Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Physica D 237 (2008) 2599–2615
PHYSICA D
www.elsevier.com/locate/physd

Computation of derivatives of the rotation number for parametric families of circle diffeomorphisms

Alejandro Luque, Jordi Villanueva*

Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain

Received 25 July 2007; received in revised form 22 February 2008; accepted 28 March 2008
Available online 4 April 2008

Communicated by J. Stark

Abstract

In this paper we present a numerical method to compute derivatives of the rotation number for parametric families of circle diffeomorphisms with high accuracy. Our methodology is an extension of a recently developed approach to compute rotation numbers based on suitable averages of iterates of the map and Richardson extrapolation. We focus on analytic circle diffeomorphisms, but the method also works if the maps are differentiable enough. In order to justify the method, we also require the family of maps to be differentiable with respect to the parameters and the rotation number to be Diophantine. In particular, the method turns out to be very efficient for computing Taylor expansions of Arnold Tongues of families of circle maps. Finally, we adapt these ideas to study invariant curves for parametric families of planar twist maps.

PACS: 02.30.Mv; 02.60.-x; 02.70.-c

Keywords: Families of circle maps; Derivatives of the rotation number; Numerical approximation

1. Introduction

The rotation number, introduced by Poincaré, is an important topological invariant in the study of the dynamics of circle maps and, by extension, invariant curves for maps or two dimensional invariant tori for vector fields. For this reason, several numerical methods for approximating rotation numbers have been developed during the last years. We refer to the works [3,4,8,13,14,21,24,31] as examples of methods of different nature and complexity. This last ranges from pure definition of the rotation number to sophisticated and involved methods like frequency analysis. The efficiency of these methods varies depending if the approximated rotation number is rational or irrational. Moreover, even though some of them can be very accurate in many cases, they are not adequate for every kind of application, for example due to violation of their assumptions or due to practical reasons, like the required amount of memory.

Recently, a new method for computing Diophantine rotation numbers of circle diffeomorphisms with high precision at low computational cost has been introduced in [27]. This method is built assuming that the circle map is conjugate to a rigid rotation in a sufficiently smooth way and, basically, it consists in averaging the iterates of the map together with Richardson extrapolation. This construction takes advantage of the geometry and dynamics of the problem, so it turns out to be very efficient in multiple applications. The method is specially suited if we are able to compute iterates of the map with a high precision, for example if we can work with computer arithmetic having a large number of decimal digits.

The goal of this paper is to extend the method of [27] in order to compute derivatives of the rotation number with respect to parameters in families of circle diffeomorphisms. We follow the same averaging-extrapolation process applied to derivatives of iterates of the map. To this end, we require the family to be differentiable with respect to parameters. Hence, we are able to obtain accurate variational information at the same time that we approximate the rotation number. Consequently, the method allows us to study parametric families of circle maps

---

* Corresponding author. Tel.: +34 934015887; fax: +34 934011713.
E-mail addresses: alejandro.luque@upc.edu (A. Luque), jordi.villanueva@upc.es (J. Villanueva).

0167-2789/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2008.03.047

from a point of view that is not given by any of the previously mentioned methods.

From a practical point of view, circle diffeomorphisms appear in the study of quasi-periodic invariant curves for maps. In particular, for planar twist maps, any such a curve induces a circle diffeomorphism in a direct way just by projecting the iterates on the angular variable. Then, using approximate derivatives of the rotation number, we can continue these invariant curves numerically with respect to parameters by means of the Newton method. The methodology presented is an alternative to more common approaches based on solving the invariance equation numerically, interpolation of the map or approximation by periodic orbits (see for example [5,7,12,28]). Furthermore, using the variational information obtained, we are able to compute asymptotic expansions relating parameters and initial conditions that correspond to curves of fixed rotation number.

Finally, we point out that the method can be formally extended to deal with maps on the torus with a Diophantine rotation vector. However, in order to apply the method to the study of quasi-periodic tori for symplectic maps in higher dimensions, there is not an analogue of the twist condition to guarantee a well defined projection of iterates on the standard torus. Then, our interest is immediately focused on a generalization of the method in the case of non-twist maps with folded invariant curves (for example, so-called meanderings [29]). These and other extensions will be object of future research [22].

Contents of the paper are organized as follows. In Section 2 we recall some fundamental facts about circle maps and we briefly review the method of [27]. In Section 3 we describe the method for the computation of derivatives of the rotation number. The rest of the paper is devoted to illustrate the method through several applications. Concretely, in Section 4 we study the Arnold family of circle maps. Finally, in Section 5 we focus on the computation and continuation of invariant curves for planar twist maps and, in particular, we present some computations for the conservative Hénon map.

2. Notation and previous results

All the results presented in this section can be found in the bibliography, but we include them for self-consistency of the text. Concretely, in Section 2.1 we recall the basic definitions, notations and properties of circle maps that we need in the paper (we refer to [9,18] for more details and proofs). On the other hand, in Section 2.2 we review briefly the method of [27] for computing rotation numbers of circle diffeomorphisms.

2.1. *Circle diffeomorphisms*

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the real circle which inherits both a group structure and a topology by means of the natural projection $\pi : \mathbb{R} \to \mathbb{T}$ (also called the universal cover of $\mathbb{T}$). We denote by $\text{Diff}^r_+(\mathbb{T})$, $r \in [0, +\infty) \cup \{\infty, \omega\}$, the group of orientation-preserving homeomorphisms of $\mathbb{T}$ of class $\mathcal{C}^r$. Concretely, if $r = 0$ it is the group of homeomorphisms of $\mathbb{T}$; if $r \geq 1$, $r \in (0, \infty) \setminus \mathbb{N}$, it is the group of $\mathcal{C}^{\lfloor r \rfloor}$-diffeomorphisms whose $\lfloor r \rfloor$-th derivative verifies a Hölder condition with exponent $r - \lfloor r \rfloor$; if $r = \omega$ it is the group of real analytic diffeomorphisms.

Given $f \in \text{Diff}^r_+(\mathbb{T})$, we can lift f to $\mathbb{R}$ by π obtaining a $\mathcal{C}^r$ map $\tilde{f}$ that makes the following diagram commute

$$\begin{tikzcd}
\mathbb{R} \arrow[r, "\tilde{f}"] \arrow[d, "\pi"'] & \mathbb{R} \arrow[d, "\pi"] \\
\mathbb{T} \arrow[r, "f"] & \mathbb{T}
\end{tikzcd}
\qquad \pi \circ \tilde{f} = f \circ \pi.$$

Moreover, we have $\tilde{f}(x + 1) - \tilde{f}(x) = 1$ (since f is orientation-preserving) and the lift is unique if we ask for $\tilde{f}(0) \in [0, 1)$. Accordingly, from now on we choose the lift with this normalization so we can omit the tilde without any ambiguity.

Definition 2.1. Let f be the lift of an orientation-preserving homeomorphism of the circle such that $f(0) \in [0, 1)$. Then the *rotation number of* f is defined as the limit

$$\rho(f) := \lim_{|n| \to \infty} \frac{f^n(x_0) - x_0}{n},$$

that exists for all $x_0 \in \mathbb{R}$, is independent of x_0 and satisfies $\rho(f) \in [0, 1)$.

Let us remark that the rotation number is invariant under orientation-preserving conjugation, i.e., for every $f, h \in \text{Diff}^0_+(\mathbb{T})$ we have that $\rho(h^{-1} \circ f \circ h) = \rho(f)$. Furthermore, given $f \in \text{Diff}^2_+(\mathbb{T})$ with $\rho(f) \in \mathbb{R} \setminus \mathbb{Q}$, Denjoy's theorem ensures that f is topologically conjugate to the rigid rotation $R_{\rho(f)}$, where $R_\theta(x) = x + \theta$. That is, there exists $\eta \in \text{Diff}^0_+(\mathbb{T})$ making the following diagram commute

$$\begin{tikzcd}
\mathbb{T} \arrow[r, "f"] & \mathbb{T} \\
\mathbb{T} \arrow[r, "R_{\rho(f)}"] \arrow[u, "\eta"] & \mathbb{T} \arrow[u, "\eta"']
\end{tikzcd}
\qquad f \circ \eta = \eta \circ R_{\rho(f)}. \tag{1}$$

In addition, if we require $\eta(0) = x_0$, for fixed x_0, then the conjugacy η is unique.

All the ideas and algorithms described in this paper make use of the existence of such conjugation and its regularity. Let us remark that, although smooth or even finite differentiability is enough, in this paper we are concerned with the analytic case. Moreover, it is well known that regularity of the conjugation depends also on the rational approximation properties of $\rho(f)$, so we will focus on Diophantine numbers.

Definition 2.2. Given $\theta \in \mathbb{R}$, we say that θ is a *Diophantine number* of (C, τ) type if there exist constants $C > 0$ and $\tau \geq 1$ such that

$$\left|1 - \mathrm{e}^{2\pi i k\theta}\right|^{-1} \leq C|k|^\tau, \quad \forall k \in \mathbb{Z}_*.$$

We will denote by $\mathcal{D}(C, \tau)$ the set of such numbers and by $\mathcal{D}$ the set of Diophantine numbers of any type.

Although Diophantine sets are Cantorian (i.e., compact, perfect and nowhere dense) a remarkable property is that $\mathbb{R}\setminus\mathcal{D}$ has zero Lebesgue measure. For this reason, this condition fits very well in practical issues and we do not resort to other weak conditions on small divisors such as the Brjuno condition (see [33]).

The first result on the regularity of conjugacy (1) is due to Arnold [2] but we also refer to [16,19,30,33] for later contributions. In particular, theoretical support of the methodology is provided by the following result:

Theorem 2.3 (*Katznelson and Ornstein [19]*). *If $f \in \mathrm{Diff}^r_+(\mathbb{T})$ has Diophantine rotation number $\rho(f) \in \mathcal{D}(C,\tau)$ for $\tau + 1 < r$, then f is conjugated to $R_{\rho(f)}$ by means of a conjugacy $\eta \in \mathrm{Diff}^{r-\tau-\varepsilon}_+(\mathbb{T})$, for any $\varepsilon > 0$. Note that $\mathrm{Diff}^{\omega}_+(\mathbb{T}) = \mathrm{Diff}^{\omega-\tau-\varepsilon}_+(\mathbb{T})$ while the domain of analyticity is reduced.*

2.2. Computing rotation numbers by averaging and extrapolation

We review here the method developed in [27] for computing Diophantine rotation numbers of analytic circle diffeomorphisms (the $\mathcal{C}^r$ case is similar). This method is highly accurate with low computational cost and it turns out to be very efficient when combined with multiple precision arithmetic routines. The reader is referred there for a detailed discussion and several applications.

Let us consider $f \in \mathrm{Diff}^{\omega}_+(\mathbb{T})$ with rotation number $\theta = \rho(f) \in \mathcal{D}$. Notice that we can write the conjugacy of Theorem 2.3 as $\eta(x) = x + \xi(x)$, ξ being a 1-periodic function normalized in such a way that $\xi(0) = x_0$, for a fixed $x_0 \in [0,1)$. Now, by using the fact that η conjugates f to a rigid rotation, we can write the following expression for the iterates under the lift:

$$f^n(x_0) = f^n(\eta(0)) = \eta(n\theta) = n\theta + \sum_{k\in\mathbb{Z}} \hat{\xi}_k \mathrm{e}^{2\pi\mathrm{i}kn\theta}, \tag{2}$$

$\forall n \in \mathbb{Z}$, where the sequence $\{\hat{\xi}_k\}_{k\in\mathbb{Z}}$ denotes the Fourier coefficients of ξ. Then, the above expression gives us the following formula

$$\frac{f^n(x_0) - x_0}{n} = \theta + \frac{1}{n}\sum_{k\in\mathbb{Z}_*} \hat{\xi}_k (\mathrm{e}^{2\pi\mathrm{i}kn\theta} - 1),$$

to compute θ modulo terms of order $\mathcal{O}(1/n)$. Unfortunately, this order of convergence is very slow for practical purposes, since it requires a huge number of iterates if we want to compute θ with high precision. Nevertheless, by averaging the iterates $f^n(x_0)$ in a suitable way, we can manage to decrease the order of this quasi-periodic term.

As a motivation, let us start by considering the sum of the first N iterates under f, that has the following expression (we use (2) to write the iterates)

$$S^1_N(f) := \sum_{n=1}^{N}(f^n(x_0) - x_0) = \frac{N(N+1)}{2}\theta - N\sum_{k\in\mathbb{Z}_*}\hat{\xi}_k + \sum_{k\in\mathbb{Z}_*}\hat{\xi}_k \frac{\mathrm{e}^{2\pi\mathrm{i}k\theta}(1-\mathrm{e}^{2\pi\mathrm{i}kN\theta})}{1-\mathrm{e}^{2\pi\mathrm{i}k\theta}},$$

and we observe that the new factor multiplying θ grows quadratically with the number of iterates, while it appears a linear term in N with constant $A_1 = -\sum_{k\in\mathbb{Z}_*}\hat{\xi}_k$. Moreover, the quasi-periodic sum remains uniformly bounded since θ is Diophantine and η is analytic (use Lemma 2.4 with $p = 1$). Thus, we obtain

$$\frac{2}{N(N+1)} S^1_N(f) = \theta + \frac{2}{N+1}A_1 + \mathcal{O}(1/N^2), \tag{3}$$

that allows us to extrapolate the value of θ with an error $\mathcal{O}(1/N^2)$ if, for example, we compute $S_N(f)$ and $S_{2N}(f)$.

In general, we introduce the following *recursive sums* for $p \in \mathbb{N}$

$$S^0_N(f) := f^N(x_0) - x_0, \qquad S^p_N(f) := \sum_{j=1}^{N} S^{p-1}_j(f). \tag{4}$$

Then, the result presented in [27] says that under the above hypotheses, the following *averaged sums of order p*

$$\tilde{S}^p_N(f) := \binom{N+p}{p+1}^{-1} S^p_N(f) \tag{5}$$

satisfy the expression

$$\tilde{S}^p_N(f) = \theta + \sum_{l=1}^{p} \frac{A^p_l}{(N+p-l+1)\cdots(N+p)} + E^p(N), \tag{6}$$

where the coefficients A^p_l depend on f and p but are independent of N. Furthermore, we have the following expressions for them

$$A^p_l = (-1)^l (p-l+2)\cdots(p+1)\sum_{k\in\mathbb{Z}_*}\hat{\xi}_k \frac{\mathrm{e}^{2\pi\mathrm{i}k(l-1)\theta}}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^{l-1}},$$

$$E^p(N) = (-1)^{p+1}\frac{(p+1)!}{N\cdots(N+p)} \times \sum_{k\in\mathbb{Z}_*}\hat{\xi}_k \frac{\mathrm{e}^{2\pi\mathrm{i}kp\theta}(1-\mathrm{e}^{2\pi\mathrm{i}kN\theta})}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^p}.$$

Finally, the remainder $E^p(N)$ is uniformly bounded by an expression of order $\mathcal{O}(1/N^{p+1})$. This follows immediately from the next standard lemma on small divisors.

Lemma 2.4. *Let $\xi \in \mathrm{Diff}^{\omega}_+(\mathbb{T})$ be a circle map that can be extended analytically to a complex strip $B_\Delta = \{z \in \mathbb{C} : |\mathrm{Im}(z)| < \Delta\}$, with $|\xi(z)| \le M$ up to the boundary of the strip. If we denote $\{\hat{\xi}_k\}_{k\in\mathbb{Z}}$ the Fourier coefficients of ξ and consider $\theta \in \mathcal{D}(C,\tau)$, then for any fixed $p \in \mathbb{N}$ we have*

$$\left|\sum_{k\in\mathbb{Z}_*}\hat{\xi}_k \frac{\mathrm{e}^{2\pi\mathrm{i}kp\theta}(1-\mathrm{e}^{2\pi\mathrm{i}kN\theta})}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^p}\right| \le \frac{\mathrm{e}^{-\pi\Delta}}{1-\mathrm{e}^{-\pi\Delta}} 4MC^p\left(\frac{\tau p}{\pi\Delta\mathrm{e}}\right)^{\tau p}.$$

To conclude this survey, we describe implementation of the method and discuss the expected behavior of the extrapolation error. In order to make Richardson extrapolation we assume, for simplicity, that the total number of iterates is a power of two. Concretely, we select an averaging order $p \in \mathbb{N}$, a maximum number of iterates $N = 2^q$, for some $q \ge p$, and compute the averaged sums $\{\tilde{S}^p_{N_j}(f)\}_{j=0,\ldots,p}$ with $N_j = 2^{q-p+j}$. Then, we can use formula (6) to obtain θ by neglecting the remainders $E^p(N_j)$ and solving the resulting linear set of equations for the unknowns $\theta, A^p_1, \ldots, A^p_p$.

However, let us point out that, due to the denominators $(N_j + p - l + 1)\cdots(N_j + p)$, the matrix of this linear system depends on q, and this is inconvenient if we want to repeat the computations using a different number of iterates. Nevertheless, we note that expression (6) can be written alternatively as

$$\tilde{S}^p_N(f) = \theta + \sum_{l=1}^{p} \frac{\hat{A}^p_l}{N^l} + \hat{E}^p(N), \tag{7}$$

for certain $\{\hat{A}^p_l\}_{l=1,\ldots,p}$, also independent of N, and with a new remainder $\hat{E}^p(N)$ that differs from $E^p(N)$ only by terms of order $\mathcal{O}(1/N^{p+1})$. Then, by neglecting the remainder $\hat{E}^p(N)$ in Eq. (7), we can obtain θ by solving a new $(p+1)$-dimensional system of equations, independent of q, for the unknowns $\theta, \hat{A}^p_1/2^{1(q-p)}, \ldots, \hat{A}^p_p/2^{p(q-p)}$. Therefore, the rotation number can be computed as follows

$$\theta = \Theta_{q,p}(f) + \mathcal{O}(2^{-(p+1)q}), \tag{8}$$

where $\Theta_{q,p}$ is an *extrapolation operator*, that is given by

$$\Theta_{q,p}(f) := \sum_{j=0}^{p} c^p_j \tilde{S}^p_{2^{q-p+j}}(f), \tag{9}$$

and the coefficients $\{c^p_j\}_{j=0,\ldots,p}$ are

$$c^p_l = (-1)^{p-l} \frac{2^{l(l+1)/2}}{\delta(l)\delta(p-l)}, \tag{10}$$

where we define $\delta(n) := (2^n - 1)(2^{n-1} - 1)\cdots(2^1 - 1)$ for $n \geq 1$ and $\delta(0) := 1$.

Remark 2.5. Note that the dimension of this linear system and the asymptotic behavior of the error only depend on the averaging order p. For this reason, in [27] p is called the extrapolation order. However, this is not always the case when computing derivatives of the rotation number. As we discuss in Section 3, the extrapolation order is in general less than the averaging order.

As far as the behavior of the error is concerned, using (8) we have that

$$|\theta - \Theta_{q,p}(f)| \leq c/2^{q(p+1)},$$

for certain constant c, independent of q, that we estimate heuristically as follows. Let us compute $\Theta_{q-1,p}(f)$ and $\Theta_{q,p}(f)$. Since $\Theta_{q,p}(f)$ is a better approximation of θ, it turns out that

$$c \sim 2^{(q-1)(p+1)}|\Theta_{q,p}(f) - \Theta_{q-1,p}(f)|.$$

Then, we obtain the expression

$$|\theta - \Theta_{q,p}(f)| \leq \frac{\nu}{2^{p+1}}|\Theta_{q,p}(f) - \Theta_{q-1,p}(f)|, \tag{11}$$

where ν is a "safety parameter" whose role is to prevent oscillations in the error as a function of q due to the quasi-periodic part. In every numerical computation we take $\nu = 10$. For more details on the behavior of the error we refer to [27].

Now, we comment on two sources of error to take into account in the implementation of the method:

- The sums $S^p_{N_j}(f)$ are evaluated using the lift rather than the map itself. Of course, this makes the sums $S^p_{N_j}(f)$ increase (actually they are of order $\mathcal{O}(N^{p+1})$) and is recommended to store their integer and decimal parts separately in order to keep the desired precision.
- If the required number of iterates increases, we have to be aware of round-off errors in evaluation of the iterates. For this reason, when implementing the above scheme in a computer, we use multiple-precision arithmetics. The computations presented in this paper have been performed using a C++ compiler and multiple arithmetic has been provided by the routines *double–double and quad-double package* of [17], which include a *double–double* data type of approximately 32 decimal digits and a *quadruple-double* data type of approximately 64 digits.

Along this section we have required the rotation number to be Diophantine. Of course, if $\theta \in \mathbb{Q}$ Eq. (6) is not valid since, in general, the dynamic of f are not conjugate to a rigid rotation. Anyway, we can compute the sums $S^p_N(f)$ and it turns out that the method works as well as for Diophantine numbers. We can justify this behavior from the known fact that, for any circle homeomorphism of rational rotation number, every orbit is either periodic or its iterates converge to a periodic orbit (see [9,18]). Then the iterates of the map tend toward periodic points, and for such points, one can see that the averaged sums $\tilde{S}^p_N(f)$ also satisfy an expression like (7) with an error of the same order, and this is all we need to perform the extrapolation. In fact, the worst situation appears when computing irrational rotation numbers that are "close" to rational ones (see also the discussion in Section 4.1).

3. Derivatives of the rotation number with respect to parameters

Now we adapt the method already described in Section 2 in order to compute derivatives of the rotation number with respect to parameters (assuming that they exist). For the sake of simplicity, we introduce the method for one-parameter families of circle diffeomorphisms, albeit the construction can be adapted to deal with multiple parameters (we discuss this situation in Remark 3.3). Thus, consider $\mu \in I \subset \mathbb{R} \mapsto f_\mu \in \mathrm{Diff}^\omega_+(\mathbb{T})$ depending on μ in a regular way. Rotation numbers of the family $\{f_\mu\}_{\mu \in I}$ induce a function $\theta : I \to [0, 1)$ given by $\theta(\mu) = \rho(f_\mu)$. Then our goal is to numerically approximate the derivatives of θ at a given point μ_0.

Let us remark that the function θ is only continuous in $\mathcal{C}^0$-topology and, actually, the rotation number depends on μ in a very non-smooth way: generically, there exists a family of disjoint open intervals of I, with dense union, such that θ

takes constant rational values on these intervals (a so-called Devil's Staircase). However, $\theta'(\mu)$ is defined for any μ such that $\theta(\mu) \notin \mathbb{Q}$ (see [15]).

Higher order derivatives are defined in "many" points in the sense of Whitney. Concretely, let $J \subset I$ be the subset of parameters such that $\theta(\mu) \in \mathcal{D}$ (typically a Cantor set). Then, from Theorem 2.3, there exists a family of conjugacies $\mu \in J \mapsto \eta_\mu \in \mathrm{Diff}^\omega_+(\mathbb{T})$, satisfying $f_\mu \circ \eta_\mu = \eta_\mu \circ R_{\theta(\mu)}$, that is unique if we fix $\eta_\mu(0) = x_0$. Then, if f_μ is $\mathcal{C}^d$ with respect to μ, the Whitney derivatives $D^j_\mu \eta_\mu$ and $D^j_\mu \theta$, for $j = 1, \ldots, d$, can be computed by taking formal derivatives with respect to μ on the conjugacy equation and solving small divisor equations thus obtained. Actually, we know that if we define $J(C,\tau)$ as the subset of J such that $\theta(\mu) \in \mathcal{D}(C,\tau)$, for certain $C > 0$ and $\tau \geq 1$, then the maps $\mu \in J(C,\tau) \mapsto \eta_\mu$ and $\mu \in J(C,\tau) \mapsto \theta$ can be extended to $\mathcal{C}^s$ functions on I, where s depends on d and τ, provided that d is big enough (see [32]).

As it is shown in Section 3.1, when we extend the method for computing the d-th derivative of θ, in general, we are forced to select an averaging order $p > d$ and the remainder turns out to be of order $\mathcal{O}(1/N^{p-d+1})$. Nevertheless, if the rotation number is known to be constant as a function of the parameters, we can avoid previous limitations. Concretely, in this case we can select any averaging order p, independent of d, since the remainder is now of order $\mathcal{O}(1/N^{p+1})$. Of course, if the rotation number is constant, then the derivatives of θ are all zero and the fact that we can obtain them with better precision seems to be irrelevant. However, from the computation of these vanishing derivatives, we can derive information about other involved objects. This is the case of many applications in which this methodology turns out to be very useful (two examples are worked out in Sections 4.3 and 5.3).

3.1. *Computation of the first derivative*

We start by explaining how to compute the first derivative of θ. For notational convenience, from now on we fix μ_0 such that $\theta(\mu_0) \in \mathcal{D}$ and we omit the dependence on μ as a subscript in families of circle maps. In addition, let us recall that we can write any conjugation as $\eta(x) = x + \xi(x)$ and denote by $\{\hat{\xi}_k\}_{k\in\mathbb{Z}}$ the Fourier coefficients of ξ. Finally, we denote the first derivatives as $\theta' = D_\mu\theta$ and $\hat{\xi}'_k = D_\mu\hat{\xi}_k$.

As we did in Section 2.2, we begin by computing the first averages (of the derivatives of the iterates) in order to illustrate the idea of the method. Thus, we proceed by formally taking derivatives with respect to μ at both sides of Eq. (2)

$$D_\mu f^n(x_0) = n\theta' + \sum_{k\in\mathbb{Z}} \hat{\xi}'_k \mathrm{e}^{2\pi \mathrm{i} k n\theta} + 2\pi \mathrm{i} n\theta' \sum_{k\in\mathbb{Z}} k\hat{\xi}_k \mathrm{e}^{2\pi \mathrm{i} k n\theta}.$$

Then, notice that a factor n appears, multiplying the second quasi-periodic sum. However, if we perform recursive sums, we can still manage to control the growth of this term due to the quasi-periodic part. Let us compute the sum

$$\begin{aligned} D_\mu S^1_N(f) &:= \sum_{n=1}^{N} D_\mu(f^n(x_0) - x_0) \\ &= \frac{N(N+1)}{2}\theta' - N\sum_{k\in\mathbb{Z}_*}\hat{\xi}'_k \\ &\quad + \sum_{k\in\mathbb{Z}_*}\hat{\xi}'_k \frac{\mathrm{e}^{2\pi\mathrm{i}k\theta}(1-\mathrm{e}^{2\pi\mathrm{i}kN\theta})}{1-\mathrm{e}^{2\pi\mathrm{i}k\theta}} \\ &\quad + 2\pi\mathrm{i}\theta'\sum_{k\in\mathbb{Z}_*} k\hat{\xi}_k \\ &\quad \times \frac{N\mathrm{e}^{2\pi\mathrm{i}k(N+2)\theta} - (N+1)\mathrm{e}^{2\pi\mathrm{i}k(N+1)\theta} + \mathrm{e}^{2\pi\mathrm{i}k\theta}}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^2}. \end{aligned}$$

Hence, we observe that the method is still valid, even though for $\theta' \neq 0$ the quasi-periodic sum is bigger than expected a priori. Indeed, we obtain the following formula

$$\frac{2}{N(N+1)} D_\mu S^1_N(f) = \theta' + \mathcal{O}(1/N), \tag{12}$$

that is similar to Eq. (3), but notice that the term $2A_1/(N+1)$ has been included in the remainder since there are oscillatory terms of the same order. Proceeding as in Section 2.2, we introduce *recursive sums* for derivatives of the iterates

$$D_\mu S^p_N(f) := D_\mu(f^N(x_0) - x_0),$$
$$D_\mu S^p_N(f) := \sum_{j=1}^{N} D_\mu S^{p-1}_j(f),$$

and their corresponding *averaged sums of order* p

$$D_\mu \tilde{S}^p_N(f) := \binom{N+p}{p+1}^{-1} D_\mu S^p_N(f).$$

Finding an expression like (12) for $p > 1$ is quite cumbersome to do directly, since the computations are very involved. However, the computation is straightforward if we take formal derivatives at both sides of Eq. (6). The resulting expression reads as

$$\begin{aligned} D_\mu \tilde{S}^p_N(f) &= \theta' + \sum_{l=1}^{p} \frac{D_\mu A^p_l}{(N+p-l+1)\cdots(N+p)} \\ &\quad + D_\mu E^p(N), \end{aligned}$$

where the new coefficients are $D_\mu A^p_l = (-1)^l(p-l+2)\cdots(p+1)D_\mu A_l$ with

$$D_\mu A_l = \sum_{k\in\mathbb{Z}_*} \frac{\mathrm{e}^{2\pi\mathrm{i}k(l-1)\theta}}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^{l-1}} \left(\hat{\xi}'_k + \frac{2\pi\mathrm{i}k(l-1)\hat{\xi}_k\theta'}{1-\mathrm{e}^{2\pi\mathrm{i}k\theta}} \right),$$

and the new remainder is

$$\begin{aligned} D_\mu E^p(N) &= (-1)^{p+1} \frac{(p+1)!}{N\cdots(N+p)} \sum_{k\in\mathbb{Z}_*} \frac{\mathrm{e}^{2\pi\mathrm{i}kp\theta}}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^p} \\ &\qquad \times \left\{ \hat{\xi}'_k(1-\mathrm{e}^{2\pi\mathrm{i}kN\theta}) + 2\pi\mathrm{i}k\hat{\xi}_k\theta' \right. \\ &\qquad \left. \times \left(p\frac{1-\mathrm{e}^{2\pi\mathrm{i}kN\theta}}{1-\mathrm{e}^{2\pi\mathrm{i}k\theta}} - N\mathrm{e}^{2\pi\mathrm{i}kpN\theta} \right) \right\}. \end{aligned}$$

Assuming that $\theta(\mu_0) \in \mathcal{D}$ and $\theta'(\mu_0) \neq 0$, we can obtain bounds analogous to those of Lemma 2.4 and conclude that the remainder satisfies $D_\mu E^p(N) = \mathcal{O}(1/N^p)$. Moreover, we observe that the coefficient $D_\mu A_p^p$ corresponds to a term of the same order, so we have to redefine the remainder in order to include this term. Hence, as we did in Eq. (7), we can arrange the unknown terms and obtain

$$D_\mu \tilde{S}_N^p(f) = \theta' + \sum_{l=1}^{p-1} \frac{D_\mu \hat{A}_l^p}{N^l} + \mathcal{O}(1/N^p),$$

where $\{D_\mu \hat{A}_l^p\}_{l=1,\dots,p-1}$ are derivatives of $\{\hat{A}_l^p\}_{l=1,\dots,p-1}$ that appear in Eq. (7).

Finally, we can extrapolate an approximation to θ' using Richardson's method of order $p-1$ as in Section 2.2. Concretely, if we compute $N = 2^q$ iterates, we can approximate the derivative of the rotation number by means of the following formula

$$\theta' = \sum_{j=0}^{p-1} c_j^{p-1} D_\mu \tilde{S}^p_{2^{q-p+1+j}}(f) + \mathcal{O}(2^{-pq}), \tag{13}$$

where the coefficients $\{c_j^{p-1}\}_{j=0,\dots,p-1}$ are given by (10).

3.2. *Computation of higher order derivatives*

The goal of this section is to generalize formula (13) to approximate $D_\mu^d \theta$ for any d, when they exist. Then, we assume that the family $\mu \mapsto f \in \mathrm{Diff}_+^\omega(\mathbb{T})$ depends $\mathcal{C}^d$-smoothly with respect to the parameter. As usual, we define the recursive sums for the d-derivative and their averages of order p as

$$D_\mu^d S_N^0(f) := D_\mu^d(f^n(x_0) - x_0),$$

$$D_\mu^d S_N^p(f) := \sum_{j=0}^{N} D_\mu^d S_j^{p-1}(f),$$

and

$$D_\mu^d \tilde{S}_N^p(f) := \binom{N+p}{p+1}^{-1} D_\mu^d S_N^p(f),$$

respectively. As before, we relate these sums to $D_\mu^d \theta$ by taking formal derivatives in Eq. (6), thus obtaining

$$D_\mu^d \tilde{S}_N^p(f) = D_\mu^d \theta + \sum_{l=1}^{p} \frac{D_\mu^d A_l^p}{(N+p-l+1)\cdots(N+p)} + D_\mu^d E^p(N). \tag{14}$$

It is immediate to check that, if $\theta(\mu_0) \in \mathcal{D}$ and $D_\mu^d\theta(\mu_0) \neq 0$, the remainder $D_\mu^d E^p(N)$ is of order $\mathcal{O}(1/N^{p-d+1})$, so this expression makes sense if the averaging order satisfies $p > d$.

Remark 3.1. Notice that in order to work with reasonable computational time and round-off errors, p cannot be taken arbitrarily large. Consequently, there is a (practical) limitation in the computation of high order derivatives.

In addition, as it was done for the first derivative, the remainder $D_\mu^d E^p(N)$ must be redefined in order to include the terms corresponding to $l \geq p-d+1$ in Eq. (14). Then we can extrapolate $D_\mu^d \theta$ by computing $N = 2^q$ iterates and solving the linear $(p-d+1)$-dimensional system associated to the following rearranged equation

$$D_\mu^d \tilde{S}_N^p(f) = D_\mu^d \theta + \sum_{l=1}^{p-d} \frac{D_\mu^d \hat{A}_l^p}{N^l} + \mathcal{O}(1/N^{p-d+1}). \tag{15}$$

Since the averaging order p and the extrapolation order $p-d$ do not coincide, let us define the *extrapolation operator of order m for the d-derivative* as

$$\Theta^d_{q,p,m}(f) := \sum_{j=0}^{m} c_j^m D_\mu^d \tilde{S}^p_{2^{q-m+j}}(f), \tag{16}$$

where coefficients $\{c_j^m\}_{j=0,\dots,m}$ are given by (10). Therefore, according to the formula (15), we can approximate the d-th derivative of the rotation number as

$$D_\mu^d \theta = \Theta^d_{q,p,p-d}(f) + \mathcal{O}(2^{-(p-d+1)q}).$$

Furthermore, as explained in Section 2.2, by comparing the approximations that correspond to 2^{q-1} and 2^q iterates, we obtain the following heuristic formula for the extrapolation error:

$$|D_\mu^d\theta - \Theta^d_{q,p,p-d}(f)| \leq \frac{\nu}{2^{p-d+1}} |\Theta^d_{q,p,p-d}(f) - \Theta^d_{q-1,p,p-d}(f)|, \tag{17}$$

where, once again, ν is a "safety parameter" that we take as $\nu = 10$.

Remark 3.2. Up to this point we have assumed that $D_\mu^d\theta \neq 0$ at the computed point. However, if we know a priori that $D_\mu^r \theta = 0$ for $r = 1,\dots,d$, then Eq. (14) holds with the following expression for the remainder:

$$D_\mu^d E^p(N) = (-1)^{p+1} \frac{(p+1)!}{N\cdots(N+p)} \times \sum_{k\in\mathbb{Z}_*} D_\mu^d \hat{\xi}_k \frac{\mathrm{e}^{2\pi \mathrm{i} k p\theta}(1-\mathrm{e}^{2\pi\mathrm{i}kN\theta})}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^p},$$

which now is of order $\mathcal{O}(1/N^{p+1})$. As in Section 2, this allows us to approximate $D_\mu^d\theta$ with the same extrapolation order as the averaging order p. Indeed, we obtain

$$0 = D_\mu^d \theta = \Theta^d_{q,p,p}(f) + \mathcal{O}(2^{-(p+1)q}),$$

and we observe that the order d is not limited by p.

The case remarked above is very interesting since we know that many applications can be modeled as a family of circle diffeomorphisms of fixed rotation number. The possibilities of this approach are illustrated by computing the Taylor expansion of Arnold Tongues (Section 4.3) and the continuation of invariant curves for the Hénon map (Section 5.3).

3.3. Scheme for evaluating the derivatives of the averaged sums

Let us introduce a recursive way for computing the sums $D_\mu^d \tilde{S}_N^p(f)$ required to evaluate the extrapolation operator (16). First of all, notice that by linearity it suffices to compute $D_\mu^d(f^n(x_0))$ for any $n \in \mathbb{N}$.

To compute the derivatives of $f^n = f \circ \overset{(n)}{\cdots} \circ f$, we proceed inductively with respect to n and d. Thus, let us assume that the derivatives $D_\mu^r(f^{n-1}(x_0))$ are known for a given $n \geq 1$ and for any $r \leq d$. Then, if we denote $z := f^{n-1}(x_0)$, our goal is to compute $D_\mu^r(f(z))$ for $r \leq d$ by using the known derivatives of z.

For $d = 1$, a recursive formula appears directly by applying the chain rule

$$D_\mu(f(z)) = \partial_\mu f(z) + \partial_x f(z) D_\mu(z). \tag{18}$$

This formula can be implemented provided the partial derivatives $\partial_\mu f$ and $\partial_x f$ can be numerically evaluated at the point z.

In general, we can perform higher order derivatives and obtain the following expression

$$\begin{aligned} D_\mu^d(f(z)) &= D_\mu^{d-1}\left(\partial_\mu f(z) + \partial_x f(z) D_\mu(z)\right) \\ &= D_\mu^{d-1}(\partial_\mu f(z)) + \sum_{r=0}^{d-1} \binom{d-1}{r} D_\mu^r(\partial_x f(z)) D_\mu^{d-r}(z). \end{aligned}$$

This motivates the extension of recurrence (18), since for evaluating the previous formula we require to know the derivatives $D_\mu^r(\partial_x f(z))$ for $r < d$ and $D_\mu^{d-1}(\partial_\mu f(z))$. We note that these derivatives can also be computed recursively using similar expressions for the maps $\partial_x f$ and $\partial_\mu f$, respectively. Concretely, assuming that we can evaluate $\partial_{\mu,x}^{i,j} f(z)$ for any $(i, j) \in \mathbb{Z}_+^2$ such that $i + j \leq d$, we can use the following recurrences

$$\begin{aligned} D_\mu^r(\partial_{\mu,x}^{i,j} f(z)) = {} & D_\mu^{r-1}(\partial_{\mu,x}^{i+1,j} f(z)) \\ & + \sum_{s=0}^{r-1} \binom{r-1}{s} D_\mu^s(\partial_{\mu,x}^{i,j+1} f(z)) D_\mu^{r-s}(z), \end{aligned}$$

to compute in a tree-like order the corresponding derivatives. To prevent redundant computations in implementation of the method, we store the value of "intermediate" derivatives $D_\mu^r(\partial_{\mu,x}^{i,j} f(z))$ so they only have to be computed once. For this reason, this scheme turns out to be more efficient than evaluating explicit expressions such as Faà di Bruno formulas (see for example [20]). Fig. 1 summarizes the recursive computations required and the convenience of storing these intermediate computations.

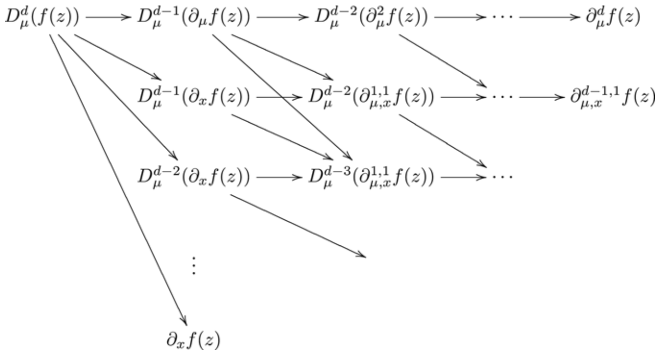

Fig. 1. Schematic representation of recurrent computations performed to evaluate $D_\mu^d(f(z))$.

Remark 3.3. The above scheme can be generalized immediately to the case of several parameters. For example, consider a two-parameter family $(\mu_1, \mu_2) \mapsto f_{\mu_1,\mu_2} \in \mathrm{Diff}_+^\omega(\mathbb{T})$ whose rotation number induces a map $(\mu_1, \mu_2) \mapsto \theta(\mu_1, \mu_2)$. Then, if $\theta(\mu_1^0, \mu_2^0) \in \mathcal{D}$, we can obtain a similar scheme to approximate $D_{\mu_1,\mu_2}^{d_1,d_2}\theta(\mu_1^0, \mu_2^0)$. In this context, note that the operator $\Theta_{q,p,p-d_1-d_2}^{d_1,d_2}$ can be defined as (16), but averaging the derivatives $D_{\mu_1,\mu_2}^{d_1,d_2}(f^n(x_0))$. Finally, if we write $z := f^{n-1}(x_0)$, we can compute inductively the derivatives $D_{\mu_1,\mu_2}^{m,l}(f(z))$, for $m \leq d_1$ and $l \leq d_2$, using the following recurrences

$$\begin{aligned} D_{\mu_1,\mu_2}^{m,l}(\partial_{\mu_1,\mu_2,x}^{i,j,k} f(z)) = {} & D_{\mu_1,\mu_2}^{m-1,l}(\partial_{\mu_1,\mu_2,x}^{i+1,j,k} f(z)) \\ & + \sum_{s=0}^{m-1}\sum_{r=0}^{l} \binom{m-1}{s}\binom{l}{r} D_{\mu_1,\mu_2}^{s,r}(\partial_{\mu_1,\mu_2,x}^{i,j,k+1} f(z)) \\ & \times D_{\mu_1,\mu_2}^{m-s,l-r}(z), \end{aligned}$$

if $m \neq 0$ and

$$\begin{aligned} D_{\mu_1,\mu_2}^{0,l}(\partial_{\mu_1,\mu_2,x}^{i,j,k} f(z)) = {} & D_{\mu_1,\mu_2}^{0,l-1}(\partial_{\mu_1,\mu_2,x}^{i,j+1,k} f(z)) \\ & + \sum_{r=0}^{l-1} \binom{l-1}{r} D_{\mu_1,\mu_2}^{0,r}(\partial_{\mu_1,\mu_2,x}^{i,j,k+1} f(z)) D_{\mu_1,\mu_2}^{0,l-r}(z), \end{aligned}$$

if $l \neq 0$. Of course, $D_{\mu_1,\mu_2}^{0,0}(\partial_{\mu_1,\mu_2,x}^{i,j,k} f(z)) = \partial_{\mu_1,\mu_2,x}^{i,j,k} f(z)$ corresponds to evaluation of the partial derivative of the map.

4. Application to the Arnold family

As a first example, let us consider the Arnold family of circle maps, given by

$$\begin{aligned} f_{\alpha,\varepsilon} : \mathbb{S} &\longrightarrow \mathbb{S} \\ x &\longmapsto x + 2\pi\alpha + \varepsilon \sin(x), \end{aligned} \tag{19}$$

where $(\alpha, \varepsilon) \in [0, 1) \times [0, 1)$ are parameters and $\mathbb{S} = \mathbb{R}/(2\pi\mathbb{Z})$. Notice that this family satisfies $f_{\alpha,\varepsilon} \in \mathrm{Diff}_+^\omega(\mathbb{S})$ for any value of the parameters. Let us remark that (19) allows to illustrate the method in a direct way, since there are explicit formulas for the partial derivatives $\partial_{\alpha,\varepsilon,x}^{i,j,k} f(x)$ of the map, for any $(i, j, k) \in \mathbb{Z}_+^3$. In Section 5 we will consider another interesting application in which the studied family is not given explicitly.

For this family of maps, it is convenient to take the angles modulo 2π just to avoid the loss of significant digits due to the factors $(2\pi)^{d-1}$ that would appear in the d-derivative of the map.

The contents of this section are organized as follows. First, in Section 4.1 we compute the derivative of a Devil's Staircase, that corresponds to the variation of the rotation number of (19) with respect to α for a fixed ε. In Section 4.2 we use the computation of derivatives of the rotation number

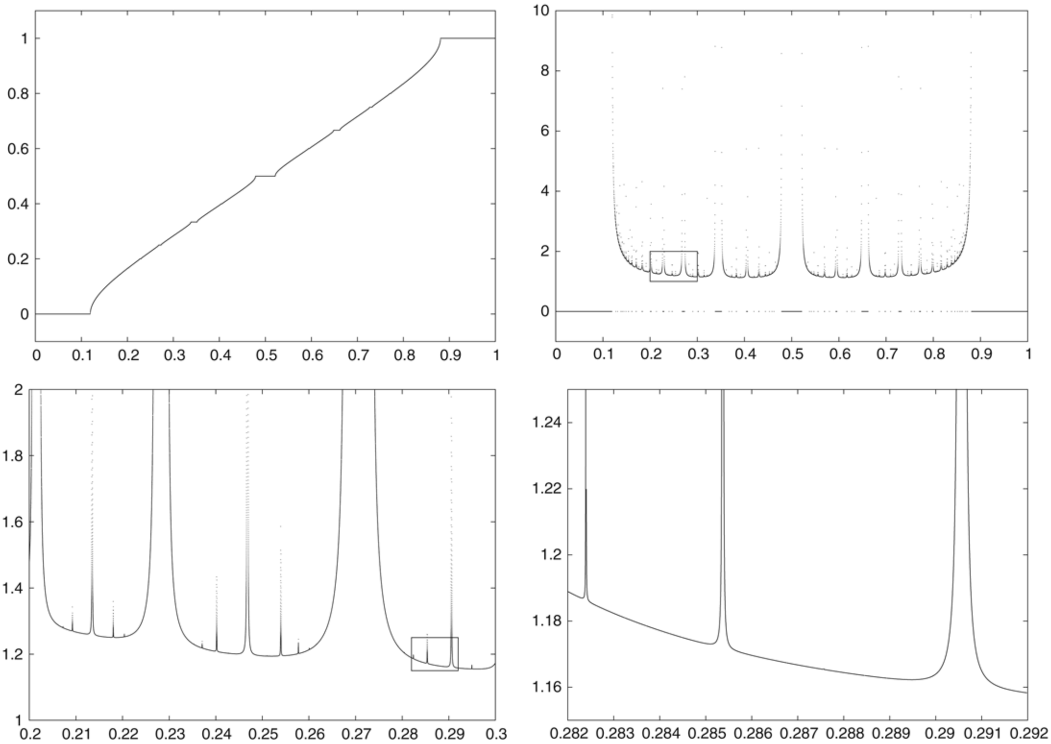

Fig. 2. Devil's Staircase $\alpha \mapsto \rho(f_\alpha)$ (top-left) and its derivative (top-right) for the Arnold family with $\varepsilon = 0.75$. The plots in the bottom correspond to some magnifications of the top-right one.

to approximate the Arnold Tongues of the family (19) by means of the Newton method. Furthermore, we compute the asymptotic expansion of these tongues and obtain pseudo-analytical expressions for the first coefficients, as a function of the rotation number.

4.1. Stepping up to a Devil's staircase

Let us fix the value of $\varepsilon \in [0, 1)$ and consider the one-parameter family $\{f_\alpha\}_{\alpha\in[0,1)}$ given by Eq. (19), i.e. $f_\alpha := f_{\alpha,\varepsilon}$. Let us recall that we can establish an ordering in this family since the normalized lifts satisfy $f_{\alpha_1}(x) < f_{\alpha_2}(x)$ for all $x \in \mathbb{R}$ if and only if $\alpha_1 < \alpha_2$. Then, we conclude that the function $\alpha \mapsto \rho(f_\alpha)$ is monotone increasing. In particular, for $\alpha_1 < \alpha_2$ such that $\rho(f_{\alpha_1}) \in \mathbb{R} \setminus \mathbb{Q}$ we have $\rho(f_{\alpha_1}) < \rho(f_{\alpha_2})$. On the other hand, if $\rho(f_{\alpha_1}) \in \mathbb{Q}$, there is an interval containing α_1 giving the same rotation number. As the values of α for which f_α has a rational rotation number are dense in [0, 1) (the complement is a Cantor set), there are infinitely many intervals where $\rho(f_\alpha)$ is locally constant. Therefore, the map $\alpha \mapsto \rho(f_\alpha)$ gives rise to a "staircase" with a dense number of stairs, that is usually called a Devil's Staircase (we refer to [9,18] for more details).

To illustrate the behavior of the method we have computed the above staircase for $\varepsilon = 0.75$. The computations have been performed by taking 10^4 points of $\alpha \in [0, 1)$, using 32-digit arithmetics (*double–double* data type from [17]), and a fixed averaging order $p = 8$. In addition, we estimate the error in the approximation of $\rho(f_\alpha)$ and $D_\alpha\rho(f_\alpha)$ using formulas (11) and (17), respectively. Then, we stop the computations for a tolerance of 10^{-26} and 10^{-24}, respectively, using at most $2^{22} = 4194304$ iterates.

Let us discuss the results obtained. First, we point out that only 11.4% of the selected points have not reached the previous tolerances for 2^{22} iterates. Moreover, we observe that the rotation number for 98.8% of the points has been obtained with an error less that 10^{-20}, while the estimated error in the derivatives is less than 10^{-18} for 97.7% of the points. Let us focus in $\alpha = 0.3377$, that is one of the "bad" points. The estimated errors for the rotation number and the derivative at this point are of order 10^{-18} and 10^{-9}, respectively. We observe that, even though this rotation number is irrational (the derivative does not vanish), it is very close to the rational 105/317, since $|317 \cdot \Theta_{22,9}(f_{0.3377}) - 105| \simeq 4.2 \times 10^{-6}$.

In Fig. 2 we show $\alpha \mapsto \rho(f_\alpha)$ and its derivative $\alpha \mapsto D_\alpha\rho(f_\alpha)$ for those points that satisfy that the estimated error is less than 10^{-18} and 10^{-16}, respectively. We recall that rational values of the rotation number correspond to constant intervals in the top-left plot, and note that by looking at the derivative (top-right plot) we can visualize the density of the stairs better than looking at the staircase itself. We remark that both these rational rotation numbers and their vanishing derivatives have been computed as well as in the Diophantine case.

Moreover, at the bottom of the same figure, we plot some magnifications of the derivative to illustrate non-smoothness of a Devil's Staircase. Concretely, the plot in the bottom-left corresponds to 10^5 values of $\alpha \in [0.2, 0.3]$ using the same implementation parameters as before. Once again, if the estimated error is bigger than 10^{-16} the point is not plotted. Finally, on the right plot we give another magnification for 10^6

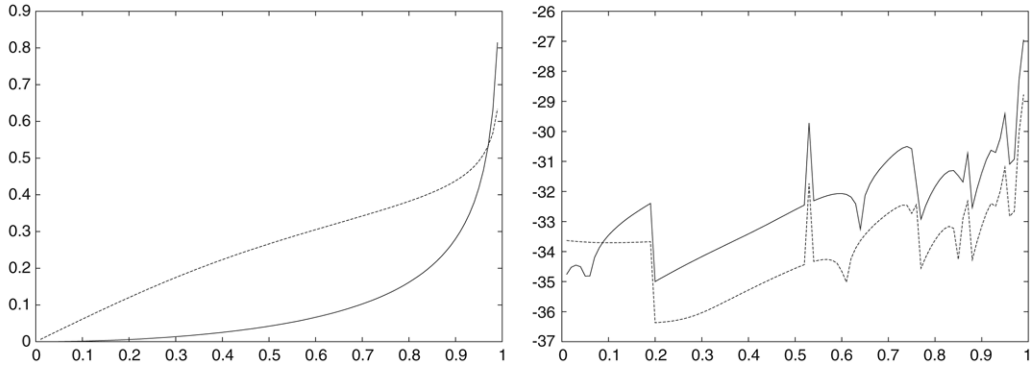

Fig. 3. Left: Graph of the derivatives $\varepsilon \mapsto D_\alpha \rho(\alpha(\varepsilon), \varepsilon)$ and $\varepsilon \mapsto D_\varepsilon \rho(\alpha(\varepsilon), \varepsilon)$ along T_θ, for $\theta = (\sqrt{5} - 1)/2$. The solid curve corresponds to $(D_\alpha \rho - 1)$ and the dashed one to $(20 \cdot D_\varepsilon \rho)$. Right: error (estimated using (11)) in $\log_{10}$ scale in the computation of these derivatives.

values of $\alpha \in [0.282, 0.292]$ that are computed with $p = 7$, and allowing at most $2^{21} = 2097152$ iterates. In this case, points that correspond to the branch in the left (i.e. close to $\alpha = 0.2825$), are typically computed with an error 10^{-10}.

4.2. Newton method for computing the Arnold Tongues

Since $f_{\alpha,\varepsilon} \in \mathrm{Diff}_+^\omega(\mathbb{S})$, we obtain a function $(\alpha, \varepsilon) \mapsto \rho(\alpha, \varepsilon) := \rho(f_{\alpha,\varepsilon})$ given by the rotation number. Then, the Arnold Tongues of (19) are defined as the sets $T_\theta = \{(\alpha, \varepsilon) : \rho(\alpha, \varepsilon) = \theta\}$, for any $\theta \in [0, 1)$. It is well known that if $\theta \in \mathbb{Q}$, then T_θ is a set with interior; otherwise, T_θ is a continuous curve which is the graph of a function $\varepsilon \mapsto \alpha(\varepsilon)$, with $\alpha(0) = \theta$. In addition, if $\theta \in \mathcal{D}$, the corresponding tongue is given by an analytic curve (see [25]).

Using the method described in Section 2.2, some Arnold Tongues T_θ of Diophantine rotation number were approximated in [27] by means of the secant method. Now, since we can compute derivatives of the rotation number, we are able to repeat the computations using a Newton method. To do that, we fix $\theta \in \mathcal{D}$ and solve the equation $\rho(\alpha, \varepsilon) - \theta = 0$ by continuing the known solution $(\theta, 0)$ with respect to ε. Indeed, we fix a partition $\{\varepsilon_j\}_{j=0,\ldots,K}$ of $[0, 1)$, and compute a numerical approximation α_j^* for every $\alpha(\varepsilon_j)$.

To this end, assume that we have a good approximation α_{j-1}^* to $\alpha(\varepsilon_{j-1})$ and let us first compute an initial approximation for $\alpha(\varepsilon_j)$. Taking derivatives in the equation $\rho(\alpha(\varepsilon), \varepsilon) - \theta = 0$ we obtain

$$D_\alpha \rho(\alpha(\varepsilon), \varepsilon)\alpha'(\varepsilon) + D_\varepsilon \rho(\alpha(\varepsilon), \varepsilon) = 0. \tag{20}$$

Thus, we can approximate $\alpha'(\varepsilon_{j-1})$ by computing numerically the derivatives $D_\alpha \rho$ and $D_\varepsilon \rho$ at $(\alpha_{j-1}^*, \varepsilon_{j-1})$. Hence, we obtain an approximated value $\alpha_j^{(0)} = \alpha_{j-1}^* + \alpha'(\varepsilon_{j-1})(\varepsilon_j - \varepsilon_{j-1})$ for $\alpha(\varepsilon_j)$. Next, we apply the Newton method

$$\alpha_j^{(n+1)} = \alpha_j^{(n)} - \frac{\rho(\alpha_j^{(n)}, \varepsilon_j) - \theta}{D_\alpha \rho(\alpha_j^{(n)}, \varepsilon_j)},$$

and stop when we converge to a value α_j^* that approximates $\alpha(\varepsilon_j)$.

Computations are performed using 64 digits (*quadruple-double* data type from [17]) and, in order to compare with the results obtained in [27], we select the same parameters in the implementation. In particular, we take a partition $\varepsilon_j = j/K$ with $K = 100$ of the interval $[0, 1)$, we select an averaging order $p = 9$ and allow at most $2^{23} = 8388608$ iterates of the map. The required tolerances are taken as 10^{-32} for computation of the rotation number (we use (11) to estimate the error) and 10^{-30} for convergence of the Newton method. Let us remark that computations are done without any prescribed tolerance for computation of derivatives $D_\alpha \rho$ and $D_\varepsilon \rho$, even though we check, using (17), that the extrapolation is done correctly.

Let us discuss the results obtained for $\theta = (\sqrt{5} - 1)/2$. As expected, the number of iterates of the Newton method is less than the ones required by the secant method. Concretely, we perform from 2 to 3 corrections as we approach the critical value $\varepsilon = 1$, while using the secant method we need at least 4 steps to converge. However, we observe that computation of the derivatives $D_\alpha \rho$ and $D_\varepsilon \rho$ fails if we take $\varepsilon = 1$, even though the secant method converges after 18 iterations. This is totally consistent since we know that $f_{\alpha,1} \in \mathrm{Diff}_+^0(\mathbb{T})$ but is still an analytic map, and that the conjugation to a rigid rotation is only Hölder continuous (see [8,34]).

In Fig. 3(left) we plot the derivatives $\varepsilon \mapsto D_\alpha \rho(\alpha(\varepsilon), \varepsilon)$ and $\varepsilon \mapsto D_\varepsilon \rho(\alpha(\varepsilon), \varepsilon)$ evaluated on the previous tongue. We observe that the derivatives have been normalized in order to fit together in the same plot. On the other hand, in the right plot we show the estimated error in the computation of these derivatives (obtained from Eq. (17)). In the worst case, $\varepsilon = 0.99$, we obtain errors of order 10^{-27} and 10^{-29} for $D_\alpha \rho$ and $D_\varepsilon \rho$, respectively.

4.3. Computation of the Taylor expansion of the Arnold Tongues

As we have mentioned in Section 4.2, if $\theta \in \mathcal{D}$ then the Arnold Tongue T_θ of (19) is given by the graph of an analytic function $\alpha(\varepsilon)$, for $\varepsilon \in [0, 1)$. Then, we can expand α at the origin as

$$\alpha(\varepsilon) = \theta + \frac{\alpha'(0)}{1!}\varepsilon + \frac{\alpha''(0)}{2!}\varepsilon^2 + \cdots + \frac{\alpha^{(d)}(0)}{d!}\varepsilon^d + \mathcal{O}(\varepsilon^{d+1}), \tag{21}$$