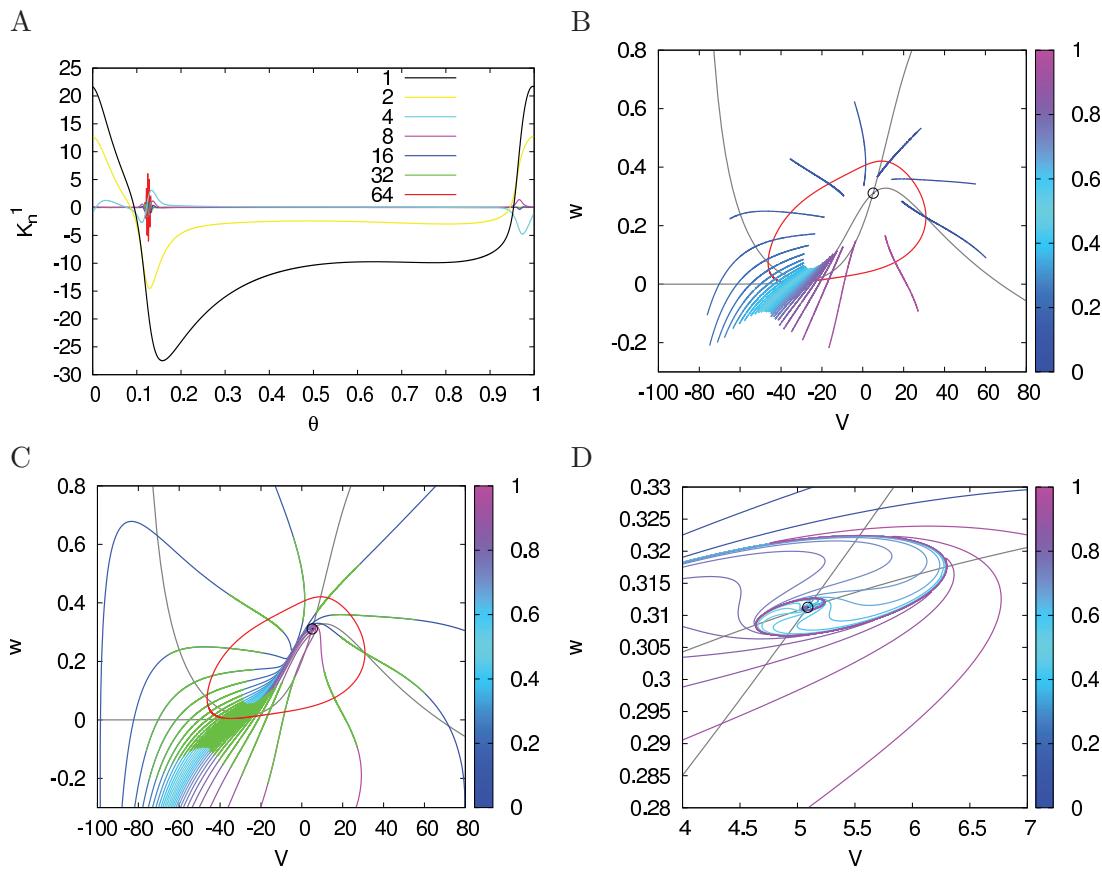
phasize certain aspects of the algorithm. The Rayleigh oscillator is quite simple, and it serves as a good first test for our algorithm. We use it to discuss the effect in the local approximation of increasing the order  $L$  of the Taylor expansion of  $K(\theta, s)$ , as well as the role of the parameter  $b$  (see section 2.2) in the numerical computation. The other two examples are models from computational neuroscience. The Morris–Lecar model with three different sets of parameters has been chosen to illustrate the isochrons in different types of bifurcations and also in the presence of other invariant objects that confine the region where the isochrons are defined. The  $(I_{Na,p} + I_K)$ -model has been chosen to illustrate how the algorithm performs in the case of a strong time-scale separation between the two variables.

For the Rayleigh oscillator, we have used  $N = 2^{11} = 2048$  Fourier modes for each  $K_n$ , for the Morris–Lecar models we have used  $N = 2^{12} = 4096$  Fourier modes, and for the  $(I_{Na,p} + I_K)$ -model we have used  $N = 2^{13} = 8192$  Fourier modes. We have computed Taylor expansions up to order  $L = 2^6 = 64$  for all of them. The program takes only a few seconds (around 10s) on a regular laptop to compute the local approximation  $K$ .



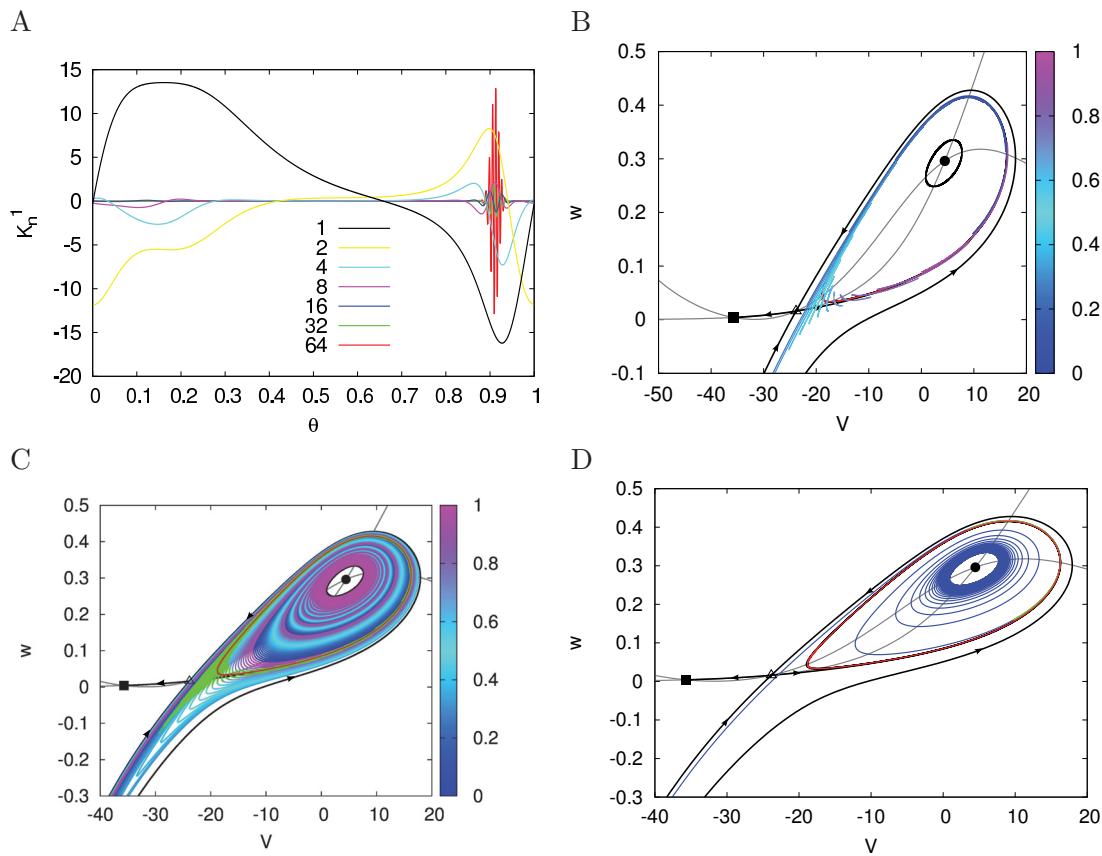
**Figure 3.** Morris–Lecar near a Hopf bifurcation. Set of parameters as in Table 1 under the Hopf column. (A) First component of some representative  $K_n$ , namely  $n = 1, 2, 4, 8, 16, 32, 64$ . (Inset)  $K_n^1$  functions were zoomed in on the range  $[0.9, 1]$ . (B) 16 isochrons uniformly distributed in time along the limit cycle (red) computed using the functions  $K_n$  up to order  $L = 64$  shown in (A) for some representative values of  $n$ . (C) Globalization of the local isochrons shown in (B) (green curves in (C)). (D) Zoom of the global isochrons around the unstable limit cycle (solid black curve). Global isochrons spiral around it. In panels (B)–(D) isochrons are colored according to their phases. The  $V$ - and  $w$ -nullclines are shown in gray, the fixed point at  $(-26.3, 0.13)$  corresponding to a stable focus is represented by a filled circle, and the unstable limit cycle is shown in black.

We already mentioned that the solution  $K$  is not unique; indeed, if  $K(\theta, \sigma)$  is a solution of the invariance equation, so is  $K(\theta + \theta_0, b\sigma)$  for any  $\theta_0 \in [0, 1]$  and  $b \in \mathbb{R}$ . We choose  $\theta_0$  so that the zero phase for the oscillator corresponds to the maximum value of the  $x$ - or  $V$ -coordinate (in neuroscience, the peak of the spike). Notice than when  $K_1$  is multiplied by a constant  $b$ , the subsequent  $K_n$  functions will be multiplied by  $b^n$ . Theoretically, this scaling will scale only the domain of  $K$  and therefore the radius of convergence of the Taylor series but will not affect the computed isochron. Computationally, it is well known that the roundoff becomes very problematic if we are working with numbers that are some orders of magnitude apart. By choosing  $b$  appropriately, it is possible to aim for having the successive  $K_n$  more or less constant in size so that roundoff is greatly reduced. We discuss the choice of  $b$  in detail in the next section for the Rayleigh example.



**Figure 4.** Morris–Lecar near an SNIC bifurcation. Set of parameters as in Table 1 under the SNIC column. (A) First component of some representative  $K_n$ , namely  $n = 1, 2, 4, 8, 16, 32, 64$ . (B) 32 isochrons uniformly distributed in time along the limit cycle (red) computed using the functions  $K_n$  up to order  $L = 64$  shown in (A) for some representative values of  $n$ . (C) Globalization of the local isochrons shown in (B) (green curves in (C)). (D) Zoom of the global isochrons around the unstable fixed point (open circle). The  $V$ - and  $w$ -nullclines are shown in gray, and the fixed point  $(5.09, 0.31)$  which corresponds to an unstable focus is represented by an open circle.

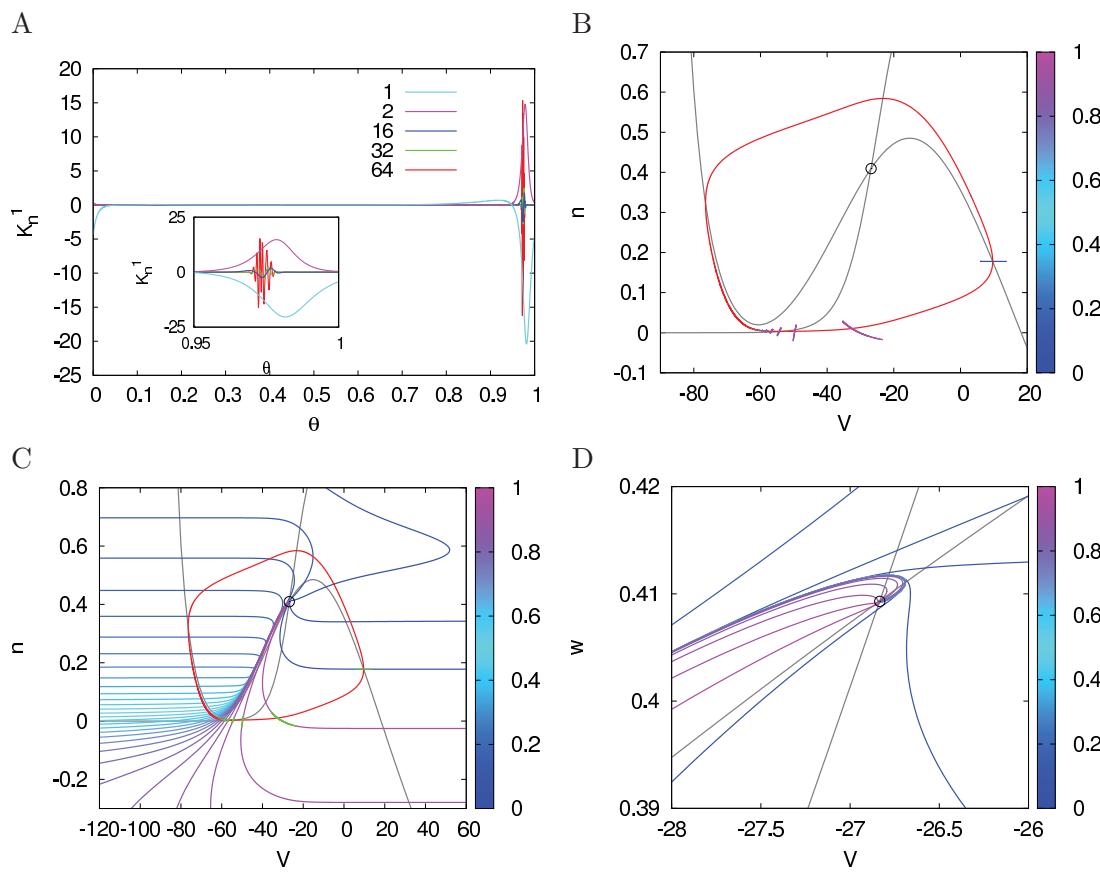
**8.1.1. The Rayleigh oscillator and the role of the parameter  $b$ .** We explore the role of  $b$  with the Rayleigh example using  $N = 2048$  Fourier coefficients. In Figure 1(A)–(C) we show the first component of the  $K_n$  functions computed up to  $n = L = 64$  for three different values of  $b$ . For small values of  $b$ ,  $K_n$  go to zero very fast as  $n$  increases, with the result that increasing the order  $L$  does not provide any extra information since we are just adding terms that are 0 (see Figure 1(A), where  $b = 0.25$ , and the computed local isochron for this value of  $b$  in Figure 1(D)). For large values of  $b$ ,  $K_n$  functions blow up as  $n$  increases and, eventually, those numbers will be too large to computationally operate with them causing large errors (see Figure 1(C), where  $b = 1.5$ ). If we increase  $n$  even more, the results are not reliable. For moderate values of  $b$ ,  $K_n$  functions can be kept at a reasonable range and the local approximation extends to a larger neighborhood (see Figure 1(B) and (D)). It is worth remarking that the range of suitable values of  $b$  shrinks as  $n$  increases. Meanwhile, for  $n$



**Figure 5.** Morris–Lecar near a homoclinic bifurcation. Set of parameters as in Table 1 under the Homoclinic column. (A) First component of some representative  $K_n$ , namely  $n = 1, 2, 4, 8, 16, 32, 64$ . (B) 32 isochrons uniformly distributed in time along the limit cycle (red) computed using the functions  $K_n$  up to order  $L = 64$  shown in (A) for some representative values of  $n$ . (C) Globalization of the local isochrons shown in (B) (green curves in (C)). Global isochrons spiral around the unstable limit cycle (solid black) and are confined by the stable manifold of the saddle point (black curve with arrows pointing towards the saddle point represented by an open triangle). (D) The isochron corresponding to phase 0. Local approximation is shown in green and globalization in blue. In panels (B)–(D) isochrons are colored according to their phases. The  $V$ - and  $w$ -nullclines are shown in gray, the fixed points at  $(-35.7, 0)$ ,  $(-23.8, 0.02)$ , and  $(4.5, 0.3)$  correspond to a stable node (filled square), a saddle (open triangle), and a stable focus (filled circle), respectively, and the stable and unstable manifolds of the saddle point as well as the unstable limit cycle are shown in black.

small, many values of  $b$  can be appropriate to keep the coefficients  $K_n$  to a moderate size; as  $n$  increases some of the values of  $b$  which were appropriate for low  $n$  may cause  $K_n$  functions to blow up or vanish. Take, for instance, the case shown in Figure 1(C), where  $b = 1.5$ . Up to  $n = 16$ ,  $K_n$  are of the order  $10^2$  at most. Hence, this value of  $b$  would be good if we had to compute  $K_n$  up to order 16. In the case shown in Figure 1(A),  $K_n$  are smaller than  $10^{-4}$  for  $n \geq 6$ . So,  $b$  would be good enough if we just want to compute  $n$  up to order 6 or smaller. In all the examples, we have done this adjustment manually. We think that an automatic implementation could be designed.

Another issue about  $b$  is that since we impose a uniform contraction rate  $\lambda$  for the nor-



**Figure 6.** The  $(I_{Na,p} + I_K)$ -model near an SNIC bifurcation. (A) First component of some representative  $K_n$ , namely  $n = 1, 2, 16, 32, 64$ . (Inset)  $K_n$  functions were zoomed in on the range  $[0.95, 1]$ . (B) 32 isochrons uniformly distributed in time along the limit cycle (red) computed using the functions  $K_n$  up to order  $L = 64$  shown in (A) for some representative values of  $n$ . (C) Globalization of the local isochrons shown in (B) (green curves in (C)). (D) Zoom of the global isochrons around the unstable focus (open circle). In panels (B)–(D) isochrons are colored according to their phases. The  $V$ - and  $n$ -nullclines are shown in gray, and the fixed point at  $(-26.8, 0.4)$  corresponding to an unstable focus is represented by an open circle.

mal variable, the change of coordinates  $K$  may take values that are some magnitudes apart, especially when the stable limit cycle has different contraction rates along it. In these cases, it is not possible to choose a uniform  $b$  such that  $K_n$  have order 1 for all values of  $\theta$  and the successive  $K_n$  go to zero for some range of values of  $\theta$  (see Figure 1(B) for  $\theta$  in the ranges  $[0, 0.3]$  and  $[0.5, 0.8]$ ). Thus, for values of  $\theta$  in the mentioned range, increasing the degree  $L$  of the Taylor expansion does not have any effect on the growth of the local domain where the isochron can be computed semianalytically (see Figure 2(A)). Indeed, for values of  $\theta$  in the mentioned range, which correspond to the isochrons on the right and left portions of the limit cycle, the local isochron does not grow further when we increase  $L$  above 32 (in some cases even 16). However, for those values of  $\theta$  with nonzero  $K_n$  for large  $n$  ( $\theta$  values around 0.4 and 0.9; see Figure 1(B)), the local isochrons increase when we increase the order of the local approximation. See Figure 2(A). This situation becomes more dramatic for systems that have

an accentuated slow-fast dynamics such as the  $(I_{Na,p} + I_K)$ -model.

**8.1.2. The Morris–Lecar model.** For the ML-Hopf model, the oscillations bifurcate from the fixed point through a subcritical Hopf bifurcation. Thus, for a small range of values of  $I_{app}$  the system shows bistability between a stable steady state (focus) and a stable limit cycle. The regions of attraction of these two objects are separated by an unstable limit cycle. See Figure 3(B). Hence, the isochrons of the stable limit cycle are confined by the unstable limit cycle and they spiral around it. See Figure 3(C) and (D).

For the ML-SNIC model, the oscillations bifurcate from the fixed point through an SNIC bifurcation. That is, a saddle and a node coalesce and disappear, giving rise to a stable limit cycle that bifurcates from the homoclinic of the saddle point. Thus, for values of  $I_{app}$  close to the critical value, the stable limit cycle inherits the slow-fast dynamics of the homoclinic connection. Indeed, the system moves slowly near the “ghost” of the saddle-node fixed point. One can observe the slow-fast dynamics by looking at the spatial distance between isochrons that are equally spaced in time. See Figure 4(B) and (C). In the regions where isochrons stay close to each other, the dynamics are slow, whereas where they are separated, the dynamics are fast. In this case, the extended isochrons spiral around the unstable fixed point (focus) inside the limit cycle. See Figure 4(C) and (D). Notice that the local approximation extends to a large neighborhood. See Figure 4(B) and (C).

For the ML-Hom model, the oscillations bifurcate from a homoclinic, but in this case the node and the saddle persist. See Figure 5(B). Moreover, for a certain range of parameters the system presents tristability: two stable fixed points (a node and a focus) and one stable limit cycle. See Figure 5(B). In this case, the region where the isochrons for the stable limit cycle are defined is confined between the stable manifold of the nearby saddle and the unstable limit cycle lying in the interior. See Figure 5(C) and (D). Notice that the isochrons also spiral around the unstable limit cycle.

**8.1.3. The  $(I_{Na,p} + I_K)$ -model and the slow-fast dynamics.** We have chosen this example because it presents a clear time-scale separation between the two variables, so it challenges our numerical algorithms. The oscillations emerge through an SNIC bifurcation as in the ML-SNIC example. We have chosen a value of  $I_{app}$  that is slightly above the critical value. Notice that the limit cycle clearly tracks the left and right branches of the  $V$ -nullcline, and the jumps between those two branches occur at a fast time scale. See Figure 6(B). This time-scale separation causes a strong difference in the contraction rates towards the limit cycle: while the limit cycle is strongly contractive along the left and right branches, it shows a small contraction along the upper and lower portions. An immediate consequence of this fact is that one cannot find a value of  $b$  that keeps  $K_n$  in a good computational range for all phases. See Figure 6(A). Hence, for most values of  $\theta$  (in the range 0–0.9), the local approximation of the isochrons is barely observable (the local neighborhood size is tiny), whereas for the other values one can obtain a pretty large local isochron (values between 0.9–1). See Figure 6(B), and compare with Figure 6(A). Notice that even with a small approximation we can apply the globalization process, so it does not affect the isochron globalization. See Figure 6(C) and (D).

Moreover, since we need to capture the behavior of the  $K_n$  function in a small piece of the domain (range 0.9–1) so that the invariance equation can be solved up to high order, we need to increase the size of the discretization grid (equivalently, the number of Fourier coefficients).

Even if for certain regions of the domain a few points would be enough, the fact that the FFT requires equidistant points forces us to pay the price of increasing the size  $N$  of the grid. So, we used  $N = 2^{13} = 8192$  and  $L = 2^8 = 64$  to compute  $K_n$  with an error of order  $10^{-10}$ . Fortunately, the algorithm described here allows us to perform this computations in seconds on a regular laptop.

**9. Discussion.** We presented efficient algorithms to compute the limit cycle and the local isochrons up to high order. The algorithms are supplemented with rigorous theoretical results of convergence and with numerical implementation details. We have tested them in different relevant examples for which we have computed local isochrons up to order 64.

We call attention to the fact that the main theorem (Theorem 3.2) is formulated in the *a posteriori* format. That is, Theorem 3.2 shows that given an approximate solution of the invariance equation (2.4), which satisfies some explicit nondegeneracy conditions, then there is a true solution nearby. Theorems in an *a posteriori* format can be used to validate the computations and allow us to be confident about the calculations even when they are close to breakdown. Hence, the calculations obtained using Algorithm 4.4 can easily be turned into computer assisted proofs. It suffices to estimate rigorously the error and the condition numbers.

From the analytic point of view, we stress that the method also leads to several other consequences, such as uniqueness and smooth dependence on parameters. Some of them can be obtained by the standard methods of ODEs (notably the smooth dependence on parameters for the limit cycle). We obtain analytic regularity of the foliation by isochrons (note that the general theory of normally hyperbolic invariant manifolds produces only finitely differentiable results and that this is sharp in examples which are excluded by our restricted class of models).

The numerical method works well irrespective of other invariant objects in the system that confine the basin of attraction to the limit cycle. Winfree referred to these sets as phaseless sets [47]. We have shown examples with an unstable limit cycle of a different period inside, so that the isochrons spiral around it (Figures 3 and 5), and with a saddle point in the vicinity of the limit cycle whose invariant manifolds then bound the domain where the isochrons are extended (Figure 5).

An important point for our method is to find a value of  $b$  such that the coefficients  $K_n$  can be kept at order 1, so that one can avoid the roundoff errors. However, quite often, orbits do not approach the limit cycle uniformly along it; indeed, the contraction is stronger around certain phases than others. The immediate consequence of this nonuniform contraction is that one cannot find a value of  $b$  that keeps  $K_n$  in a suitable range for all values of  $\theta$ . Instead, for a given  $b$  only for a fraction of the domain, the function  $K_n$  takes values in a suitable computational range, while for the rest of the domain,  $K_n$  goes to zero. This is even more dramatic for stiff systems, as in the  $(I_{Na,p} + I_K)$ -model.

Moreover, in the case of stiff systems one has to approximate functions that, although they are analytic, have Fourier coefficients that decrease very slowly and not uniformly. Therefore, one needs to use a large number of Fourier coefficients to approximate them correctly. We think that for these cases, other methods of discretization which are more adaptive, like splines, could give some improvements.

Finally, we want to emphasize the advantages of computing a local approximation of the

isochrons up to high order, beyond using them to compute the global isochrons. Actually, the globalization of the isochrons could be performed just with the linear approximation ( $L = 1$ ) (see, for instance, [33]). However, the fact that we have a parameterization  $K$  computed “semianalytically” in a large neighborhood around the limit cycle provides us with relevant information about the oscillator.

On one hand, the value of the PRC not only on the limit cycle ( $s = 0$ ) but also in a neighborhood of it ( $s > 0$ ) immediately follows from  $K$ . So, the effects of the perturbation on the normal variable  $s$  are also known. This allows us to study perturbations that displace the trajectory away from the limit cycle as well as perturbations that occur when the system is not on the limit cycle [35, 7, 46].

On the other hand, one can obtain high order PRCs from the parameterization  $K$  or, even better, go beyond PRCs and use the fact that when the parameterization  $K$  is known we can obtain the new phase just inverting the map  $K$ . Thus, if a brief (but not necessarily weak) perturbation displaces the point  $(x, y)$  to  $(x', y')$ , the new phase  $\theta'$  and normal variable  $s'$  are given just by  $K^{-1}(x', y')$ . Fast and accurate methods for computing isochrons are needed to study coupled oscillators beyond the restriction of weak perturbations.

## REFERENCES

- [1] L. V. AHLFORS, *Complex Analysis: An Introduction to the Theory of Analytic Functions of One Complex Variable*, 3rd ed., International Series in Pure and Applied Mathematics, McGraw-Hill, New York, 1978.
- [2] R. BROUKE AND K. GARTHWAITE, *A programming system for analytical series expansions on a computer*, Celestial Mech., 1 (1969), pp. 271–284.
- [3] X. CABRÉ, E. FONTICH, AND R. DE LA LLAVE, *The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces*, Indiana Univ. Math. J., 52 (2003), pp. 283–328.
- [4] X. CABRÉ, E. FONTICH, AND R. DE LA LLAVE, *The parameterization method for invariant manifolds. II. Regularity with respect to parameters*, Indiana Univ. Math. J., 52 (2003), pp. 329–360.
- [5] X. CABRÉ, E. FONTICH, AND R. DE LA LLAVE, *The parameterization method for invariant manifolds. III. Overview and applications*, J. Differential Equations, 218 (2005), pp. 444–515.
- [6] R. CALLEJA AND R. DE LA LLAVE, *A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification*, Nonlinearity, 23 (2010), pp. 2029–2058.
- [7] O. CASTEJÓN, A. GUILLAMON, AND G. HUGUET, *Phase-amplitude response functions for transient-state stimuli*, J. Math. Neurosci., (2013), 3:13 doi:10.1186/2190-8567-3-13.
- [8] C. DE BOOR, *A Practical Guide to Splines*, Appl. Math. Sci. 27, Springer-Verlag, New York, 2001.
- [9] A. DEPRIT, *Canonical transformations depending on a small parameter*, Celestial Mech., 1 (1969/1970), pp. 12–30.
- [10] J. W. EATON, *GNU Octave 3.6.2*, <http://www.octave.org> (2012).
- [11] G. B. ERMENTROUT, *Type I membranes, phase resetting curves, and synchrony*, Neural Comput., 8 (1996), pp. 979–1001.
- [12] G. B. ERMENTROUT AND N. KOPELL, *Multiple pulse interactions and averaging in systems of coupled neural oscillators*, J. Math. Biol., 29 (1991), pp. 195–217.
- [13] G. B. ERMENTROUT AND D. H. TERMAN, *Mathematical Foundations of Neuroscience*, Interdiscip. Appl. Math. 35, Springer, New York, 2010.
- [14] M. FRIGO AND S. G. JOHNSON, *The design and implementation of FFTW3*, Proc. IEEE, 93 (2005), pp. 216–231.
- [15] M. GALASSI, J. DAVIS, J. THAILER, B. GOUGH, G. JUNGMAN, P. ALKEN, M. BOOTH, AND F. ROSSI, *GSL, GNU Scientific Library 1.15*, <http://www.gnu.org/software/gsl> (2011).
- [16] M. GASTINEAU AND J. LASKAR, *TRIP 1.2.26*, TRIP Reference manual, IMCCE, Paris Observatory, <http://www.imcce.fr/trip/> (2012).

- [17] R. M. GHIGLIAZZA AND P. HOLMES, *A minimal model of a central pattern generator and motoneurons for insect locomotion*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 671–700.
- [18] J. GUCKENHEIMER, *Isochrones and phaseless sets*, J. Math. Biol., 1 (1974/75), pp. 259–273.
- [19] A. GUILLAMON AND G. HUGUET, *A computational and geometric approach to phase resetting curves and surfaces*, SIAM J. Appl. Dyn. Syst., 8 (2009), pp. 1005–1042.
- [20] A. HARO, *Automatic Differentiation Tools in Computational Dynamical Systems*, manuscript, 2008.
- [21] E. M. IZHKEVICH, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*, Comput. Neurosci., MIT Press, Cambridge, MA, 2007.
- [22] À. JORBA AND M. ZOU, *A software package for the numerical integration of ODEs by means of high-order Taylor methods*, Experiment. Math., 14 (2005), pp. 99–117.
- [23] D. E. KNUTH, *The Art of Computer Programming. Volume 2: Seminumerical Algorithms*, Addison-Wesley, Reading, MA, 1997.
- [24] A. N. KOLMOGOROV, *On conservation of conditionally periodic motions for a small change in Hamilton's function*, Dokl. Akad. Nauk SSSR (N.S.), 98 (1954), pp. 527–530 (in Russian); Stochastic Behavior in Classical and Quantum Hamiltonian Systems (Volta Memorial Conf., Como, 1977), Lecture Notes in Phys. 93, Springer, Berlin, 1979, pp. 51–56 (in English).
- [25] R. DE LA LLAVE, *A tutorial on KAM theory*, in Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999), AMS, Providence, RI, 2001, pp. 175–292.
- [26] R. DE LA LLAVE, A. GONZÁLEZ, À. JORBA, AND J. VILLANUEVA, *KAM theory without action-angle variables*, Nonlinearity, 18 (2005), pp. 855–895.
- [27] S. K. MARAN AND C. C. CANAVIER, *Using phase resetting to predict 1:1 and 2:2 locking in two neuron networks in which firing order is not always preserved*, J. Comput. Neurosci., 24 (2008), pp. 37–55.
- [28] A. MAUROY AND I. I. MEZIĆ, *On the use of Fourier averages to compute the global isochrons of (quasi)periodic dynamics*, Chaos, 22 (2012), 033112.
- [29] C. MORRIS AND H. LECAR, *Voltage oscillations in the barnacle giant muscle fiber*, Biophys. J., 35 (1981), pp. 193–213.
- [30] J. MOSER, *A rapidly convergent iteration method and non-linear differential equations. II*, Ann. Scuola Norm. Sup. Pisa (3), 20 (1966), pp. 499–535.
- [31] J. MOSER, *A rapidly convergent iteration method and non-linear partial differential equations. I*, Ann. Scuola Norm. Sup. Pisa (3), 20 (1966), pp. 265–315.
- [32] J. MOSER, *Convergent series expansions for quasi-periodic motions*, Math. Ann., 169 (1967), pp. 136–176.
- [33] H. M. OSINGA AND J. MOEHLIS, *Continuation-based computation of global isochrons*, SIAM J. Appl. Dyn. Syst., 9 (2010), pp. 1201–1228.
- [34] R. B. PLATTE AND L. N. TREFETHEN, *Chebfun: A new kind of numerical computing*, in Progress in Industrial Mathematics at ECMI 2008, Math. Ind. 15, Springer, Heidelberg, 2010, pp. 69–87.
- [35] A. RABINOVITCH AND M. FRIEDMAN, *Fixed points of two-dimensional maps obtained under rapid stimulations*, Phys. Lett. A, 355 (2006), pp. 319–325.
- [36] L. RAYLEIGH, *On maintained vibrations*, Phil. Mag., 15 (1883), pp. 229–235.
- [37] J. RINZEL AND G. B. ERMENTROUT, *Analysis of neural excitability and oscillations*, in Methods in Neuronal Modelling: From Synapses to Networks, 2nd ed., C. Koch and I. Segev, eds., MIT Press, Cambridge, MA, 1998, pp. 251–291.
- [38] W. RUDIN, *Real and Complex Analysis*, 2nd ed., McGraw-Hill, New York, 1974.
- [39] S. SAKS AND A. ZYGMUND, *Analytic Functions*, 2nd ed., Monografie Matematyczne, Tom 28, Państwowe Wydawnictwo Naukowe, Warsaw, 1965.
- [40] D. S. SCHMIDT, *Polypak: An algebraic processor for computations in celestial mechanics*, in Computer Algebra, D. Chudnovsky and R. Jenks, eds., Marcel Dekker, New York, 1989, pp. 111–120.
- [41] W. E. SHERWOOD AND J. GUCKENHEIMER, *Dissecting the phase response of a model bursting neuron*, SIAM J. Appl. Dyn. Syst., 9 (2010), pp. 659–703.
- [42] C. SIMÓ, *On the analytical and numerical approximation of invariant manifolds*, in Les Méthodes Modernes de la Mécanique Céleste, Editions Frontières, D. Benest and C. Froeschlé, eds., 1990, pp. 285–329.
- [43] D. TAKESHITA AND R. FERES, *Higher order approximation of isochrons*, Nonlinearity, 23 (2010), pp. 1303–1323.
- [44] B. VAN DER POL, *On relaxation oscillations*, The London, Edinburgh and Dublin Phil. Mag. and J. of Sci., 2 (1927), pp. 978–992.

- [45] C. VAN VREESWIJK, L. F. ABBOTT, AND G. B. ERMENTROUT, *When inhibition not excitation synchronizes neural firing*, J. Comput. Neurosci., 1 (1994), pp. 313–321.
- [46] K. C. WEDGWOOD, K. K. LIN, R. THUL, AND S. COOMBES, *Phase-amplitude descriptions of neural oscillator models*, J. Math. Neurosci., 3 (2013), 2.
- [47] A. T. WINFREE, *Patterns of phase compromise in biological cycles*, J. Math. Biol., 1 (1974/75), pp. 73–95.
- [48] E. ZEHNDER, *Generalized implicit function theorems with applications to some small divisor problems. I*, Comm. Pure Appl. Math., 28 (1975), pp. 91–140.