

# On the length and area spectrum of analytic convex domains

**Pau Martín<sup>†</sup>, Rafael Ramírez-Ros<sup>‡</sup> and Anna Tamarit-Sariol<sup>‡</sup>**

<sup>†</sup>Departament de Matemàtica Aplicada IV, Universitat Politècnica de Catalunya, Ed. C3,  
Jordi Girona 1–3, 08034 Barcelona, Spain

<sup>‡</sup>Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647,  
08028 Barcelona, Spain

E-mail: [Martin@ma4.upc.edu](mailto:Martin@ma4.upc.edu), [Rafael.Ramirez@upc.edu](mailto:Rafael.Ramirez@upc.edu),  
[Anna.Tamarit@upc.edu](mailto:Anna.Tamarit@upc.edu)

**Abstract.** Area-preserving twist maps have at least two different  $(p, q)$ -periodic orbits and every  $(p, q)$ -periodic orbit has its  $(p, q)$ -periodic action for suitable couples  $(p, q)$ . We establish an exponentially small upper bound for the differences of  $(p, q)$ -periodic actions when the map is analytic on a  $(m, n)$ -resonant rotational invariant curve (resonant RIC) and  $p/q$  is “sufficiently close” to  $m/n$ . The exponent in this upper bound is closely related to the analyticity strip width of a suitable angular variable. The result is obtained in two steps. First, we prove a Neishtadt-like theorem, in which the  $n$ -th power of the twist map is written as an integrable twist map plus an exponentially small remainder on the distance to the RIC. Second, we apply the MacKay-Meiss-Percival action principle.

We apply our exponentially small upper bound to several billiard problems. The resonant RIC is a boundary of the phase space in almost all of them. For instance, we show that the lengths (respectively, areas) of all the  $(1, q)$ -periodic billiard (respectively, dual billiard) trajectories inside (respectively, outside) analytic strictly convex domains are exponentially close in the period  $q$ . This improves some classical results of Marvizi, Melrose, Colin de Verdière, Tabachnikov, and others about the smooth case.

AMS classification scheme numbers: 37E10, 37E40, 37J40, 52A10

*Keywords:* twist maps, invariant curves, exponential smallness, billiards, dual billiards

## 1. Introduction

Billiards were introduced by Birkhoff [4]. Let  $\Gamma$  be a smooth strictly convex curve in the plane, oriented counterclockwise, and let  $\Omega$  be the billiard table enclosed by  $\Gamma$ . Billiard trajectories inside  $\Omega$  consist of polygonal lines inscribed in  $\Gamma$  whose consecutive sides obey to the rule “the angle of reflection is equal to the angle of incidence.” See [18, 20, 36] for a general description.

A  $(p, q)$ -periodic billiard trajectory forms a closed polygon with  $q$  sides that makes  $p$  turns inside  $\Gamma$ . Birkhoff [4] proved that there are at least two different Birkhoff  $(p, q)$ -periodic billiard trajectories inside  $\Omega$  for any relatively prime integers  $p$  and  $q$  such that  $1 \leq p < q$ .

Let  $\mathcal{L}^{(p,q)}$  be the supremum of the absolute values of the differences of the lengths of all such trajectories. The quantities  $\mathcal{L}^{(p,q)}$  were already studied by Marvizi and Melrose [25]

and Colin de Verdière [6] for smooth tables. The former authors produced an asymptotic expansion of the lengths for  $(p, q)$ -periodic billiard trajectories approaching  $\Gamma$  when  $p$  is fixed and  $q \rightarrow +\infty$ . They saw that there exists a sequence  $(l_k)_{k \geq 1}$ , depending only on  $p$  and  $\Gamma$ , such that, if  $L^{(p,q)}$  is the length of any  $(p, q)$ -periodic trajectory, then

$$L^{(p,q)} \asymp p \text{Length}[\Gamma] + \sum_{k \geq 1} \frac{l_k}{q^{2k}}, \quad q \rightarrow \infty, \quad (1)$$

where  $l_1 = l_1(\Gamma, p) = -\frac{1}{24} \left( p \int_{\Gamma} \kappa^{2/3}(s) ds \right)^3$ , and  $\kappa(s)$  is the curvature of  $\Gamma$  as a function of the arc-length parameter  $s$ . The symbol  $\asymp$  means that the series in the right hand side is asymptotic to  $L^{(p,q)}$ . The asymptotic coefficients  $l_k = l_k(\Gamma, p)$  can be explicitly written in terms of the curvature  $\kappa(s)$ . For instance, the explicit formulas for  $l_1, l_2, l_3$ , and  $l_4$  can be found in [35]. Since the expansion of the lengths in powers of  $q^{-1}$  coincides for all these  $(p, q)$ -periodic trajectories,  $\mathcal{L}^{(p,q)} = O(q^{-\infty})$  for smooth strictly convex tables when  $p$  is fixed and  $q \rightarrow +\infty$ . Colin de Verdière studied the lengths of periodic trajectories close to an elliptic  $(1, 2)$ -periodic trajectory on a smooth axisymmetric billiard table, and found that the quantities  $\mathcal{L}^{(p,q)}$  are again beyond all order with respect to  $q$ .

These works suggest that the supremum length differences  $\mathcal{L}^{(p,q)}$  are exponentially small in the period  $q$  for analytic strictly convex tables. Indeed, we have proved that if  $\Gamma$  is analytic and  $p$  is a fixed positive integer, then there exist  $K, \alpha > 0$  such that

$$\mathcal{L}^{(p,q)} \leq K e^{-2\pi\alpha q/p}, \quad (2)$$

for all integer  $q \geq 2$  relatively prime with  $p$ . The exponent  $\alpha$  is related to the width of a complex strip where a certain 1-periodic angular coordinate is analytic. A more precise statement is given in Theorem 5. The search for exponentially small asymptotic formulas is the natural challenge after obtaining the exponentially small upper bound (2). This problem has been numerically tackled in [24], where we have conjectured that, if  $\Gamma$  is a generic axisymmetric strictly convex algebraic curve, then

$$\mathcal{L}^{(1,q)} \asymp A(q) q^{-3} e^{-rq}, \quad q \rightarrow +\infty,$$

where the positive exponent  $r$  is half of the radius of convergence of the Borel transform of the asymptotic series (1) and  $A(q)$  is either a constant or an oscillating function. The proof of this asymptotic formula is a work in progress.

Similar exponentially small upper bounds hold in other billiard problems. We mention two examples. First, for  $(p, q)$ -periodic billiard trajectories inside strictly convex analytic tables of constant width when  $p/q \rightarrow 1/2$ . Second, for  $(p, q)$ -periodic billiard trajectories inside strictly convex analytic tables in surfaces of constant curvature when  $p/q \rightarrow 0$ .

The billiard dynamics close to the boundary has also been studied from the point of view of KAM theory. Lazutkin [21] proved that there are infinitely many caustics inside any  $C^{555}$  strictly convex table. These caustics accumulate at the boundary of the table, and have Diophantine rotation numbers. Douady [11] improved the result to  $C^7$  billiard tables.

A special remark on the relevance of these results is the following. Kac [17] formulated the inverse spectral problem for planar domains. That is, to study how much geometric information about  $\Omega$  can be obtained from the Laplacian spectrum with homogeneous

Dirichlet conditions on  $\Gamma$ . Andersson and Melrose [1] gave an explicit relation between the length spectrum and the Laplacian spectrum. The length spectrum of  $\Omega$  is the union of the lengths of all its  $(p, q)$ -periodic billiard trajectories and all the integer multiples of  $\text{Length}[\Gamma]$ . See also [25, 6].

Our results also apply to the dual billiards introduced by Day [9] and popularized by Moser [28] as a crude model for planetary motion. Some general references are [15, 5, 37, 36]. Let  $\mathcal{U}$  be unbounded component of  $\mathbb{R}^2 \setminus \Gamma$ . The dual billiard map  $f : \mathcal{U} \rightarrow \mathcal{U}$  is defined as follows:  $f(z)$  is the reflection of  $z$  in the tangency point of the oriented tangent line to  $\Gamma$  through  $z$ . Billiards and dual billiards are projective dual in the sphere [37].

A  $(p, q)$ -periodic dual billiard trajectory forms a closed circumscribed polygon with  $q$  sides that makes  $p$  turns outside  $\Gamma$ . The area of a  $(p, q)$ -periodic trajectory is the area enclosed by the corresponding polygon, taking into account some multiplicities if  $p \geq 2$ . There are at least two different Birkhoff  $(p, q)$ -periodic dual billiard trajectories outside  $\Gamma$  for any relatively prime integers  $p$  and  $q$  such that  $q \geq 3$  and  $1 \leq p < q$ .

Tabachnikov [36, 37] studied the supremum  $\mathcal{A}^{(1,q)}$  of the absolute value of the differences of the areas enclosed by all such  $(1, q)$ -periodic trajectories for smooth tables. He proved that there is a sequence  $(a_k)_{k \geq 1}$ , depending only on  $\Gamma$ , such that, if  $A^{(1,q)}$  is the area enclosed by any  $(1, q)$ -periodic dual billiard trajectory, then

$$A^{(1,q)} \asymp \text{Area}[\Omega] + \sum_{k \geq 1} \frac{a_k}{q^{2k}}, \quad q \rightarrow \infty, \quad (3)$$

where  $a_1 = a_1(\Gamma) = \frac{1}{24} \int_{\Gamma} \kappa^{1/3}(s) ds$ . Hence, the expansion of the areas in powers of  $q^{-1}$  coincides for all these  $(1, q)$ -periodic trajectories, and so,  $\mathcal{A}^{(1,q)} = O(q^{-\infty})$  for smooth strictly convex dual tables when  $q \rightarrow +\infty$ . Douady [11] found the existence of infinitely many invariant curves outside any  $C^7$  strictly convex dual table. These invariant curves accumulate at the boundary of the dual table and have Diophantine rotation numbers.

In a completely analogous way to (classical) billiards, we have proved that, once fixed any positive integer  $p$ , if  $\Gamma$  is analytic, then there exist  $K, \alpha > 0$  such that

$$\mathcal{A}^{(p,q)} \leq K e^{-2\pi\alpha q/p}, \quad (4)$$

for all integer  $q \geq 3$  relatively prime with  $p$ . Once more, the exponent  $\alpha$  is related to the width of a complex strip where a certain 1-periodic angular coordinate is analytic. The precise statement is given in Theorem 8.

Still in the context of dual billiards, the points at infinity can be seen as  $(1, 2)$ -periodic points, hence they form a  $(1, 2)$ -resonant RIC. Douady [11] found the existence of infinitely many invariant curves outside any  $C^8$  strictly convex dual table. These invariant curves accumulate at infinity and have Diophantine rotation numbers. We have proved that, once fixed any constant  $L \geq 1$ , if  $\Gamma$  is analytic, then there exist  $K, \alpha > 0$  such that

$$\mathcal{A}^{(p,q)} \leq K \exp\left(-\frac{\pi\alpha}{|p/q - 1/2|}\right), \quad (5)$$

for all relatively prime integers  $p$  and  $q$  such that  $1 \leq |2p - q| \leq L$  and  $q \geq 3$ . See Theorem 9.

The three exponents  $\alpha$  that appear in the exponentially small upper bounds (2), (4), and (5) may be different, since each one is associated to a different analyticity strip width.

Besides, all of these upper bounds follow directly from a general upper bound about analytic area-preserving twist maps with analytic resonant RICs. Let us explain it.

Classical and dual billiard maps are exact twist maps defined on an open cylinder when written in suitable coordinates. Exact twist maps have been vastly studied. They satisfy a Lagrangian formulation and their orbits are stationary points of the action functional. See for instance [4, 27, 18].

Birkhoff [4] showed that the minima and minimax points of the  $(p, q)$ -periodic action correspond to two different Birkhoff  $(p, q)$ -periodic orbits of the twist map. A Birkhoff  $(p, q)$ -periodic orbit is an orbit such that, after  $q$  iterates, performs exactly  $p$  revolutions around the cylinder and its points are ordered in the base  $\mathbb{T}$  as the ones following a rigid rotation of angle  $p/q$ . Since there exist at least two different Birkhoff  $(p, q)$ -periodic orbits, we consider the supremum  $\Delta^{(p,q)}$  of the absolute value of the differences of the actions among all of them. The quantity  $\Delta^{(p,q)}$  coincides with  $\mathcal{L}^{(p,q)}$  and  $\mathcal{A}^{(p,q)}$  for classical and dual billiards, respectively.

Let  $\Delta W_{p/q}$  be the difference of actions between the minimax and minima  $(p, q)$ -periodic orbits. Note that  $\Delta^{(p,q)}$  is an upper bound of  $\Delta W_{p/q}$ . Mather [26] used  $\Delta W_{p/q}$  as a criterion to prove the existence of RICs of given irrational rotation numbers. More concretely, he proved that there exists a RIC with irrational rotation number  $\varrho$  if and only if  $\lim_{p/q \rightarrow \varrho} \Delta W_{p/q} = 0$ .

Another criterion related to the destruction of RICs, in this case empirical, was proposed by Greene. The destruction of a RIC with Diophantine rotation number  $\varrho$  under perturbation is related to a ‘‘sudden change from stability to instability of the nearby periodic orbits’’ [13]. The stability of a periodic orbit is measured by the residue. MacKay [22] proved the criterion in some contexts. In particular, for an analytic area-preserving twist map, the residue of a sequence of periodic orbits with rotation numbers  $p/q \rightarrow \varrho$  decays exponentially in  $|\varrho - p/q|$ . The same proof leads to a similar exponentially small bound of Mather’s  $\Delta W_{p/q}$  as  $p/q \rightarrow \varrho$ . Delshams and de la Llave [10] studied similar problems for analytic area-preserving nontwist maps.

Generically RICs with a rational rotation number break under perturbation [33, 31]. Nevertheless, there are situations in which some distinguished resonant RICs always exist. See Sections 3 and 4 for several examples related to billiard and dual billiard maps.

Let us assume that we have an analytic exact twist map with a  $(m, n)$ -resonant RIC. That is, a RIC whose points are  $(m, n)$ -periodic. Then there exist some variables  $(x, y)$  in which the resonant RIC is located at  $\{y = 0\}$  and the  $n$ -th power of the exact twist map is a small perturbation of the integrable twist map  $(x_1, y_1) = (x + y, y)$ . To be precise, it has the form

$$x_1 = x + y + O(y^2), \quad y_1 = y + O(y^3).$$

Since the  $n$ -th power map is real analytic, it can be extended to a complex domain of the form

$$D_{a_*, b_*} := \{(x, y) \in \mathbb{C}/\mathbb{Z} \times \mathbb{C} : |\Im x| < a_*, |y| < b_*\}.$$

The quantity  $a_*$  plays a more important role than  $b_*$ . To be precise, we have proved that, once fixed any  $\alpha \in (0, a_*)$  and  $L \geq 1$ , there exists  $K > 0$  such that

$$\Delta^{(p,q)} \leq K \exp\left(-\frac{2\pi\alpha q}{|np - mq|}\right),$$

for any relatively prime integers  $p$  and  $q$  such that  $1 \leq |np - mq| \leq L$  and  $q \geq 1$ . See Theorem 3 for a more detailed statement. This upper bound is optimal because  $\alpha \in (0, a_*)$ . That is, the exponent  $\alpha$  can be taken as close to the analyticity strip width  $a_*$  as desired. The constant  $K$  may explode when  $\alpha$  tends to  $a_*$ , so, in general, we can not take  $\alpha = a_*$ . A similar optimal exponentially small upper bound was obtained in [12] in the setting of the splitting of separatrices of weakly hyperbolic fixed points of analytical area-preserving maps. The proof of this optimal bound adds some extra technicalities, but we feel that the effort is worth it.

The proof is based on two facts. First, we write the  $n$ -th power of the exact twist map as the integrable twist map  $(x_1, y_1) = (x + y, y)$  plus an exponentially small remainder on the distance to the RIC. See Theorem 2. The size of the remainder is reduced by performing a finite sequence of changes of variables, but the number of such changes increases when we approach to the resonant RIC. This is a classical Neishtadt-like argument [29]. Second, we apply the MacKay-Meiss-Percival action principle [23], in which the difference of actions of  $(p, q)$ -periodic orbits is interpreted as an area on the phase space.

The structure of the paper is the following. Section 2 is devoted to state our results in the general context of analytic exact twist maps. In Sections 3 and 4, we present the different billiard maps and show how the results in Section 2 apply. Sections 5 and 6 contain the technical proofs.

## 2. Main theorems

### 2.1. Dynamics close to an analytic resonant RIC

We consider analytic maps defined on a neighbourhood (in the cylinder) of a resonant RIC. We also assume that these maps have the intersection property, at least in one side of the RIC. Our goal is to show that these maps are exponentially close to an integrable one in the distance to the resonant RIC. No exactness or area-preserving condition is required.

Let  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ . Let  $I$  be an interval of the real line.

**Definition 1.** A continuous map  $g : \mathbb{T} \times I \rightarrow \mathbb{T} \times \mathbb{R}$  has the intersection property if the image of any closed homotopically non trivial loop of the cylinder  $\mathbb{T} \times I$  intersects the loop.

The intersection property is preserved under global changes of variables.

**Definition 2.** Let  $g : \mathbb{T} \times I \rightarrow \mathbb{T} \times \mathbb{R}$  be a continuous map. A rotational invariant curve (RIC) of  $g$  is a closed loop  $C \subset \mathbb{T} \times I$  homotopically non trivial such that  $g(C) = C$ . Let  $m$  and  $n$  be two relatively prime integers such that  $n \geq 1$ . We say that  $C$  is  $(m, n)$ -resonant when  $G^n(s, r) = (s + m, r)$  for all  $(s, r) \in C$ , where  $G(s, r)$  is a lift of  $g$ .

We want to study the dynamics of an analytic map  $g$  with the intersection property in a neighbourhood of an  $(m, n)$ -resonant RIC  $C$ . First, we note that all points on the RIC  $C$  remain fixed under the power map  $f = g^n$ . Second, we recall a classical lemma that appears in several papers about billiards [21, 37]. We present the proof for completeness.

**Lemma 1** (cf. [37]). *Let  $\beta > 0$ . Let  $f : \mathbb{T} \times (-\beta, \beta) \rightarrow \mathbb{T} \times \mathbb{R}$ ,  $f(s, r) = (s_1, r_1)$ , be a real analytic map with the intersection property of the form*

$$s_1 = s + \varphi(s)r + O(r^2), \quad r_1 = r + \psi(s)r^2 + O(r^3), \quad (6)$$

*for some real analytic 1-periodic functions  $\varphi(s)$  and  $\psi(s)$ , and  $\varphi(s)$  has no zeros. Then there exist some new analytic coordinates  $(x, y)$  in which  $f$  has the form  $(x_1, y_1) = f(x, y)$ , with*

$$x_1 = x + y + O(y^2), \quad y_1 = y + O(y^3). \quad (7)$$

*Besides, there exist some analyticity strip width  $a_* > 0$  and some analyticity radius  $b_* > 0$  such that  $f$  is real analytic on  $\mathbb{T} \times (-b_*, b_*)$  and can be analytically extended to the domain*

$$D_{a_*, b_*} = \{(x, y) \in (\mathbb{C}/\mathbb{Z}) \times \mathbb{C} : |\Im x| < a_*, |y| < b_*\}. \quad (8)$$

*The same result holds when  $f : \mathbb{T} \times [0, \beta) \rightarrow \mathbb{T} \times \mathbb{R}$  or when  $f : \mathbb{T} \times (-\beta, 0] \rightarrow \mathbb{T} \times \mathbb{R}$  is a real analytic map with the intersection property of the form (6).*

*Proof.* The intersection property implies that the quotient  $\psi(s)/\varphi(s)$  has zero average. Let us prove this claim. Set  $\mu = \int_0^1 (\psi(s)/\varphi(s)) ds$ . We consider the coordinates  $(x, y)$  defined by

$$x = a(s), \quad y = b(s)r, \quad (9)$$

where the real analytic functions  $a(s)$  and  $b(s)$  have the form

$$a(s) = \int_0^s \frac{b(t)}{\varphi(t)} dt, \quad b(s) = \nu \exp \left[ - \int_0^s \left( \frac{\psi(t)}{\varphi(t)} - \mu \right) dt \right] \quad (10)$$

for some constant  $\nu \neq 0$ . The function  $b(s)$  is 1-periodic and has no zeros. The constant  $\nu$  is determined in such a way that  $a(s+1) = a(s) + 1$ , so the new angular coordinate  $x$  is defined modulus one:  $x \in \mathbb{T}$ . A straightforward computation shows that

$$x_1 = x + y + O(y^2), \quad y_1 = y + \mu\varphi(s)y^2/b(s) + O(y^3).$$

If  $\mu \neq 0$ , then the map  $(x, y) \mapsto (x_1, y_1)$  does not have the intersection property. Hence,  $\mu = 0$  and the map  $f$  has the form (7) in the analytic variables  $(x, y)$ . These coordinates  $(x, y)$  cover an open complex neighbourhood of  $\{y = 0\}$ . In particular,  $f$  can be analytically extended to the complex domain  $D_{a_*, b_*}$  for some  $a_*, b_* > 0$ .  $\square$

The map (7) can be viewed as a perturbation of the integrable twist map

$$x_1 = x + y, \quad y_1 = y. \quad (11)$$

We want to reduce the size of the nonintegrable terms  $O(y^2)$  and  $O(y^3)$  as much as possible. We can reduce them through normal form steps up to any desired order; see Lemma 10. Thus, the nonintegrable part of the dynamics is beyond all order in  $y$ ; that is, in the distance to  $C$ . A general principle in conservative dynamical systems states that beyond all order phenomena are often exponentially small in the analytic category. Our goal is to write the map as an exponentially small perturbation in  $y$  of the integrable twist map (11). The final result is stated in the following theorem. If  $h$  is a real-valued smooth function,  $\partial_i h$  denotes the derivative with respect to the  $i$ -th variable.

**Theorem 2.** Let  $a_*, b_* > 0$ . Let  $f : \mathbb{T} \times (-b_*, b_*) \rightarrow \mathbb{T} \times \mathbb{R}$ ,  $(x_1, y_1) = f(x, y)$ , be a real analytic map of the form (7) with the intersection property on  $\mathbb{T} \times [0, b_*)$  or  $\mathbb{T} \times (-b_*, 0]$  that can be analytically extended to the complex domain (8). Let  $m \geq 2$  be an arbitrary order. Let  $\alpha \in (0, a_*)$ . There exist constants  $K > 0$  and  $b'_* \in (0, b_*)$  such that, if  $b \in (0, b'_*)$ , then there exists an analytic change of variables  $(x, y) = \Phi(\xi, \eta)$  such that:

(i) It is uniformly (with respect to  $b$ ) close to the identity on  $\mathbb{T} \times (-b, b)$ . That is,

$$x = \xi + O(\eta), \quad y = \eta + O(\eta^2), \quad \det[D\Phi(\xi, \eta)] = 1 + O(\eta),$$

for all  $(\xi, \eta) \in \mathbb{T} \times (-b, b)$ , where the  $O(\eta)$  and  $O(\eta^2)$  terms are uniform in  $b$ ; and

(ii) The transformed map  $(\xi, \eta) \mapsto (\xi_1, \eta_1)$  is real analytic on the cylinder  $\mathbb{T} \times (-b, b)$ . Besides, it has the form

$$\xi_1 = \xi + \eta + \eta^m g_1(\xi, \eta), \quad \eta_1 = \eta + \eta^{m+1} g_2(\xi, \eta), \quad (12)$$

where  $|g_j(\xi, \eta)| \leq K e^{-2\pi\alpha/b}$  and  $|\partial_i g_j(\xi, \eta)| \leq K b^{-2}$  for all  $(\xi, \eta) \in \mathbb{T} \times (-b, b)$ .

The proof can be found in Section 5.

*Remark 1.* We require the intersection property on one side of the resonant RIC  $C = \mathbb{T} \times \{0\}$  only. This is useful to study the billiard (and dual billiard) maps considered later on, because the boundaries of the natural phase spaces of these maps can be viewed as resonant RICs.

Consider a perturbed Hamiltonian system which is close to an integrable system. It is known that, under the appropriate nondegeneracy conditions, the measure of the set of tori which decompose under the perturbation can be bounded from above by a quantity of order  $\sqrt{\epsilon}$ ,  $\epsilon$  being the perturbation parameter [29, 32]. Neishtadt [29] also considered a context where the perturbation becomes exponentially small in some parameter  $\epsilon$  and hence the measure of the complementary set which is cut out from phase space by the invariant tori is of order  $e^{-c/\epsilon}$ ,  $c$  being a positive constant. This argument could be applied to our context. First, any neighbourhood of an analytic RIC of fixed points of an analytic map contains infinitely many RICs. Second, the area of the complementary of the RICs in any of such neighbourhoods is exponentially small in the size of the neighbourhood. Third, the gaps between the RICs are exponentially small in their distance to the RIC of fixed points. The first result follows from the classical Moser twist theorem [34]. The others follow from the ideas explained above.

## 2.2. Difference of periodic actions

In this subsection we consider real analytic exact twist maps defined on open cylinders that can be extended as rigid rotations on the boundaries of the cylinder. This has to do with the fact that we look for a global result. See Remark 2 below.

Let  $D \subset \mathbb{T} \times \mathbb{R}$  be the open cylinder enclosed between two homotopically non trivial loops,  $C_-$  and  $C_+$ . We assume that  $C_-$  is strictly below  $C_+$ . Let  $\omega = -d\lambda$  be an exact symplectic form on  $D$ . The symplectic form  $\omega$  may be degenerate or not defined on the boundaries  $\partial D = C_- \cup C_+$ .

**Definition 3.** A smooth diffeomorphism  $g : D \rightarrow D$  is an exact twist map when it preserves the exact symplectic form  $\omega = -d\lambda$ , has zero flux, and satisfies the classical twist condition  $\partial_2 s_1(s, r) > 0$ , where  $G(s, r) = (s_1, r_1)$  is a lift of  $g$ .

Henceforth, the exact symplectic form  $\omega = -d\lambda$  and the lift  $G$  remain fixed. We stress that, if  $g : D \rightarrow D$  is a diffeomorphism preserving  $\omega$  and  $g$  has a RIC  $C$ , then both the zero flux condition and the intersection property are automatically satisfied.

If an exact twist map  $g : D \rightarrow D$  can be continuously extended as rigid rotations of angles  $\varrho_-$  and  $\varrho_+$  to the boundaries  $C_-$  and  $C_+$ , then  $\varrho_- < \varrho_+$  (due to the twist condition) and there exists a function  $h : \{(s, s_1) \in \mathbb{R}^2 : \varrho_- < s_1 - s < \varrho_+\} \rightarrow \mathbb{R}$ , determined modulo an additive constant, such that

$$G^*\lambda - \lambda = dh.$$

The function  $h$  is called the *Lagrangian* or *generating function* of  $g$ .

Let  $p$  and  $q$  be two relatively prime integers such that  $\varrho_- < p/q < \varrho_+$  and  $q \geq 1$ . A point  $(s, r) \in \mathbb{R} \times I$  is  $(p, q)$ -periodic when  $G^q(s, r) = (s + p, r)$ . The corresponding point  $(s, r) \in \mathbb{T} \times I$  is a periodic point of period  $q$  by  $g$  that is translated  $p$  units in the base by the lift. A  $(p, q)$ -periodic orbit is *Birkhoff* when it is ordered around the cylinder in the same way that the orbits of the rigid rotation of angle  $p/q$ . See [18] for details. The Poincaré-Birkhoff Theorem states that there exist at least two different Birkhoff  $(p, q)$ -periodic orbits [18, 27].

Let  $O = \{(s_k, r_k)\}_{k \in \mathbb{Z}}$  be a  $(p, q)$ -periodic orbit. Its  $(p, q)$ -periodic action is

$$W^{(p,q)}[O] = h(s_0, s_1) + h(s_1, s_2) + \cdots + h(s_{q-1}, s_0 + p).$$

Our goal is to establish an exponentially small bound for the non-negative quantity

$$\Delta^{(p,q)} = \sup_{O, \bar{O} \in \mathcal{O}_g^{(p,q)}} |W^{(p,q)}[\bar{O}] - W^{(p,q)}[O]|,$$

where  $\mathcal{O}_g^{(p,q)}$  denotes the set of all Birkhoff  $(p, q)$ -periodic orbits of the map  $g : D \rightarrow D$ . The difference of  $(p, q)$ -periodic actions can be interpreted as the  $\omega$ -area of certain domains.

Let us explain it.

Let  $O = \{(s_k, r_k)\}_{k \in \mathbb{Z}}$  and  $\bar{O} = \{(\bar{s}_k, \bar{r}_k)\}_{k \in \mathbb{Z}}$  be two  $(p, q)$ -periodic orbits. We can assume, without loss of generality, that  $0 < \bar{s}_0 - s_0 < 1$ . Let  $L_0$  be a curve from  $(s_0, r_0)$  to  $(\bar{s}_0, \bar{r}_0)$  contained in  $\mathbb{R} \times I$ . Set  $L_k = G^k(L_0)$ . The curves  $L_0$  and  $L_q$  have the same endpoints in  $\mathbb{T} \times I$ . Let us assume that these two curves have no topological crossing on the cylinder  $\mathbb{T} \times I$  and let  $B \subset \mathbb{T} \times I$  be the domain enclosed between them.

Observe that  $\int_{L_{k+1}} \lambda - \int_{L_k} \lambda = \int_{L_k} (G^*\lambda - \lambda) = \int_{L_k} dh = h(\bar{s}_k, \bar{s}_{k+1}) - h(s_k, s_{k+1})$ . Hence,  $\sum_{k=0}^{q-1} (h(\bar{s}_k, \bar{s}_{k+1}) - h(s_k, s_{k+1})) = \int_{L_q} \lambda - \int_{L_0} \lambda = \int_{L_q - L_0} \lambda = \pm \int_B \omega$ , where the sign  $\pm$  depends on the orientation of the closed path  $g^q(L_0) - L_0$ , but we do not need it, because we take absolute values in both sides of the previous relation:

$$|W^{(p,q)}[\bar{O}] - W^{(p,q)}[O]| = \int_B \omega =: \text{Area}_\omega[B]. \quad (13)$$

These arguments go back to the *MacKay-Meiss-Percival action principle* [23, 27]. If the curves  $L_0$  and  $g^q(L_0)$  have some topological crossing, then the domain  $B$  has several



connected components, in which case  $|W^{(p,q)}[\bar{O}] - W^{(p,q)}[O]| \leq \int_B \omega =: \text{Area}_\omega[B]$ , because the sign in front of the integral of the area form  $\omega$  depends on the connected component.

If the analytic exact twist map  $g$  has a  $(m, n)$ -resonant RIC, then

$$\Delta^{(m,n)} = 0.$$

Indeed, we can take a segment of the RIC as the curve  $L_0$  used in the previous construction in such a way that  $g^n(L_0) = L_0$  and  $\text{Area}_\omega[B] = 0$ . It turns out that the differences of  $(p, q)$ -periodic actions of  $g$  are exponentially small when  $p/q$  is ‘‘sufficiently close’’ to  $m/n$ . The meaning of ‘‘sufficiently close’’ is clarified in the following theorem. See also Remark 3.

**Theorem 3.** *Let  $g : D \rightarrow D$  be an analytic exact twist map that can be continuously extended as rigid rotations of angles  $\varrho_-$  and  $\varrho_+$  to the boundaries  $C_-$  and  $C_+$ , respectively. Let  $m$  and  $n$  be two relatively prime integers such that  $n \geq 1$ . Let  $a_*, b_* > 0$ ,  $\alpha \in (0, a_*)$ , and  $L \geq 1$ . If  $\varrho_- = m/n$ ,  $g$  can be analytically extended to  $C_-$ , and there exist some analytic coordinates  $(x, y)$  such that  $C_- \equiv \{y = 0\}$  and the power map  $f = g^n$  has the form (7) and can be analytically extended to the complex domain (8), then there exists  $K > 0$  such that*

$$\Delta^{(p,q)} \leq K \exp\left(-\frac{2\pi\alpha q}{|np - mq|}\right), \quad (14)$$

for all relatively prime integers  $p$  and  $q$  with  $1 \leq np - mq \leq L$  and  $q \geq 1$ . The same upper bound holds interchanging the roles of  $C_+$  and  $C_-$ , but in this case  $1 \leq mq - np \leq L$ .

The proof has been placed at Section 6.

*Remark 2.* This theorem requires some global hypotheses because, by definition, the computation of  $\Delta^{(p,q)}$  involves all  $(p, q)$ -periodic orbits of the map  $g : D \rightarrow D$  and not only the ones close to the resonant boundary. The global twist condition and the continuous extension as a rigid rotation to the other boundary imply that all  $(p, q)$ -periodic orbits are close to the resonant boundary. Clearly, there exists a local version of the exponentially small upper bound (14) when the analytic exact twist map  $g$  is defined on a small neighbourhood of the resonant RIC.

*Remark 3.* Condition  $|np - mq| \leq L$  implies that  $|p/q - m/n| = O(1/q)$  as  $q \rightarrow +\infty$ .

*Remark 4.* In many applications,  $\varrho_- = 0$ , so  $C_-$  is a RIC of fixed points. Then Theorem 3 implies that  $\Delta^{(p,q)}$  is exponentially small in the period  $q$  when  $p$  remains uniformly bounded. To be precise, if  $\alpha \in (0, a_*)$  and  $L \geq 1$ , there exists  $K > 0$  such that

$$\Delta^{(p,q)} \leq K e^{-2\pi\alpha q/p}$$

for all relatively prime integers  $p$  and  $q$  with  $q \geq 1$  and  $1 \leq p \leq L$ .

### 3. On the length spectrum of analytic convex domains

#### 3.1. Convex billiards

We recall some well-known results about billiards that can be found in [20, 36, 18].

Let  $\Gamma$  be a closed strictly convex curve in the Euclidean plane  $\mathbb{R}^2$ . We assume, without loss of generality, that  $\Gamma$  has length one. Let  $s \in \mathbb{T}$  be an arc-length parameter on  $\Gamma$ . Let  $f : \mathbb{T} \times (0, \pi) \rightarrow \mathbb{T} \times (0, \pi)$ ,  $(s_1, r_1) = f(s, r)$ , be the map that models the billiard dynamics inside  $\Gamma$  using the *Birkhoff coordinates*  $(s, r)$ , where  $s \in \mathbb{T}$  determines the impact point on the curve, and  $r \in (0, \pi)$  denotes the impact angle.

The map  $f$  preserves the exact symplectic form  $\omega = \sin r \, ds \wedge dr$  and has the intersection property. Indeed,  $f : \mathbb{T} \times (0, \pi) \rightarrow \mathbb{T} \times (0, \pi)$  is an exact twist map with boundary rotation numbers  $\varrho_- = 0$  and  $\varrho_+ = 1$ . Besides, its Lagrangian is the distance between consecutive impact points. Finally,  $f$  is analytic when  $\Gamma$  is analytic.

Any  $(p, q)$ -periodic orbit on the billiard map forms a closed inscribed polygon with  $q$  sides that makes  $p$  turns inside  $\Gamma$ . Since the Lagrangian of the billiard map is the distance between consecutive impact points, the periodic action of a periodic orbit is just the total length of the corresponding polygon. Therefore, the supremum action difference among  $(p, q)$ -periodic billiard orbits is the supremum length difference among inscribed billiard  $(p, q)$ -polygons.

### 3.2. Study close to the boundary of the billiard table

We note that  $r \rightarrow 0^+$  when a billiard trajectory approaches  $\Gamma$ . Thus, in order to study the billiard dynamics close to  $\Gamma$ , we must study the billiard map  $f : \mathbb{T} \times (0, \pi) \rightarrow \mathbb{T} \times (0, \pi)$  in a neighbourhood of the lower boundary  $\mathbb{T} \times \{0\}$ . We want to apply Theorem 3 to this lower boundary, so we need to check that  $f$  can be analytically extended to  $\mathbb{T} \times [0, \pi)$ . We prove a stronger result in the following proposition. To be precise, the billiard map can be considered as a real analytic diffeomorphism of a torus by identifying the upper boundary  $\mathbb{T} \times \{\pi\}$  and the lower boundary  $\mathbb{T} \times \{0\}$  and considering the impact angles on the projective line.

**Proposition 4.** *Let  $\Gamma$  be an analytic strictly convex curve in the Euclidean plane. Let  $\rho(s)$  be the radius of curvature of  $\Gamma$  as a function of the arc-length parameter  $s$ . The billiard map  $f : \mathbb{T} \times (0, \pi) \rightarrow \mathbb{T} \times (0, \pi)$  associated to  $\Gamma$  satisfies the following properties:*

(i) *It extends analytically to  $\mathbb{T} \times \mathbb{R}$ ;*

(ii) *Its analytic extension has the reversibilities and the periodicity*

$$f = R \circ f^{-1} \circ R, \quad f = T \circ f^{-1} \circ T, \quad P \circ f = f \circ P,$$

where  $R(s, r) = (s, -r)$ ,  $T(s, r) = (s, \pi - r)$ , and  $P(s, r) = (s, r + \pi)$ ;

(iii) *It can be considered as a real analytic diffeomorphism of the torus  $\mathbb{T} \times (\mathbb{R}/\pi\mathbb{Z})$ ; and*

(iv) *The first terms of its Taylor expansion for small impact angles are*

$$\begin{cases} s_1 = s + 2\rho(s)r + 4\rho(s)\rho'(s)r^2/3 + O(r^3), \\ r_1 = r - 2\rho'(s)r^2/3 + (4(\rho'(s))^2/9 - 2\rho(s)\rho''(s)/3)r^3 + O(r^4). \end{cases}$$

*Proof.* Let  $m : \mathbb{T} \rightarrow \Gamma$ ,  $\kappa : \mathbb{T} \rightarrow (0, +\infty)$ , and  $n : \mathbb{T} \rightarrow \mathbb{R}^2$  be an arc-length parametrization, the curvature, and a unit normal vector of the analytic strictly convex curve linked by the

relation  $m''(s) = \kappa(s)n(s)$ . In what follows, we consider these functions as real analytic 1-periodic functions defined on the universal cover  $\mathbb{R}$  of  $\mathbb{T}$ . We also consider the set

$$U = \{(s, s_1) \in \mathbb{R}^2 : |s_1 - s| < 1\} = U_- \cup \Delta \cup U_+,$$

where  $U_- = \{(s, s_1) \in \mathbb{R}^2 : s - 1 < s_1 < s\}$ ,  $U_+ = \{(s, s_1) \in \mathbb{R}^2 : s < s_1 < s + 1\}$ , and  $\Delta$  is the diagonal  $\{(s, s_1) \in \mathbb{R}^2 : s_1 = s\}$ . Next, we write both impact angles as functions of consecutive impact points. To be precise, there exist two analytic functions  $\bar{r}, \bar{r}_1 : U_+ \rightarrow (0, \pi)$  such that

$$(s, s_1) \in U_+ \text{ and } f(s, r) = (s_1, r_1) \iff r = \bar{r}(s, s_1) \text{ and } r_1 = \bar{r}_1(s, s_1).$$

Let us study what happens when  $s_1 \rightarrow s^+$ . We note that  $r = \bar{r}(s, s_1)$  is the angle between  $m'(s)$  and  $m(s_1) - m(s)$ . Hence,

$$\tan \bar{r}(s, s_1) = \frac{\det(m'(s), m(s_1) - m(s))}{\langle m'(s), m(s_1) - m(s) \rangle} = \frac{\int_0^1 \det(m'(s), m'(s + t(s_1 - s))) dt}{\int_0^1 \langle m'(s), m'(s + t(s_1 - s)) \rangle dt}, \quad (15)$$

so the function  $\bar{r}(s, s_1)$  can be analytically extended to the diagonal  $\Delta$ ,  $\bar{r}(s, s) \equiv 0$ , and

$$\partial_2 \bar{r}(s, s) = \lim_{s_1 \rightarrow s^+} \partial_2 \bar{r}(s, s_1) = \lim_{s_1 \rightarrow s^+} \cos^2 \bar{r}(s, s_1) \partial_2 \tan \bar{r}(s, s_1) = \kappa(s)/2,$$

for all  $s \in \mathbb{R}$ . The function  $\bar{r}_1(s, s_1)$  can also be extended analytically to  $\Delta$  by using an analogous argument. Since  $\partial_2 \bar{r}(s, s) = \kappa(s)/2 > 0$  for all  $s \in \mathbb{R}$ , there exists a function  $\bar{s}_1(s, r)$  analytic on  $\mathbb{R} \times \{0\}$  such that  $\bar{r}(s, \bar{s}_1(s, r)) = r$  and  $\bar{s}_1(s, 0) = s$ . Thus, we can write the billiard map as  $f(s, r) = (s_1, r_1) = (\bar{s}_1(s, r), \bar{r}_1(s, \bar{s}_1(s, r)))$ , which proves that  $f$  can be analytically extended to a complex neighbourhood of  $\mathbb{R} \times \{0\}$ .

Let us describe the meaning of this extension for small negative impact angles. To begin with, we model the billiard dynamics as a map that acts on consecutive impact points. That is, we consider the new billiard map  $\tilde{f} : U_+ \rightarrow U_+$  such that  $\tilde{f}(s_{-1}, s) = (s, s_1)$ . The previous arguments show that  $\tilde{f}$  can be analytically extended to a complex neighbourhood of  $\Delta$ . It is geometrically clear that relation  $\tilde{f}(s_1, s) = (s, s_{-1})$  holds for this extension. That is,  $\tilde{f} = \tilde{R} \circ \tilde{f}^{-1} \circ \tilde{R}$ , where  $\tilde{R} : U \rightarrow U$  is the reversor  $\tilde{R}(s, s_1) = (s_1, s)$ .

Next, we come back to the Birkhoff coordinates  $(s, r)$ . Let  $R(s, r) = (s, -r)$ . Since  $\bar{r}_1(s_1, s) = -\bar{r}(s, s_1)$ , the reversor  $\tilde{R}$  becomes  $\hat{R} = R \circ f^{-1}$  in the Birkhoff coordinates. Hence, relations  $f = \hat{R} \circ f^{-1} \circ \hat{R}$  and  $f = R \circ f^{-1} \circ R$  hold, at least in a neighbourhood of  $\mathbb{R} \times \{0\}$ . Thus, we can analytically extend  $f$  to  $\mathbb{R} \times (-\pi, \pi)$  by using the reversor  $R$ . We note that  $P = T \circ R$ , where  $T(s, r) = (s, \pi - r)$  is a well-known reversor of  $f$ . This implies that  $P \circ f = f \circ P$ , so we can analytically extend  $f$  to the whole space  $\mathbb{R} \times \mathbb{R}$  by periodicity. Finally, we recall that  $s$  is an angular variable, so the billiard map  $f$  is analytic in the infinite cylinder  $\mathbb{T} \times \mathbb{R}$ .

This ends the proof of the first two items. The third item is a trivial consequence of them. The Taylor expansion given in the last item was obtained by Lazutkin [21].  $\square$

From the Taylor expansion given in the previous proposition, Lazutkin deduced that the billiard map takes the form

$$x_1 = x + y + O(y^3), \quad y_1 = y + O(y^4)$$

in the analytic Lazutkin coordinates  $(x, y)$  defined by

$$x = k \int_0^s \rho^{-2/3}(t) dt, \quad y = 4k\rho^{1/3}(s) \sin(r/2), \quad k^{-1} = \int_0^1 \rho^{-2/3}(t) dt.$$

The constant  $k$  is determined in such a way that the new angular coordinate  $x$  is defined modulus one. Lazutkin's results are more refined than the ones in Lemma 1, he wrote the billiard map as a smaller perturbation of the integrable twist map  $(x_1, y_1) = (x + y, y)$ , but we do not need it. Our coordinates (9)–(10) and the Lazutkin coordinates are directly related. Namely,  $\varphi(s) = 2\rho(s)$  and  $\psi(s) = -2\rho'(s)/3$  imply that  $a(s) = \frac{\nu}{2} \int_0^s \rho^{-2/3}(t) dt$ ,  $b(s) = \nu\rho^{1/3}(s)$ , and  $\nu = 2k$ .

Let  $D = \mathbb{T} \times (0, \pi)$ ,  $C_- = \mathbb{T} \times \{0\}$ , and  $C_+ = \mathbb{T} \times \{\pi\}$ . Set  $m = 0$  and  $n = 1$ . The billiard map  $f : D \rightarrow D$  satisfies the hypotheses required in Theorem 3. Only one hypothesis can rise doubts at this point. Namely, the existence of some analytic coordinates  $(x, y)$  such that  $C_- \equiv \{y = 0\}$  and  $f$  has the form (7) and can be analytically extended to the complex domain (8) for some  $a_*, b_* > 0$ . This existence follows directly from Lemma 1 since  $f$  has the form (6) with  $\varphi(s) = 2\rho(s) > 0$  and  $\psi(s) = -2\rho'(s)/3$ . See Proposition 4. Therefore, we get the exponentially small upper bound of the quantities  $\mathcal{L}^{(p,q)}$  defined in the introduction.

**Theorem 5.** *Let  $\Gamma$  be an analytic strictly convex curve in the Euclidean plane. Let  $a_* > 0$  be the analyticity strip width of the lower boundary  $C_-$ . Let  $\alpha \in (0, a_*)$  and  $L \geq 1$ . There exists a constant  $K > 0$  such that*

$$\mathcal{L}^{(p,q)} \leq K e^{-2\pi\alpha q/p},$$

for all relatively prime integers  $p$  and  $q$  with  $q \geq 2$  and  $0 < p \leq L$ .

The same exponentially small upper bound holds for analytic geodesically strictly convex curves on surfaces of constant curvature, where the billiard trajectories are just broken geodesics. Billiard maps on the Klein model of the hyperbolic plane  $\mathbb{H}^2$  and on the positive hemisphere  $\mathbb{S}_+^2$  have been studied, for instance, in [7], where it is shown that they are exact twist maps with the same boundary rotation numbers as in the Euclidean case. Therefore, by local isometry arguments, we can write a version of Theorem 5 on any surface of constant curvature.

### 3.3. Billiard tables of constant width

**Definition 4.** *A smooth closed convex curve is of constant width if and only if it has a chord in any direction perpendicular to the curve at both ends.*

Billiards inside convex curves of constant width have a nice property [19, 14]. Let us explain it.

The billiard map associated to a smooth convex curve of constant width has the horizontal line  $\mathbb{T} \times \{\pi/2\}$  as a  $(1, 2)$ -resonant RIC. Any trajectory belonging to that RIC is orthogonal to the curve at its two endpoints. Due to the variational formulation, all the  $(1, 2)$ -periodic orbits are extrema of the  $(1, 2)$ -periodic action and, thus, all  $(1, 2)$ -periodic trajectories have the same length, which is the reason we refer to them as constant width curves.

Another characterization of constant width curves is the following. We reparametrize the curve by using the angle  $\varphi \in (0, 2\pi)$  between the tangent vector at a point in the curve and some fixed line. Let  $\rho(\varphi)$  be the radius of curvature at this point. The curve has constant width if and only if the Fourier series of  $\rho(\varphi)$  contains no other even coefficients than the constant term. Thus, the space of analytic constant width curves has infinite dimension and codimension.

Next, we apply Theorem 3 twice in order to get the exponentially small upper bound of  $\mathcal{L}^{(p,q)}$  in both sides of the  $(1, 2)$ -resonant RIC  $\mathbb{T} \times \{\pi/2\}$ . Namely, we apply it to the cylinders  $\mathbb{T} \times (\pi/2, \pi)$  and  $\mathbb{T} \times (0, \pi/2)$ . Since  $\mathbb{T} \times \{\pi/2\}$  is a  $(1, 2)$ -resonant RIC of the billiard map  $g : \mathbb{T} \times (0, \pi) \rightarrow \mathbb{T} \times (0, \pi)$ , the square billiard map  $f = g^2$  has the form

$$s_1 = s + \varphi(s)v + O(v^2), \quad v_1 = v + \psi(s)v^2 + O(v^3),$$

for some real analytic 1-periodic functions  $\varphi(s)$  and  $\psi(s)$ , where  $v = r - \pi/2$  measures the distance to the resonant RIC. The function  $\varphi(s)$  is positive because  $g$  satisfies the twist condition on the whole phase space  $\mathbb{T} \times (0, \pi)$ . Therefore, we can apply Lemma 1, so all hypotheses of Theorem 3 are satisfied.

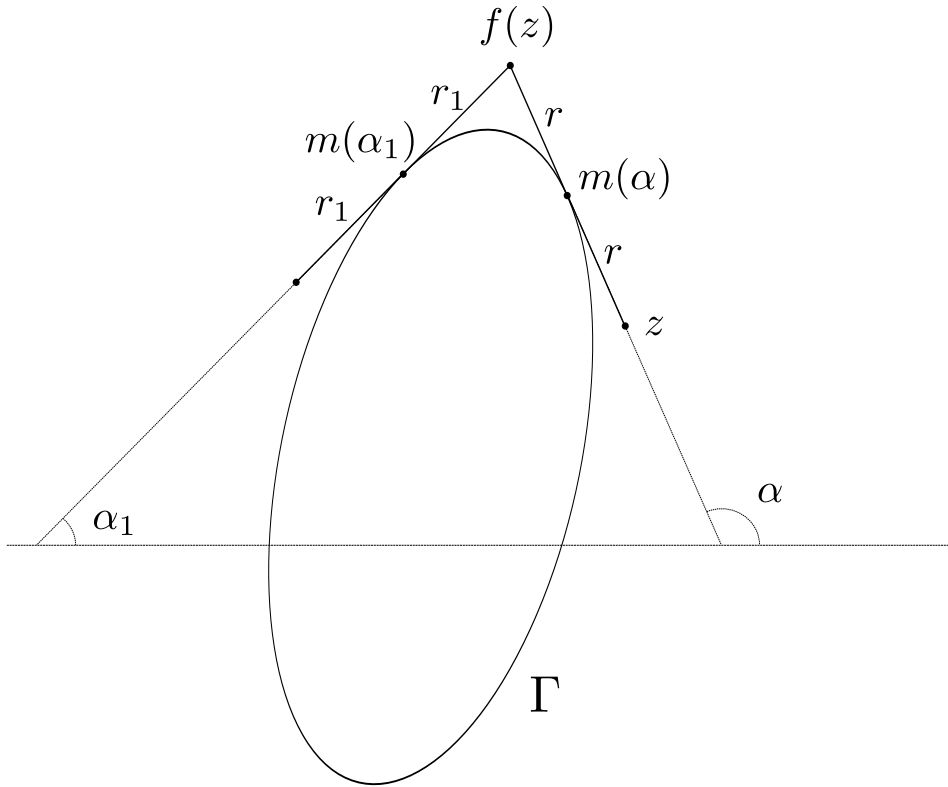
**Theorem 6.** *Let  $g : \mathbb{T} \times (0, \pi) \rightarrow \mathbb{T} \times (0, \pi)$  be the billiard map of an analytic strictly convex curve of constant width. Let  $a_* > 0$  be the analyticity strip width of the  $(1, 2)$ -resonant RIC of  $g$ . Let  $\alpha \in (0, a_*)$  and  $L \geq 1$ . There exists a constant  $K > 0$  such that*

$$\mathcal{L}^{(p,q)} \leq K \exp\left(-\frac{\pi\alpha}{|p/q - 1/2|}\right),$$

for all relatively prime integers  $p$  and  $q$  such that  $1 \leq |2p - q| \leq L$  and  $q \geq 3$ .

One could try to generalize constant width billiards, where  $\mathbb{T} \times \{\pi/2\}$  is a  $(1, 2)$ -resonant RIC, to *constant angle tables*, where  $\mathbb{T} \times \{r_0\}$  is assumed to be a  $(m, n)$ -resonant RIC. However, the only table such that  $\mathbb{T} \times \{r_0\}$  is a  $(m, n)$ -resonant RIC, with  $(m, n) \neq (1, 2)$ , is the circle. See [14, 8]. By the way, Theorem 3 applies to this case but, since the circular billiard is integrable,  $\mathcal{L}^{(m,n)} \equiv 0$ , for all  $(m, n)$ . In fact, the circular billiard map is globally conjugated to the integrable twist map (11).

There are more billiard tables with resonant RICs, but their RICs are not horizontal. For instance, the elliptic table has all possible  $(m, n)$ -resonant RICs, but the  $(1, 2)$ -resonant one. Hence, in this case,  $\mathcal{L}^{(m,n)} = 0$ . Baryshnikov and Zharnitsky [3] proved that an ellipse can be infinitesimally perturbed so that any chosen resonant RIC will persist. Innami [16] found a condition on the billiard table that guarantees the existence of a  $(1, 3)$ -resonant RIC. However, Theorem 3 can not be applied in such cases, because both the Baryshnikov-Zharnitsky and the Innami constructions are done in the smooth category, where we can only claim that  $\mathcal{L}^{(p,q)}$  is beyond all order in the difference between rotation numbers.



**Figure 1.** The envelope coordinates  $(\alpha, r)$  and the dual billiard map  $f : \mathcal{U} \rightarrow \mathcal{U}$ .

## 4. On the area spectrum of analytic convex domains

### 4.1. Dual billiards

We recall some well-known facts about dual billiards that can be found in [5, 15, 36].

Let  $\Gamma$  be a strictly convex closed curve in the Euclidean plane  $\mathbb{R}^2$ . Let  $\mathcal{U}$  be the unbounded component of  $\mathbb{R}^2 \setminus \Gamma$ . The dual billiard map  $f : \mathcal{U} \rightarrow \mathcal{U}$  is defined as follows:  $f(z)$  is the reflection of  $z$  in the tangency point of the oriented tangent line to  $\Gamma$  through  $z$ . This map is area-preserving. Next, we introduce the *envelope coordinates*  $(\alpha, r) \in \mathbb{T}_* \times (0, +\infty)$  of a point  $z \in \mathcal{U}$ . In this section,  $\mathbb{T}_* = \mathbb{R}/2\pi\mathbb{Z}$ . We recall that  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ .

Given a point  $z \in \mathcal{U}$ , let  $\alpha \in \mathbb{T}_*$  be the angle made by the positive tangent line to  $\Gamma$  in the direction of  $z$  with a fixed direction of the plane, and let  $r \in I$  be the distance along this line from  $\Gamma$  to  $z$ . See Figure 1.

The dual billiard map preserves the exact symplectic form  $\omega = r d\alpha \wedge dr$  and has the intersection property. Indeed,  $f : \mathbb{T}_* \times (0, +\infty) \rightarrow \mathbb{T}_* \times (0, +\infty)$ ,  $(\alpha_1, r_1) = f(\alpha, r)$ , is an exact twist map with boundary rotation numbers  $\varrho_- = 0$  and  $\varrho_+ = \pi$ . Its Lagrangian is the area enclosed by  $\Gamma$  and the tangent lines through the points on  $\Gamma$  with coordinates  $\alpha$  and  $\alpha_1$ .

Any  $(p, q)$ -periodic orbit on the dual billiard map forms a closed circumscribed polygon with  $q$  sides that makes  $p$  turns outside  $\Gamma$ . Since the Lagrangian of the dual billiard map is the above-mentioned area, the periodic action of a periodic orbit is just the area enclosed between the corresponding polygon and  $\Gamma$ , taking into account some multiplicities when

$p \geq 2$ . Therefore, the supremum action difference among  $(p, q)$ -periodic dual billiard orbits is the supremum area difference among circumscribed dual billiard  $(p, q)$ -polygons.

#### 4.2. Study close to the curve

We note that  $r \rightarrow 0^+$  when the point  $z \in \mathcal{U}$  approaches to the curve  $\Gamma$ . Therefore, in order to study the dual billiard dynamics close to  $\Gamma$ , we must study the dual billiard map  $f$  in a neighbourhood of the lower boundary  $C_- = \mathbb{T}_* \times \{0\}$  of  $\mathbb{T}_* \times (0, +\infty)$ .

Let us check that the dual billiard map  $f : \mathbb{T}_* \times (0, +\infty) \rightarrow \mathbb{T}_* \times (0, +\infty)$  satisfies the hypotheses required in Theorem 3. Only two of these hypotheses remain to be checked. The first one is that  $f$  can be analytically extended to the lower boundary  $C_- = \mathbb{T}_* \times \{0\}$ . The second one is that there exist some analytic coordinates  $(x, y)$  such that  $C_- \equiv \{y = 0\}$  and  $f$  has the form (7) and can be analytically extended to the complex domain (8) for some  $a_*, b_* > 0$ . This second part will follow directly from Lemma 1 once we check that

$$\alpha_1 = \alpha + 2\kappa(\alpha)r + O(r^2), \quad r_1 = r - 2\kappa'(\alpha)r^2/3 + O(r^3), \quad (16)$$

where  $\kappa(\alpha)$  is the curvature of  $\Gamma$  as a function of the envelope parameter  $\alpha$ . We prove these results (and a stronger extension one) in the following proposition.

**Proposition 7.** *Let  $\Gamma$  be an analytic strictly convex curve in the Euclidean plane. The dual billiard map  $f : \mathbb{T}_* \times (0, +\infty) \rightarrow \mathbb{T}_* \times (0, +\infty)$  associated to  $\Gamma$  satisfies the following properties:*

- (i) *It extends analytically to  $\mathbb{T}_* \times \mathbb{R}$ ;*
- (ii) *Its analytic extension has the reversibility  $f = R \circ f^{-1} \circ R$ , with  $R(\alpha, r) = (\alpha, -r)$ ; and*
- (iii) *It has the form (16) for small distances  $r$ .*

*Proof.* Following the proof of Proposition 4, we consider the coordinate  $\alpha$  in the universal cover  $\mathbb{R}$ , we write both distances as functions of consecutive tangent points:  $r = \bar{r}(\alpha, \alpha_1)$  and  $r_1 = \bar{r}_1(\alpha, \alpha_1)$ , and then we repeat the steps of the previous proof. There are only three remarkable differences between the two proofs.

First,  $U_+ = \{(\alpha, \alpha_1) \in \mathbb{R}^2 : \alpha < \alpha_1 < \alpha + \pi\}$  is the open set where we directly know that the functions  $\bar{r}(\alpha, \alpha_1)$  and  $\bar{r}_1(\alpha, \alpha_1)$  are analytic. We note that tangent lines through points with coordinates  $\alpha$  and  $\alpha + \pi$  are parallel, so  $\lim_{\alpha_1 \rightarrow (\alpha + \pi)^-} \bar{r}(\alpha, \alpha_1) = +\infty$ .

Second, to study what happens when  $\alpha_1 \rightarrow \alpha^+$ , let  $\rho(\alpha)$  be the radius of curvature of  $\Gamma$  in the envelope parameter  $\alpha$  and set  $\alpha_1 = \alpha + \delta$ . From Boyland [5], we know that

$$\bar{r}(\alpha, \alpha_1) = \frac{\int_{\alpha}^{\alpha_1} \sin(v - \alpha)\rho(v)dv}{\sin(\alpha_1 - \alpha)} = \frac{\delta}{\sin \delta} \int_0^1 \sin(\delta t)\rho(\alpha + \delta t)dt,$$

so the function  $\bar{r}(\alpha, \alpha_1)$  can be analytically extended to  $\Delta = \{(\alpha, \alpha_1) \in \mathbb{R}^2 : \alpha_1 = \alpha\}$ ,  $\bar{r}(\alpha, \alpha) = 0$ , and  $\partial_2 \bar{r}(\alpha, \alpha) = \rho(\alpha)/2 > 0$  for all  $\alpha \in \mathbb{R}$ .

Third, the dual billiard map has no periodicity in the coordinate  $r$ , which is geometrically obvious since  $r$  is no longer an angle, but a distance, in dual billiards.

The rest of the proof of the analyticity and reversibility is completely analogous. We omit the details.

Finally, the Taylor expansion (16) of the dual billiard map around  $r = 0$  was obtained by Tabachnikov in [37]. To be precise, Tabachnikov wrote this Taylor expansion in coordinates  $(s, r)$ , where  $s$  is an arc-length parameter, but his result can be easily adapted.  $\square$

From the Taylor expansion given in the previous proposition, Tabachnikov deduced that the billiard map takes the form (7) in the analytic *Tabachnikov coordinates*  $(x, y)$  defined by

$$x = k \int_0^\alpha \kappa^{-2/3}(v)dv, \quad y = 2k\kappa^{1/3}(\alpha)r, \quad k^{-1} = \int_0^{2\pi} \kappa^{-2/3}(v)dv.$$

The constant  $k$  has been determined in such a way that  $x$  is defined modulus one:  $x \in \mathbb{T}$ . The coordinates (9)–(10) coincide with the Tabachnikov coordinates. Concretely,  $\varphi(\alpha) = 2\kappa(\alpha)$  and  $\psi(\alpha) = -2\kappa'(\alpha)/3$  imply that  $a(\alpha) = \frac{\nu}{2} \int_0^\alpha \kappa^{-2/3}(t)dt$ ,  $b(\alpha) = \nu\rho^{1/3}(\alpha)$ , and  $\nu = 2k$ .

We get the following exponentially small upper bound of the quantities  $\mathcal{A}^{(p,q)}$  defined in the introduction by direct application of Proposition 7, Lemma 1, and Theorem 3.

**Theorem 8.** *Let  $\Gamma$  be an analytic strictly convex curve in the Euclidean plane. Let  $a_* > 0$  be the analyticity strip width of the lower boundary  $C_-$ . Let  $\alpha \in (0, a_*)$  and  $L \geq 1$ . There exists a constant  $K > 0$  such that*

$$\mathcal{A}^{(p,q)} \leq K e^{-2\pi\alpha q/p},$$

for all relatively prime integers  $p$  and  $q$  with  $q \geq 3$  and  $0 < p \leq L$ .

Tabachnikov [38] studied the dual billiard map in the hyperbolic plane  $\mathbb{H}^2$ , and extended the asymptotic expansion (3) to that new setting. He also claimed that there exists an analogous formula for dual billiards on the unit sphere  $\mathbb{S}^2$ . Therefore, by local isometry arguments, we can write a version of Theorem 8 on any surface of constant curvature.

#### 4.3. Study far away from the curve

We note that  $r \rightarrow +\infty$  when the point  $z \in \mathcal{U}$  moves away from the curve  $\Gamma$ . We use the coordinates  $(\alpha, v)$  to work at infinity, where  $v = 1/r$  and  $(\alpha, r) \in \mathbb{T}_* \times (0, +\infty)$  are the coordinates introduced in Subsection 4.1. Tabachnikov [37] realized that the dual billiard map at infinity can be seen as a map defined in a neighbourhood of the  $(1, 2)$ -resonant RIC  $\{v = 0\}$ . To be precise, he saw that the dual billiard map can be analytically extended to  $v \geq 0$ , and its square has the form

$$\alpha_1 = \alpha + \varphi(\alpha)v + O(v^2), \quad v_1 = v + \psi(\alpha)v^2 + O(v^3),$$

for some real analytic 1-periodic functions  $\varphi(\alpha)$  and  $\psi(\alpha)$ , and  $\varphi(\alpha)$  is negative. This is all we need to state the following theorem.



**Theorem 9.** *Let  $\Gamma$  be an analytic strictly convex curve in the Euclidean plane. Let  $a_* > 0$  be the analyticity strip width of the  $(1, 2)$ -resonant RIC  $\{v = 0\}$ . Let  $\alpha \in (0, a_*)$  and  $L \geq 1$ . There exists a constant  $K > 0$  such that*

$$\mathcal{A}^{(p,q)} \leq K \exp\left(-\frac{\pi\alpha}{|p/q - 1/2|}\right),$$

for all relatively prime integers  $p$  and  $q$  such that  $1 \leq |2p - q| \leq L$  and  $q \geq 3$ .

## 5. Proof of Theorem 2

### 5.1. Spaces, norms, and projections

Let  $\mathcal{X}_{a,b}$ , with  $a > 0$  and  $b > 0$ , be the space of all analytic functions  $g$  defined on the open set

$$D_{a,b} = \{(x, y) \in (\mathbb{C}/\mathbb{Z}) \times \mathbb{C} : |\Im x| < a, |y| < b\}$$

with bounded Fourier norm

$$\|g\|_{a,b} = \sum_{k \in \mathbb{Z}} |\hat{g}_k|_b e^{2\pi|k|a},$$

where  $\hat{g}_k(y)$  denotes the  $k$ -th Fourier coefficient of the 1-periodic function  $g(\cdot, y)$  and

$$|\hat{g}_k|_b = \sup \{|\hat{g}_k(y)| : y \in B_b\}$$

denotes its sup-norm on the complex open ball  $B_b = \{y \in \mathbb{C} : |y| < b\}$ . Let

$$|g|_{a,b} = \sup \{|g(x, y)| : (x, y) \in D_{a,b}\}$$

be the sup-norm on  $D_{a,b}$ .

Let  $\mathcal{X}_{a,b,m}$  be the space of all vectorial functions  $G : D_{a,b} \rightarrow \mathbb{C}^2$  of the form

$$G(x, y) = (y^m g_1(x, y), y^{m+1} g_2(x, y))$$

such that  $g_1, g_2 \in \mathcal{X}_{a,b}$ . The space  $\mathcal{X}_{a,b,m}$  is a Banach space with the Fourier norm  $\|G\|_{a,b,m} = \max \{\|g_1\|_{a,b}, \|g_2\|_{a,b}\}$ . The sup-norm  $|\cdot|_{a,b,m}$  on  $\mathcal{X}_{a,b,m}$  is defined analogously.

Let  $g_2^*(y) = \int_0^1 g_2(x, y) dx$  be the average of  $g_2(x, y)$ . Let  $\mathcal{X}_{a,b,m} = \mathcal{X}_{a,b,m}^* \oplus \mathcal{X}_{a,b,m}^\bullet$  be the direct decomposition where  $\mathcal{X}_{a,b,m}^*$  is the vectorial subspace of the elements of the form  $G^*(x, y) = (0, y^{m+1} g_2^*(y))$ , whereas  $\mathcal{X}_{a,b,m}^\bullet$  is the one of the elements with  $g_2^*(y) = 0$ . Let  $\pi^* : \mathcal{X}_{a,b,m} \rightarrow \mathcal{X}_{a,b,m}^*$  and  $\pi^\bullet : \mathcal{X}_{a,b,m} \rightarrow \mathcal{X}_{a,b,m}^\bullet$  be the associated projections. Thus, any  $G \in \mathcal{X}_{a,b,m}$  can be decomposed as  $G = G^* + G^\bullet$ , where

$$G^* = \pi^*(G) = (0, y^{m+1} g_2^*(y)) \in \mathcal{X}_{a,b,m}^*, \quad G^\bullet = \pi^\bullet(G) \in \mathcal{X}_{a,b,m}^\bullet.$$

Obviously,  $\|G^*\|_{a,b,m} \leq \|G\|_{a,b,m}$  and  $\|G^\bullet\|_{a,b,m} \leq \|G\|_{a,b,m}$

We will always denote the scalar functions in  $\mathcal{X}_{a,b}$  with lower-case letters, and the vectorial functions in  $\mathcal{X}_{a,b,m}$  with upper-case letters. Asterisk and bullet superscripts in upper-case letters stand for the  $\pi^*$ -projections and  $\pi^\bullet$ -projections of vectorial functions in  $\mathcal{X}_{a,b,m}$ , respectively. Asterisk superscripts in lower-case letters denote averages of scalar functions in  $\mathcal{X}_{a,b}$ . We will always write the couple of scalar functions associated to any given vectorial

function of  $\mathcal{X}_{a,b,m}$  with the corresponding lower-case letter and the subscripts  $j = 1, 2$ . Hat symbols denote Fourier coefficients.

### 5.2. The averaging and the iterative lemmas

Henceforth, let  $A(x, y) = (x + y, y)$  be the integrable twist map introduced in (11). Let  $F = F_2$  be a map satisfying the properties listed in Lemma 1, so  $F_2 = A + G_2$  for some  $G_2 \in \mathcal{X}_{a_2, b_2, 2}$ , where  $a_2 = a_*$  is the analyticity strip width in the angular variable  $x$ , and  $b_2 = b_*$  is the analyticity radius in  $y$ . Hence,  $F_2$  is a perturbation of  $A$  of order two. The following lemma allows us to increase that order as much as we want by simply losing as little analyticity strip width as we want. It is based on classical averaging methods. In particular, we see that  $F$  is a perturbation beyond all order of  $A$ .

**Lemma 10** (Averaging Lemma). *Let  $F_2 = A + G_2$ , with  $G_2 \in \mathcal{X}_{a_2, b_2, 2}$  and  $a_2 > 0$  and  $b_2 > 0$ , be a real analytic map with the intersection property on the cylinder  $\mathbb{T} \times (-b_2, b_2)$ . Let  $m \geq 3$  be an integer. Let  $a_m$  be any analyticity strip width such that  $a_m \in (0, a_2)$ .*

*There exist an analyticity radius  $b_m \in (0, b_2)$  and a change of variables of the form  $\Phi_m = I + \Psi_m$  for some  $\Psi_m \in \mathcal{X}_{a_m, b_m, 1}$  such that the transformed map  $F_m = \Phi_m^{-1} \circ F_2 \circ \Phi_m$  is real analytic, has the intersection property on the cylinder  $\mathbb{T} \times (-b_m, b_m)$ , and has the form  $F_m = A + G_m$  for some  $G_m \in \mathcal{X}_{a_m, b_m, m}$ .*

*Besides, the change of variables  $\Phi_m$  is close to the identity on  $\mathbb{T} \times (-b_m, b_m)$ . That is,*

$$\Phi_m(x, y) = (x + O(y), y + O(y^2)), \quad \det[\Phi_m(x, y)] = 1 + O(y), \quad (17)$$

*uniformly for all  $(x, y) \in \mathbb{T} \times (-b_m, b_m)$ .*

*Proof.* The change  $\Phi_m$  is the composition of  $m - 2$  changes of the form

$$\tilde{\Phi}_l = I + \tilde{\Psi}_l, \quad \tilde{\Psi}_l \in \mathcal{X}_{a_l, b_l, l-1}, \quad 2 \leq l < m,$$

where  $a_l = a_2 - (l - 2)\epsilon$ ,  $\epsilon = (a_2 - a_m)/(m - 2)$ ,  $(b_l)_{2 \leq l < m}$  is a positive decreasing sequence, and  $\tilde{\Psi}_l$  is constructed as follows to increase the order of the perturbation from  $l$  to  $l + 1$ .

Let us suppose that we have a real analytic map with the intersection property on  $\mathbb{T} \times (-b_l, b_l)$  of the form  $F_l = A + G_l$ , for some  $G_l \in \mathcal{X}_{a_l, b_l, l}$  with  $a_l, b_l > 0$  and  $l \geq 2$ .

We begin with a formal computation. We write

$$F_l(x, y) = (x + y + y^l h_1(x) + O(y^{l+1}), y + y^{l+1} h_2(x) + O(y^{l+2})),$$

where the functions  $h_1(x)$  and  $h_2(x)$  are 1-periodic and analytic on the open complex strip  $\{x \in \mathbb{C}/\mathbb{Z} : |\Im x| < a_l\}$ . We will see, by using an *a posteriori* reasoning, that the intersection property implies that  $h_2(x)$  has zero average; that is,  $h_2^* = \int_0^1 h_2(x) dx = 0$ . Nevertheless, we can not prove it yet. Thus, we will keep an eye on  $h_2^*$  in what follows.

If we take the change of variables

$$\tilde{\Phi}_l(x, y) = (x + y^{l-1} \psi_1(x), y + y^l \psi_2(x)) \quad (18)$$

for some functions  $\psi_1(x)$  and  $\psi_2(x)$ , then, after a straightforward computation, the map  $F_{l+1} = (\tilde{\Phi}_l)^{-1} \circ F_l \circ \tilde{\Phi}_l$  has the form

$$F_{l+1}(x, y) = (x + y + y^l k_1(x) + O(y^{l+1}), y + y^{l+1} k_2(x) + O(y^{l+2})),$$

with  $k_1 = \psi_2 + h_1 - \psi'_1$  and  $k_2 = h_2 - \psi'_2$ . Therefore, we take

$$\psi_2(x) = \int_0^x (h_2(s) - h_2^*) ds - h_1^*, \quad \psi_1(x) = \int_0^x (\psi_2(s) + h_1(s)) ds,$$

so that  $k_1(x) = 0$  and  $k_2(x) = h_2^*$ . These functions  $\psi_2(x)$  and  $\psi_1(x)$  are 1-periodic, because  $h_2(x) - h_2^*$  and  $\psi_2(x) + h_1(x)$  have zero average. Besides,  $\psi_1(x)$  and  $\psi_2(x)$  are analytic in the open complex strip  $\{x \in \mathbb{C}/\mathbb{Z} : |\Im x| < a_l\}$ . Indeed,  $\tilde{\Phi}_l = I + \tilde{\Psi}_l$  with  $\tilde{\Psi}_l \in \mathcal{X}_{a_l, b_l, l-1}$ .

Next, we control the domain of definition of the map  $F_{l+1}$ . The inverse change is

$$(\tilde{\Phi}_l)^{-1}(x, y) = (x - y^{l-1} \psi_1(x) + O(y^l), y - y^l \psi_2(x) + O(y^{l+1})).$$

Thus, the maps  $\tilde{\Phi}_l$ ,  $F_l$ , and  $(\tilde{\Phi}_l)^{-1}$  have the form  $(x, y) \mapsto (x + O(y), y + O(y^2))$ , since  $l \geq 2$ . Consequently, if  $b_{l+1} \leq b_l/2$  is small enough, then

$$D_{a_{l+1}, b_{l+1}} \xrightarrow{\tilde{\Phi}_l} D_{a_l - 2\epsilon/3, 4b_{l+1}/3} \xrightarrow{F_l} D_{a_l - \epsilon/3, 5b_{l+1}/3} \xrightarrow{(\tilde{\Phi}_l)^{-1}} D_{a_l, 2b_{l+1}} \subset D_{a_l, b_l},$$

so  $F_{l+1} = (\tilde{\Phi}_l)^{-1} \circ F_l \circ \tilde{\Phi}_l$  is well-defined on  $D_{a_{l+1}, b_{l+1}}$ . Now, let us check that  $h_2^* = 0$ . At this moment, we only know that

$$F_{l+1}(x, y) = (x + y + O(y^{l+1}), y + y^{l+1} h_2^* + O(y^{l+2})),$$

since the change of variables has not eliminated the average  $h_2^*$ . We recall that  $F_{l+1}$  has the intersection property on the cylinder  $\mathbb{T} \times (-b_{l+1}, b_{l+1})$ . This implies that  $h_2^* = 0$ . On the contrary, the image of the loop  $\mathbb{T} \times \{y_0\}$  does not intersect itself when  $0 < y_0 \ll 1$ .

Finally, properties (17) follow directly from the fact that we have performed a finite number of changes, all of them satisfying these same properties.  $\square$

Next, the following theorem provides the exponentially small bound for the  $\pi^\bullet$ -projection of the residue provided an initial order  $m$  big enough. It is the main tool to prove Theorem 2.

**Theorem 11.** *Let  $m \geq 6$  be an integer,  $\bar{a} > 0$ ,  $\bar{d} > 0$ , and  $r \in (0, 1)$ . There exist constants  $\bar{b} = \bar{b}(m, \bar{a}, \bar{d}, r) > 0$  and  $c_j = c_j(r) > 0$ ,  $j = 1, 2, 3$ , such that, if*

$$\bar{F} = A + \bar{G}, \quad \bar{G} \in \mathcal{X}_{\bar{a}, \bar{b}, m}, \quad \bar{d}^* = \|\pi^*(\bar{G})\|_{\bar{a}, \bar{b}, m}, \quad \bar{d}^\bullet = \|\pi^\bullet(\bar{G})\|_{\bar{a}, \bar{b}, m}, \quad (19)$$

and

$$0 < \check{a} < \bar{a}, \quad 0 < \check{b} \leq \bar{b}\sqrt{r}, \quad \bar{d}^* + (1 + c_2)\bar{d}^\bullet \leq \bar{d},$$

then there exists a change of variables  $\check{\Phi} = I + \check{\Psi}$  satisfying the following properties:

- (i)  $\check{\Psi} \in \mathcal{X}_{\check{a}, \check{b}, m-1}$  with  $|\check{\Psi}|_{\check{a}, \check{b}, m-1} \leq c_1 \bar{d}^\bullet$ ; and
- (ii) The transformed map  $\check{F} = \check{\Phi}^{-1} \circ \bar{F} \circ \check{\Phi}$  is real analytic, has the intersection property on the cylinder  $\mathbb{T} \times (-\check{b}, \check{b})$ , and has the form  $\check{F} = A + \check{G}$ ,  $\check{G} \in \mathcal{X}_{\check{a}, \check{b}, m}$ ,

$$\|\pi^*(\check{G})\|_{\check{a}, \check{b}, m} \leq \bar{d}^* + c_2 \bar{d}^\bullet, \quad \|\pi^\bullet(\check{G})\|_{\check{a}, \check{b}, m} \leq c_3 e^{-2\pi r(\bar{a} - \check{a})/\check{b}} \bar{d}^\bullet.$$

Theorem 11 is proved in Subsection 5.5.

In order to present the main ideas of the proof, let us try to completely get rid of the remainder of the map of the form  $F = A + G$ , for some  $G \in \mathcal{X}_{a,b,m}$ , with a change of variables of the form  $\Phi = I + \Psi$ , for some  $\Psi \in \mathcal{X}_{a,b,m-1}$ . Concretely, we look for  $\Phi$  such that  $A = \Phi^{-1} \circ F \circ \Phi$ , or, equivalently, we look for  $\Psi$  such that

$$\Psi \circ A - A\Psi = G \circ (I + \Psi).$$

It is not possible to solve this equation in general. Instead, we consider the linear equation

$$\Psi \circ A - A\Psi = G.$$

This vectorial equation reads as

$$\begin{cases} \psi_1(x+y, y) - \psi_1(x, y) = y(\psi_2(x, y) + g_1(x, y)), \\ \psi_2(x+y, y) - \psi_2(x, y) = yg_2(x, y). \end{cases}$$

Therefore, we need to solve two linear equations of the form

$$\psi(x+y, y) - \psi(x, y) = yg(x, y), \quad (20)$$

where  $g \in \mathcal{X}_{a,b}$  is known. If the average of  $g(x, y)$  is different from zero:  $g^*(y) = \hat{g}_0(y) \neq 0$ , then this equation can not be solved. Besides, it is a straightforward computation to check that, if  $\hat{g}_0(y) = 0$ , the formal solution of this equation in the Fourier basis is

$$\hat{\psi}_k(y) = \frac{y}{e^{2\pi k y i} - 1} \hat{g}_k(y), \quad \forall k \neq 0, \quad (21)$$

whereas the zero-th coefficient  $\hat{\psi}_0(y)$  can be chosen arbitrarily. From (21), it is clear that (20) can not be solved unless  $g$  has only a finite number of harmonics and zero average. For this reason, given a function  $g(x, y)$  with zero average, we define its  $K$ -cut off as

$$g^{<K}(x, y) = \sum_{|k| < K} \hat{g}_k(y) e^{2\pi k x i}. \quad (22)$$

Let  $K$  be such that  $|2\pi k y| < 2\pi$  for all  $|y| < b$  and  $|k| < K$ . Hence, we will take  $K = s/b$  for some fixed  $s \in (0, 1)$ , and we will actually solve truncated linear equations of the form

$$\psi(x+y, y) - \psi(x, y) = yg^{<K}(x, y). \quad (23)$$

The Fourier norm is specially suited to analyze this kind of equations; see Lemma 14.

Summarizing these ideas, we look for a change of variables of the form  $\Phi = I + \Psi$ , where  $\Psi$  satisfies the truncated linear vectorial equation

$$\Psi \circ A - A\Psi = (G^\bullet)^{<K}, \quad (24)$$

where  $(G^\bullet)^{<K}$  denotes the  $K$ -cut off of  $G^\bullet = \pi^\bullet(G)$ . The average of the first component of  $(G^\bullet)^{<K}$  may be non-zero. Equation (24) is studied in Lemma 15. This close to the identity change of variables  $\Phi = I + \Psi$  does not completely eliminate the remainder. However, if  $b$  is small enough, it reduces the size of the  $\pi^\bullet$ -projection of the remainder as the following lemma shows.

**Lemma 12** (Iterative lemma). *Let  $m \geq 6$  be an integer,  $\bar{a} > 0$ ,  $\bar{d} > 0$ ,  $\mu > 0$ , and  $\rho \in (0, 1)$ . There exist a constant  $\bar{b} > 0$  such that if*

$$F = A + G, \quad G \in \mathcal{X}_{a,b,m}, \quad d^* = \|\pi^*(G)\|_{a,b,m}, \quad d^\bullet = \|\pi^\bullet(G)\|_{a,b,m},$$

with

$$0 < a \leq \bar{a}, \quad 0 < b \leq \min\{\bar{b}, a/6\}, \quad d^* + d^\bullet \leq \bar{d},$$

then there exists a change of variables  $\Phi = I + \Psi$  satisfying the following properties:

(i)  $\Psi \in \mathcal{X}_{a,b,m-1}$  is a solution of the truncated linear equation (24) such that

$$\|\Psi\|_{a,b,m-1} \leq \|\Psi\|_{a,b,m-1} \leq \Omega d^\bullet,$$

where  $\Omega = \Omega(\sqrt{\rho})$  is defined in Lemma 15; and

(ii) The transformed map  $\tilde{F} = \Phi^{-1} \circ F \circ \Phi$  is real analytic, has the intersection property on the cylinder  $\mathbb{T} \times (-\tilde{b}, \tilde{b})$ , and has the form  $\tilde{F} = A + \tilde{G}$ ,  $\tilde{G} \in \mathcal{X}_{\tilde{a},\tilde{b},m}$ ,

$$\|\pi^*(\tilde{G})\|_{\tilde{a},\tilde{b},m} \leq d^* + e^{-12\pi\rho} d^\bullet, \quad \|\pi^\bullet(\tilde{G})\|_{\tilde{a},\tilde{b},m} \leq e^{-12\pi\rho} d^\bullet,$$

where  $\tilde{a} = a - 6b$  and  $\tilde{b} = b - \mu b^2$ .

*Remark 5.* If  $\tilde{a} = a - 6b$ , then  $e^{-2\pi\rho(a-\tilde{a})/b} = e^{-12\pi\rho}$ .

The proof of this lemma is found in Subsection 5.4. Some technicalities in the proof require the use of the sup-norm, which forces us to deal with both the Fourier norm and the sup-norm. The relations between them are stated in Lemma 13.

Finally, Theorem 11 is obtained by means of a finite sequence of changes of variables like the ones described in the iterative lemma. We want to perform as many of such changes as possible because each change reduces the size of the  $\pi^\bullet$ -projection of the remainder by the factor  $e^{-12\pi\rho}$ . Since the loss of analyticity in the angular variable is  $6b = O(b)$ , then we can at most perform a number  $O(1/b)$  of such changes. This idea goes back to Neishtadt [30].

The intersection property is used neither in the proof of the iterative lemma nor in the proof of Theorem 11, but will be essential to control the size of the  $\pi^*$ -projections of the remainders in terms of the size of their  $\pi^\bullet$ -projections later on.

### 5.3. Technical lemmas

**Lemma 13.** *Let  $0 < \alpha < \min\{a, 1/2\pi\}$ ,  $b > 0$ , and  $g \in \mathcal{X}_{a,b}$ . Let  $g^{\geq K} = g - g^{< K}$ , with  $g^{< K}$  the  $K$ -cut off of  $g$ , defined in (22). Then:*

- (i)  $\|g^{< K}\|_{a,b} \leq \|g\|_{a,b}$
- (ii)  $\|g^{\geq K}\|_{a-\alpha,b} \leq e^{-2\pi K\alpha} \|g\|_{a,b}$ ,
- (iii)  $|g|_{a,b} \leq \|g\|_{a,b}$ , and
- (iv)  $\|g\|_{a-\alpha,b} \leq \alpha^{-1} |g|_{a,b}$ .

If  $m \in \mathbb{N}$ , then these bounds also hold for any vectorial function  $G \in \mathcal{X}_{a,b,m}$ .

*Proof.* First, the Fourier norm of  $g^{<K}$  is a partial sum of the Fourier norm of  $g$ . Second,  $\|g^{\geq K}\|_{a-\alpha,b} = \sum_{|k|\geq K} |\hat{g}_k|_b e^{2\pi|k|(a-\alpha)} = e^{-2\pi K\alpha} \sum_{|k|\geq K} |\hat{g}_k|_b e^{2\pi|k|a} \leq e^{-2\pi K\alpha} \|g\|_{a,b}$ . Third,  $|g(x,y)| \leq \sum_{k\in\mathbb{Z}} |\hat{g}_k(y)| |e^{2\pi kxi}| \leq \sum_{k\in\mathbb{Z}} |\hat{g}_k|_b e^{2\pi|k|a} = \|g\|_{a,b}$ , for all  $(x,y) \in D_{a,b}$ . Fourth, we recall that the Fourier coefficients of the analytic function  $g$  satisfy the inequality  $|\hat{g}_k|_b \leq e^{-2\pi|k|a} \|g\|_{a,b}$  for all  $k \in \mathbb{Z}$ . Hence,

$$\|g\|_{a-\alpha,b} = \sum_{k\in\mathbb{Z}} |\hat{g}_k|_b e^{2\pi|k|(a-\alpha)} \leq 2\|g\|_{a,b} \sum_{k\geq 0} e^{-2\pi k\alpha} \leq \alpha^{-1} \|g\|_{a,b},$$

where we have used that  $\sum_{k\geq 0} e^{-kt} = (1 - e^{-t})^{-1} \leq e/t < \pi/t$  for all  $t \in (0, 1)$ . The last part follows from the definition of the norms  $\|\cdot\|_{a,b,m}$  and  $|\cdot|_{a,b,m}$ .  $\square$

**Lemma 14.** *If  $s \in (0, 1)$ ,  $K = s/b$ , and  $g \in \mathcal{X}_{a,b}$  is a function with zero average, then the truncated linear equation (23) has a unique solution  $\psi \in \mathcal{X}_{a,b}$  with zero average and  $\|\psi\|_{a,b} \leq \omega \|g\|_{a,b}$ , where*

$$\omega = \omega(s) = \frac{1}{2\pi} \cdot \max_{|z|\leq 2\pi s} \left| \frac{z}{e^z - 1} \right|. \quad (25)$$

*Proof.* The Fourier coefficients of  $\psi$  must satisfy (21). We note that  $\omega < \infty$  for all  $s \in (0, 1)$ , since the function  $z/(e^z - 1)$  is analytic on the open ball  $|z| < 2\pi$ . Moreover,

$$|\hat{\psi}_k|_b \leq \left( \max_{|y|\leq b} \left| \frac{y}{e^{2\pi kyi} - 1} \right| \right) |\hat{g}_k|_b \leq \frac{\omega}{|k|} |\hat{g}_k|_b \leq \omega |\hat{g}_k|_b,$$

for all  $0 < |k| < K = s/b$ . Finally, we recall that  $\hat{\psi}_0(y) \equiv 0$ . Then we obtain that  $\|\psi\|_{a,b} = \sum_{0 < |k| < K} |\hat{\psi}_k|_b e^{2\pi|k|a} \leq \omega \sum_{k\in\mathbb{Z}} |\hat{g}_k|_b e^{2\pi|k|a} = \omega \|g\|_{a,b}$ .  $\square$

*Remark 6.* We will denote by  $\psi = \mathcal{G}_K(g^{<K})$  the linear operator that sends the independent term  $g^{<K}$  of the truncated linear equation (23) to the solution  $\psi$  with zero average. Note that the solution  $\psi$  has no harmonics of order  $\geq K$ .

**Lemma 15.** *If  $m \geq 1$ ,  $s \in (0, 1)$ ,  $K = s/b$ , and  $G \in \mathcal{X}_{a,b,m}$ , then the truncated linear equation (24) has a solution  $\Psi \in \mathcal{X}_{a,b,m-1}$  such that*

$$\|\Psi\|_{a,b,m-1} \leq \Omega \|G^\bullet\|_{a,b,m},$$

where  $G^\bullet = \pi^\bullet(G)$ ,  $\Omega = \Omega(s) = (\omega(s) + 1) \max\{1, \omega(s)\}$ , and  $\omega(s)$  is defined in (25).

*Proof.* Let  $G = (y^m g_1, y^{m+1} g_2)$  and  $\Psi = (y^{m-1} \psi_1, y^m \psi_2)$ . Then the vectorial equation  $\Psi \circ A - A\Psi = (G^\bullet)^{<K}$  reads as

$$\begin{cases} \psi_1(x+y, y) - \psi_1(x, y) = y(\psi_2(x, y) + g_1^{<K}(x, y)), \\ \psi_2(x+y, y) - \psi_2(x, y) = y(g_2^{<K}(x, y) - g_2^*(y)). \end{cases}$$

Let  $\psi_2 = \mathcal{G}_K(g_2^{<K} - g_2^*) - g_1^*$  and  $\psi_1 = \mathcal{G}_K(\psi_2 + g_1^{<K})$ . These operations are well-defined since both  $g_2 - g_2^*$  and  $\psi_2 + g_1$  have zero average. As for the bounds,

$$\begin{aligned} \|\psi_2\|_{a,b} &\leq \omega \|g_2 - g_2^*\|_{a,b} + \|g_1\|_{a,b} \leq \Omega \|G^\bullet\|_{a,b,m}, \\ \|\psi_1\|_{a,b} &\leq \omega \|\psi_2 + g_1^{<K}\|_{a,b} \\ &\leq \omega \|\psi_2 + g_1^*\|_{a,b} + \omega \|g_1^{<K} - g_1^*\|_{a,b} \end{aligned}$$

$$\leq \omega^2 \|g_2 - g_2^*\|_{a,b} + \omega \|g_1\|_{a,b} \leq \Omega \|G^\bullet\|_{a,b,m},$$

where we have used Lemma 14.  $\square$

**Lemma 16.** *Let  $l, n \in \mathbb{N}$ ,  $0 < \alpha < \min\{a/3, 1/2\pi\}$ ,  $0 < \beta < b/2$ ,  $c_1, c_2 > 0$ , and  $c = c_1 + c_2$ , such that*

$$b + b^n c < \alpha + \beta, \quad b^{n+1} c < \beta. \quad (26)$$

Let  $M = M(\alpha, b, \beta, c_1, c_2, l, n) = (1 + b^n c)^{l+1} b^{l-1} (\alpha^{-1} + b\beta^{-1} + l + 1)$ . Let  $\Delta \in \mathcal{X}_{a,b,l}$ . Let  $\Gamma_1, \Gamma_2 \in \mathcal{X}_{a-2\alpha, b-2\beta, n}$  with  $\|\Gamma_j\|_{a-2\alpha, b-2\beta, n} \leq c_j$ . Let  $L(x, y) = (x + \eta y, y)$  with  $|\eta| \leq 1$ . Then,

- (i)  $\Delta \circ (L + \Gamma_1) - \Delta \circ (L + \Gamma_2) \in \mathcal{X}_{a-3\alpha, b-2\beta, n+l}$ ,
- (ii)  $|\Delta \circ (L + \Gamma_1) - \Delta \circ (L + \Gamma_2)|_{a-2\alpha, b-2\beta, n+l} \leq M |\Delta|_{a,b,l} |\Gamma_1 - \Gamma_2|_{a-2\alpha, b-2\beta, n}$ , and
- (iii)  $\|\Delta \circ (L + \Gamma_1) - \Delta \circ (L + \Gamma_2)\|_{a-3\alpha, b-2\beta, n+l} \leq M \alpha^{-1} \|\Delta\|_{a,b,l} \|\Gamma_1 - \Gamma_2\|_{a-2\alpha, b-2\beta, n}$ .

*Proof.* Let  $\Gamma = \Gamma_1 - \Gamma_2$ . Then  $|\Gamma|_{a-2\alpha, b-2\beta, n} \leq \|\Gamma\|_{a-2\alpha, b-2\beta, n} \leq c$  and

$$\Delta \circ (L + \Gamma_1) - \Delta \circ (L + \Gamma_2) = \int_0^1 (D\Delta \circ (L + t\Gamma)) \cdot \Gamma dt.$$

Let  $(x_t, y_t) = (L + t\Gamma)(x, y) = (x + \eta y + ty^n \gamma_1(x, y), y + ty^{n+1} \gamma_2(x, y))$ , with  $t \in [0, 1]$  fixed. We deduce from conditions (26) that  $(x_0, y_0)$  and  $(x_1, y_1)$  belong to  $D_{a-\alpha, b-\beta}$  for all  $(x, y) \in D_{a-2\alpha, b-2\beta}$ . Therefore,  $(x_t, y_t) \in D_{a-\alpha, b-\beta}$  by convexity of the domain, and so, the composition  $\Delta \circ (L + t\Gamma)$  is well-defined on the domain  $D_{a-2\alpha, b-2\beta}$ .

A simple computation shows that the product  $D\Delta(x_t, y_t) \cdot \Gamma(x, y)$  is equal to

$$\left( \begin{array}{l} y_t^{l-1} y^{n+1} (y^{-1} y_t \partial_1 \delta_1(x_t, y_t) \gamma_1(x, y) + (y_t \partial_2 \delta_1(x_t, y_t) + l \delta_1(x_t, y_t)) \gamma_2(x, y)) \\ y_t^l y^{n+1} (y^{-1} y_t \partial_1 \delta_2(x_t, y_t) \gamma_1(x, y) + (y_t \partial_2 \delta_2(x_t, y_t) + (l+1) \delta_2(x_t, y_t)) \gamma_2(x, y)) \end{array} \right).$$

Let us bound the elements above. On the one hand,  $|\gamma_i(x, y)| \leq |\Gamma|_{a-2\alpha, b-2\beta, n} \leq c$ ,  $|y_t| = |y + ty^{n+1} \gamma_2(x, y)| \leq (1 + b^n c) |y|$ , and  $|y| < b$  for all  $(x, y) \in D_{a-2\alpha, b-2\beta}$ . On the other hand,  $|\delta_i(x_s, y_s)| \leq |\Delta|_{a,b,l}$  and the Cauchy estimates imply that

$$|\partial_1 \delta_i(x_t, y_t)| \leq \alpha^{-1} |\Delta|_{a,b,l}, \quad |\partial_2 \delta_i(x_t, y_t)| \leq \beta^{-1} |\Delta|_{a,b,l}.$$

From the previous bounds and the definitions of both norms, we deduce that

$$|(D\Delta \circ (L + t\Gamma)) \cdot \Gamma|_{a-2\alpha, b-2\beta, n+l} \leq M' |\Delta|_{a,b,l} |\Gamma|_{a-2\alpha, b-2\beta, n},$$

for all  $t \in [0, 1]$ , where  $M' = (1 + b^n c)^{l+1} (\alpha^{-1} + b\beta^{-1} + l + 1)$ . Thus,

$$\begin{aligned} |\Delta \circ (L + \Gamma_1) - \Delta \circ (L + \Gamma_2)|_{a-2\alpha, b-2\beta, n+l} &\leq M' |\Delta|_{a,b,l} |\Gamma|_{a-2\alpha, b-2\beta, n}, \\ \|\Delta \circ (L + \Gamma_1) - \Delta \circ (L + \Gamma_2)\|_{a-3\alpha, b-2\beta, n+l} &\leq \alpha^{-1} M' \|\Delta\|_{a,b,l} \|\Gamma\|_{a-2\alpha, b-2\beta, n}. \end{aligned}$$

This proves the first item. The other items follow from the bounds  $|\cdot|_{a,b,n+1} \leq b^{l-1} |\cdot|_{a,b,n+l}$  and  $\|\cdot\|_{a,b,n+1} \leq b^{l-1} \|\cdot\|_{a,b,n+l}$ , since  $M = M' b^{l-1}$ .  $\square$

**Lemma 17.** *Let  $n \in \mathbb{N}$ ,  $0 < \alpha < \min\{a/3, 1/2\pi\}$ ,  $0 < \beta < b/2$ , and  $p > 0$  such that conditions (26) hold with  $c = 2p$ . Let  $\Phi = I + \Psi$ , with  $\Psi \in \mathcal{X}_{a,b,n}$  and  $\|\Psi\|_{a,b,n} \leq p$ . Then  $\Phi(D_{a',b'}) \subset D_{a'+\alpha,b'+\beta}$  for all  $0 < a' \leq a$  and  $0 < b' \leq b$ .*

*Let  $M_* = M(\alpha, b, \beta, p, p, n, n)$ , where  $M$  is defined in Lemma 16. If  $M_*bp < 1$ , then  $\Phi$  is invertible and the inverse change  $\Phi^{-1}$  satisfies the following properties:*

- (i)  $\Phi^{-1} = I + \Upsilon$  for some  $\Upsilon \in \mathcal{X}_{a-2\alpha,b-2\beta,n}$  such that  $|\Upsilon|_{a-2\alpha,b-2\beta,n} \leq |\Psi|_{a,b,n}$ ,
- (ii)  $\Phi^{-1}(D_{a',b'}) \subset D_{a'+\alpha,b'+\beta}$  for all  $0 < a' \leq a - 2\alpha$  and  $0 < b' \leq b - 2\beta$ , and
- (iii)  $\|\Upsilon + \Psi\|_{a-3\alpha,b-2\beta,n+1} \leq M_*\alpha^{-1}\|\Psi\|_{a,b,n}^2$ .

*Proof.* Note that  $|\Psi|_{a,b,n} \leq \|\Psi\|_{a,b,n} \leq p$ . Conditions (26) imply that  $b^n p < \alpha$  and  $b^{n+1} p < \beta$ . Therefore,  $\Phi(D_{a',b'}) \subset D_{a'+\alpha,b'+\beta}$  for all  $0 < a' \leq a$  and  $0 < b' \leq b$ . Analogously, if  $\Upsilon \in \mathcal{X}_{a-2\alpha,b-2\beta,n}$  and  $|\Upsilon|_{a-2\alpha,b-2\beta,n} \leq p$ , then  $(I + \Upsilon)(D_{a',b'}) \subset D_{a'+\alpha,b'+\beta}$  for all  $0 < a' \leq a - 2\alpha$ ,  $0 < b' \leq b - 2\beta$ .

We denote by  $\mathcal{B}$  the closed ball in  $\mathcal{X}_{a-2\alpha,b-2\beta,n}$  of radius  $p$  in the sup-norm. Let us prove that the functional  $\mathcal{P} : \mathcal{B} \rightarrow \mathcal{B}$ ,  $\mathcal{P}(\Upsilon) = -\Psi \circ (I + \Upsilon)$ , is a well-defined contraction with Lipschitz constant

$$\text{Lip } \mathcal{P} \leq M_*b|\Psi|_{a,b,n} \leq M_*bp < 1. \quad (27)$$

First, we observe that

$$|\mathcal{P}(\Upsilon)|_{a-2\alpha,b-2\beta,n} \leq |\Psi|_{a,b,n} \leq p, \quad \forall \Upsilon \in \mathcal{B}, \quad (28)$$

so  $\mathcal{P}(\mathcal{B}) \subset \mathcal{B}$ . Second, we bound  $\mathcal{P}(\Upsilon) - \mathcal{P}(\Xi) = \Psi \circ (I + \Xi) - \Psi \circ (I + \Upsilon)$  as follows:

$$\begin{aligned} |\mathcal{P}(\Upsilon) - \mathcal{P}(\Xi)|_{a-2\alpha,b-2\beta,n} &\leq b|\mathcal{P}(\Upsilon) - \mathcal{P}(\Xi)|_{a-2\alpha,b-2\beta,n+1} \\ &\leq M_*b|\Psi|_{a,b,n}|\Xi - \Upsilon|_{a-2\alpha,b-2\beta,n}. \end{aligned} \quad (29)$$

The first inequality is direct, and the second comes from Lemma 16 with  $\Delta = \Psi$ ,  $\Gamma_1 = \Xi$ ,  $\Gamma_2 = \Upsilon$ ,  $L = I$ ,  $c_j = p$ , and  $l = n$ . This proves that  $\mathcal{P}$  is a contraction with Lipschitz constant (27). Thus,  $\mathcal{P}$  has a unique fixed point  $\Upsilon \in \mathcal{B}$  which satisfies that

$$(I + \Psi) \circ (I + \Upsilon) = I + \Upsilon + \Psi \circ (I + \Upsilon) = I + \Upsilon - \mathcal{P}(\Upsilon) = I$$

on  $D_{a-2\alpha,b-2\beta}$ . Therefore, the inverse map  $\Phi^{-1}$  exists and equals  $I + \Upsilon$ . Furthermore,  $|\Upsilon|_{a-2\alpha,b-2\beta,n} \leq |\Psi|_{a,b,n}$  follows from (28). Finally,

$$\begin{aligned} \|\Upsilon + \Psi\|_{a-3\alpha,b-2\beta,n+1} &\leq \alpha^{-1}|\Upsilon + \Psi|_{a-2\alpha,b-2\beta,n+1} \\ &\leq \alpha^{-1}|\mathcal{P}(\Upsilon) - \mathcal{P}(\mathbf{0})|_{a-2\alpha,b-2\beta,n+1} \\ &\leq M_*\alpha^{-1}|\Psi|_{a,b,n}|\Upsilon|_{a-2\alpha,b-2\beta,n} \\ &\leq M_*\alpha^{-1}|\Psi|_{a,b,n}^2 \leq M_*\alpha^{-1}\|\Psi\|_{a,b,n}^2. \end{aligned}$$

We have used the second inequality of equation (29) with  $\Xi = \mathbf{0}$ . □



## 5.4. Proof of Lemma 12

We recall that  $F = A + G$  with  $G = \pi^*(G) + \pi^\bullet(G) = G^* + G^\bullet \in \mathcal{X}_{a,b,m}$ ,  $d^* = \|G^*\|_{a,b,m}$ ,  $d^\bullet = \|G^\bullet\|_{a,b,m}$ ,  $d^* + d^\bullet \leq \bar{d}$ , and  $m \geq 6$ . Let  $s = \sqrt{\rho} \in (\rho, 1)$ . Let  $\Omega = \Omega(s) = \Omega(\sqrt{\rho}) > \rho$  be the constant introduced in Lemma 15 and  $\sigma = 1 + 2\Omega$ . Let  $\Phi = I + \Psi$  be the change of variables where  $\Psi \in \mathcal{X}_{a,b,m-1}$  is the solution given in Lemma 15 of the truncated linear equation  $\Psi \circ A - A\Psi = (G^\bullet)^{<K}$  with  $K = s/b$ , so that

$$\|\Psi\|_{a',b',m-1} \leq \|\Psi\|_{a',b',m-1} \leq \Omega \|G^\bullet\|_{a',b',m} \leq \Omega d^\bullet, \quad (30)$$

for all  $0 < a' \leq a$  and  $0 < b' \leq b$ . Let  $\Phi^{-1} = I + \Upsilon$  be the inverse change studied in Lemma 17. Let  $\tilde{F} = \Phi^{-1} \circ F \circ \Phi$  be the transformed map. Let  $\tilde{G} = \tilde{F} - A$  be the new remainder.

Henceforth, we will assume that  $\alpha$ ,  $b$ , and  $\beta$  are some positive constants such that

$$b \leq \alpha < \min\{a/6, 1/2\pi\}, \quad 0 < \beta < b/4, \quad b + \sigma b^{m-1} \bar{d} < \alpha + \beta, \quad \sigma b^m \bar{d} < \min(\alpha, \beta). \quad (31)$$

We split the proof in four steps.

*Step 1: Control of the domains.* Note that  $\tilde{F}(D_{a',b'}) \subset D_{a'+4\alpha, b'+3\beta}$  for all  $0 < a' \leq a - 4\alpha$  and  $0 < b' \leq b - 4\beta$ . Indeed,

$$\tilde{F} : D_{a',b'} \xrightarrow{\Phi} D_{a'+\alpha, b'+\beta} \xrightarrow{F} D_{a'+3\alpha, b'+2\beta} \xrightarrow{\Phi^{-1}} D_{a'+4\alpha, b'+3\beta}.$$

The behaviours of the changes  $\Phi$  and  $\Phi^{-1}$  follow directly from Lemma 17, which can be applied since conditions (31) are more restrictive than the ones required in Lemma 17 when  $p = \Omega \bar{d}$  and  $n = m - 1$ . We also need that  $a' + 2\alpha \leq a - 2\alpha$  and  $b' + 2\beta \leq b - 2\beta$  in order to control the inverse  $\Phi^{-1}$ , which explains the restrictions on  $a'$  and  $b'$ .

The behaviour of the map  $F = A + G$  follows from the bound

$$|G|_{a,b,m} \leq \|G\|_{a,b,m} \leq \|G^*\|_{a,b,m} + \|G^\bullet\|_{a,b,m} = d^* + d^\bullet \leq \bar{d}$$

and conditions  $b + b^m \bar{d} < 2\alpha$  and  $b^{m+1} \bar{d} < \beta$ , which are also a consequence of (31).

*Step 2: Decomposition of the new remainder.* It turns out that  $\tilde{G} = G^* + \sum_{j=1}^4 \tilde{G}_j$ , where

$$\begin{aligned} \tilde{G}_1 &= (G^\bullet)^{\geq K} = G^\bullet - (G^\bullet)^{<K}, & \tilde{G}_2 &= G \circ \Phi - G, \\ \tilde{G}_3 &= \Psi \circ A - \Psi \circ F \circ \Phi, & \tilde{G}_4 &= (\Upsilon + \Psi) \circ (F \circ \Phi). \end{aligned} \quad (32)$$

Indeed,  $G^* + \tilde{G}_1 + \tilde{G}_2 = G \circ \Phi - (G^\bullet)^{<K}$  and  $\tilde{G}_3 + \tilde{G}_4 = \Psi \circ A + \Upsilon \circ (F \circ \Phi)$ , so

$$\begin{aligned} G^* + \sum_{j=1}^4 \tilde{G}_j &= G \circ \Phi + A\Psi + \Upsilon \circ (F \circ \Phi) \\ &= (F - A) \circ \Phi + A(\Phi - I) + (\Phi^{-1} - I) \circ (F \circ \Phi) \\ &= \Phi^{-1} \circ F \circ \Phi - A = \tilde{G}. \end{aligned}$$

Finally, let  $\tilde{G}^* = \pi^*(\tilde{G}) = G^* + \sum_{j=2}^4 \pi^*(G_j)$  and  $\tilde{G}^\bullet = \pi^\bullet(\tilde{G}) = \sum_{j=1}^4 \pi^\bullet(G_j)$ .

*Step 3: Bounds of the projections of the new remainder.* Lemma 13 and the bound (30) will be used several times in what follows. Below, we apply Lemma 16 (twice) and Lemma 17 (once). The required hypotheses in each case are satisfied due to conditions (31).

- If  $\tilde{a} < a$  and  $\tilde{b} \leq b$ , then

$$\|\tilde{G}_1\|_{\tilde{a},\tilde{b},m} = \|(G^\bullet)^{\geq K}\|_{\tilde{a},\tilde{b},m} \leq e^{-2\pi K(a-\tilde{a})} \|G^\bullet\|_{a,\tilde{b},m} \leq e^{-2\pi K(a-\tilde{a})} d^\bullet.$$

- If  $\tilde{a} \leq a - 3\alpha$  and  $\tilde{b} \leq b - 2\beta$ , then

$$\begin{aligned} \|\tilde{G}_2\|_{\tilde{a},\tilde{b},m} &\leq M_2 \alpha^{-1} \|G\|_{\tilde{a}+3\alpha,\tilde{b}+2\beta,m} \|\Psi\|_{\tilde{a}+\alpha,\tilde{b},m-1} \\ &\leq \Omega M_2 \alpha^{-1} \|G\|_{\tilde{a}+3\alpha,\tilde{b}+2\beta,m} \|G^\bullet\|_{\tilde{a}+\alpha,\tilde{b},m} \\ &\leq \Omega M_2 \alpha^{-1} \bar{d} d^\bullet. \end{aligned}$$

The first inequality follows from Lemma 16 with  $\Delta = G$ ,  $L = \mathbf{I}$ ,  $\Gamma_1 = \Psi$ ,  $\Gamma_2 = \mathbf{0}$ ,  $l = m$ , and  $n = m - 1$ , so that  $M_2 = M(\alpha, b, \beta, \Omega \bar{d}, 0, m, m - 1)$ .

- If  $\tilde{a} \leq a - 2\alpha$  and  $\tilde{b} \leq b - \beta$ , then  $\|F \circ \Phi - A\|_{\tilde{a},\tilde{b},m-1} \leq \sigma \|G\|_{a,b,m}$ . Indeed,  $F \circ \Phi - A = A\Psi + G \circ \Phi$  and

$$\begin{aligned} \|A\Psi\|_{\tilde{a},\tilde{b},m-1} &\leq 2\|\Psi\|_{\tilde{a},\tilde{b},m-1} \leq 2\Omega \|G^\bullet\|_{\tilde{a},\tilde{b},m} \leq 2\Omega d^\bullet \leq 2\Omega \bar{d}, \\ \|G \circ \Phi\|_{\tilde{a},\tilde{b},m-1} &\leq \alpha^{-1} |G \circ \Phi|_{\tilde{a}+\alpha,\tilde{b},m-1} \leq \alpha^{-1} |G|_{\tilde{a}+2\alpha,\tilde{b}+\beta,m-1} \\ &\leq \alpha^{-1} \|G\|_{\tilde{a}+2\alpha,\tilde{b}+\beta,m-1} \leq b\alpha^{-1} \|G\|_{\tilde{a}+2\alpha,\tilde{b}+\beta,m} \\ &\leq b\alpha^{-1} \|G\|_{a,b,m} \leq \|G\|_{a,b,m} \leq \bar{d}. \end{aligned}$$

We have used that  $\Phi(D_{\tilde{a}+\alpha,\tilde{b}}) \subset D_{\tilde{a}+2\alpha,\tilde{b}+\beta}$  to bound  $|G \circ \Phi|_{\tilde{a},\tilde{b},m-1}$ .

- If  $\tilde{a} \leq a - 3\alpha$  and  $\tilde{b} \leq b - 2\beta$ , then

$$\begin{aligned} \|\tilde{G}_3\|_{\tilde{a},\tilde{b},m} &\leq M_3 \alpha^{-1} \|\Psi\|_{\tilde{a}+3\alpha,\tilde{b}+2\beta,m-1} \|F \circ \Phi - A\|_{\tilde{a}+\alpha,\tilde{b},m-1} \\ &\leq M_3 \alpha^{-1} \Omega \|G^\bullet\|_{\tilde{a}+3\alpha,\tilde{b}+2\beta,m} \sigma \bar{d} \\ &\leq \Omega \sigma M_3 \alpha^{-1} \bar{d} d^\bullet. \end{aligned}$$

The first inequality follows from Lemma 16 with  $\Delta = \Psi$ ,  $L = A$ ,  $\Gamma_1 = \mathbf{0}$ ,  $\Gamma_2 = F \circ \Phi - A$ ,  $l = m - 1$ , and  $n = m - 1$ , so that  $M_3 = M(\alpha, b, \beta, 0, \sigma \bar{d}, m - 1, m - 1)$ .

- If  $\tilde{a} \leq a - 6\alpha$  and  $\tilde{b} \leq b - 4\beta$ , then

$$\begin{aligned} \|\tilde{G}_4\|_{\tilde{a},\tilde{b},m} &\leq \alpha^{-1} |(\Upsilon + \Psi) \circ (F \circ \Phi)|_{\tilde{a}+\alpha,\tilde{b},m} \leq \alpha^{-1} |\Upsilon + \Psi|_{\tilde{a}+3\alpha,\tilde{b}+2\beta,m} \\ &\leq \alpha^{-1} \|\Upsilon + \Psi\|_{\tilde{a}+\alpha,\tilde{b},m} \leq M_4 \alpha^{-2} \|\Psi\|_{\tilde{a}+6\alpha,\tilde{b}+4\beta,m-1}^2 \\ &\leq M_4 \alpha^{-2} (\Omega \|G^\bullet\|_{\tilde{a}+6\alpha,\tilde{b}+4\beta,m})^2 \\ &\leq \Omega^2 M_4 \alpha^{-2} (d^\bullet)^2 \leq \Omega^2 M_4 \alpha^{-2} \bar{d} d^\bullet. \end{aligned}$$

The second inequality uses the inclusion  $(F \circ \Phi)(D_{\tilde{a}+\alpha,\tilde{b}}) \subset D_{\tilde{a}+3\alpha,\tilde{b}+2\beta}$ . The fourth one follows from Lemma 17 with  $M_4 = M(\alpha, b, \beta, \Omega \bar{d}, \Omega \bar{d}, m - 1, m - 1)$ . We need to verify the hypothesis  $M_* b p < 1$  in this last lemma. It turns out that  $M_* b p = M_4 b \Omega \bar{d} = O(b^{m-2})$ , so it suffices to take  $0 < b \leq \bar{b}$ , with  $\bar{b}$  small enough.

- If  $\tilde{a} \leq a - 6\alpha$  and  $\tilde{b} \leq b - 4\beta$ , then

$$\begin{aligned} \|\tilde{G}^*\|_{\tilde{a},\tilde{b},m} &\leq \|G^*\|_{\tilde{a},\tilde{b},m} + \sum_{j=2}^4 \|\tilde{G}_j\|_{\tilde{a},\tilde{b},m} \leq d^* + \tilde{M} \bar{d} d^\bullet, \\ \|\tilde{G}^\bullet\|_{\tilde{a},\tilde{b},m} &\leq \sum_{j=1}^4 \|\tilde{G}_j\|_{\tilde{a},\tilde{b},m} \leq \left( e^{-2\pi K(a-\tilde{a})} + \tilde{M} \bar{d} \right) d^\bullet, \end{aligned}$$

where  $\tilde{M} = \Omega \alpha^{-1} (M_2 + \sigma M_3 + \Omega \alpha^{-1} M_4)$  and the constants  $M_j$ ,  $j = 2, 3, 4$ , have been defined previously.

*Step 4: Choice of the loss of analyticity domain.* We set  $\alpha = b$  and  $\beta = \mu b^2/4$ . If  $\bar{b} > 0$  is small enough, then conditions (31) hold for all  $0 < b \leq \bar{b}$ . In addition,

$$M = M(\alpha, b, \beta, c_1, c_2, l, n) = O(b^{l-2}) \quad \text{as } b \rightarrow 0^+,$$

where  $M$  is the expression introduced in Lemma 16. If we take  $\tilde{a} = a - 6\alpha$  and  $\tilde{b} = b - 4\beta$ , then the bounds of the previous step imply that

$$\|\tilde{G}^*\|_{\tilde{a}, \tilde{b}, m} \leq d^* + \tilde{M} \bar{d} d^\bullet, \quad \|\tilde{G}^\bullet\|_{\tilde{a}, \tilde{b}, m} \leq \left( e^{-12\pi K b} + \tilde{M} \bar{d} \right) d^\bullet,$$

where  $\tilde{M} = \tilde{M}(b; d, m, s) = \Omega b^{-1} (M_2 + \sigma M_3 + \Omega b^{-1} M_4) = O(b^{m-5})$ . We recall that  $m \geq 6$ ,  $0 < \rho < s < 1$ , and  $K = s/b$ . Hence, if  $0 < b \leq \bar{b}$  and  $\bar{b}$  is small enough, then

$$\|\tilde{G}^*\|_{\tilde{a}, \tilde{b}, m} \leq d^* + e^{-12\pi\rho} d^\bullet, \quad \|\tilde{G}^\bullet\|_{\tilde{a}, \tilde{b}, m} \leq e^{-12\pi\rho} d^\bullet.$$

Indeed,  $\tilde{M} \bar{d} \leq e^{-12\pi\rho} - e^{-12\pi s} \leq e^{-12\pi\rho}$  if we take a small enough value of  $\bar{b}$ .

This ends the proof of the iterative lemma.

### 5.5. Proof of Theorem 11

Set  $\rho = \sqrt{r} \in (0, 1)$ ,  $\mu = 6(1 - \rho)/(\bar{a} - \check{a})$ , and  $\Omega = \Omega(\sqrt{\rho})$ , where the function  $\Omega(s)$  is defined in Lemma 15. Let  $\bar{b}$  be the positive constant associated to the integer  $m \geq 6$  in Lemma 12, the numbers  $\bar{a}, \bar{d}, \mu > 0$ , and the exponent  $\rho \in (0, 1)$ . Let  $c_1 = c_1(r) = \Omega \sum_{n \geq 0} e^{-12\pi\rho n}$ ,  $c_2 = c_2(r) = \sum_{n \geq 1} e^{-12\pi\rho n}$ , and  $c_3 = c_3(r) = e^{12\pi\rho}$ .

Let us check that  $\bar{b}$ ,  $c_1$ ,  $c_2$ , and  $c_3$  satisfy the properties given in Theorem 11.

Let  $a_0 = \bar{a}$ ,  $d_0^* = \bar{d}^*$ ,  $d_0^\bullet = \bar{d}^\bullet$ ,  $0 < b_0 = \check{b}/\rho \leq \bar{b}$ ,  $F_0 = \bar{F} = A + \bar{G}$  be the map given in (19),  $G_0^* = \pi^*(\bar{G})$ , and  $G_0^\bullet = \pi^\bullet(\bar{G})$ . By recursively applying Lemma 12, we obtain a sequence of changes of variables  $\Phi_n = I + \Psi_n$ , with  $\Psi_n \in \mathcal{X}_{a_{n-1}, b_{n-1}, m-1}$ , and a sequence of maps  $F_n = A + G_n$ , with  $G_n = G_n^* + G_n^\bullet$ ,  $G_n^* \in \mathcal{X}_{a_n, b_n, m}^*$  and  $G_n^\bullet \in \mathcal{X}_{a_n, b_n, m}^\bullet$ , such that

$$\|\Psi_n\|_{a_{n-1}, b_{n-1}, m-1} \leq \Omega d_{n-1}^\bullet, \quad \|G_n^*\|_{a_n, b_n, m} \leq d_n^*, \quad \|G_n^\bullet\|_{a_n, b_n, m} \leq d_n^\bullet,$$

with  $a_{n+1} = a_n - 6b_n$ ,  $b_{n+1} = b_n - \mu b_n^2$ ,  $d_{n+1}^* = d_n^* + e^{-12\pi\rho} d_n^\bullet$ , and  $d_{n+1}^\bullet = e^{-12\pi\rho} d_n^\bullet$ .

Let  $N$  be the biggest integer satisfying  $Nb_0 \leq (\bar{a} - \check{a})/6$ . The sequences  $(a_n)_{0 \leq n \leq N}$ ,  $(b_n)_{0 \leq n \leq N}$ , and  $(d_n^\bullet)_{0 \leq n \leq N}$  are decreasing. The sequence  $(d_n^*)_{0 \leq n \leq N}$  is increasing. Indeed,

$$\begin{aligned} a_N &= a_{N-1} - 6b_{N-1} \geq a_{N-1} - 6b_0 \geq \dots \geq a_0 - 6Nb_0 \geq \check{a}, \\ b_N &= b_{N-1} - \mu b_{N-1}^2 \geq b_{N-1} - \mu b_0^2 \geq \dots \geq (1 - \mu N b_0) b_0 \geq \rho b_0 = \check{b}, \\ d_N^\bullet &\leq e^{-12\pi\rho} d_{N-1}^\bullet \leq \dots \leq e^{-12\pi\rho N} d_0^\bullet \leq c_3 e^{-2\pi r (\bar{a} - \check{a}) / \check{b}} \bar{d}^\bullet, \\ d_N^* &\leq d_{N-1}^* + e^{-12\pi\rho} d_{N-1}^\bullet \leq \dots \leq d_0^* + \left( \sum_{n=1}^N e^{-12\pi\rho n} \right) d_0^\bullet \leq \bar{d}^* + c_2 \bar{d}^\bullet, \end{aligned}$$

and  $d_n^* + d_n^\bullet \leq d_N^* + d_0^\bullet \leq \bar{d}^* + (1 + c_2) \bar{d}^\bullet \leq \bar{d}$  for all  $n = 0, \dots, N$ .

We can apply  $N$  times the iterative lemma. Let  $\check{F} = A + \check{G} = A + G_N = F_N$  be the map obtained after those  $N$  steps. Then

$$\begin{aligned} \|\pi^*(\check{G})\|_{\check{a}, \check{b}, m} &\leq \|\pi^*(G_N)\|_{a_N, b_N, m} \leq d_N^* \leq \bar{d}^* + c_2 \bar{d}^\bullet, \\ \|\pi^\bullet(\check{G})\|_{\check{a}, \check{b}, m} &\leq \|\pi^\bullet(G_N)\|_{a_N, b_N, m} \leq d_N^\bullet \leq c_3 e^{-2\pi r (\bar{a} - \check{a}) / \check{b}} \bar{d}^\bullet. \end{aligned}$$

Finally, let  $\check{\Phi} = \Phi_N \circ \cdots \circ \Phi_1$  be the change of variables such that  $\check{F} = \check{\Phi}^{-1} \circ \bar{F} \circ \check{\Phi}$ . We want to check that  $\check{\Phi} = I + \check{\Psi}$  for some  $\check{\Psi} \in \mathcal{X}_{\check{a}, \check{b}, m-1}$  such that  $|\check{\Psi}|_{\check{a}, \check{b}, m-1} \leq c_1 \check{d}^\bullet$ . We note that

$$\check{\Psi} = \Psi_1 + \cdots + \Psi_N,$$

where each term of the above summation is evaluated at a different argument. Nevertheless, those arguments are not important when computing the sup-norm:

$$|\check{\Psi}|_{\check{a}, \check{b}, m-1} \leq \sum_{n=1}^N |\Psi_n|_{a_{n-1}, b_{n-1}, m-1} \leq \Omega \sum_{n=0}^{N-1} d_n^\bullet \leq \Omega d_0^\bullet \sum_{n=0}^{N-1} e^{-12\pi\rho n} = c_1 \bar{d}^\bullet.$$

This ends the proof of Theorem 11.

### 5.6. Proof of Theorem 2

Let us begin with a simple, but essential, chain of inequalities associated to certain analyticity strip widths that will appear along the proof. If  $\alpha \in (0, a_*)$ , then there exists  $r \in (0, 1)$ ,  $\bar{b} > 0$ , and some analyticity strip widths  $a_2$ ,  $\bar{a} = a_m$ , and  $\check{a} = \bar{b}$ , such that

$$0 < \bar{b} =: \check{a} := \bar{a} - (1 + \bar{b})\alpha/r < \bar{a} := a_m < a_2 < a_*. \quad (33)$$

The first two reductions (that is, from  $a_*$  to  $a_2$  and from  $a_2$  to  $a_m$ ) are as small as we want. The third reduction (from  $\bar{a} = a_m$  to  $\check{a} = \bar{a} - (1 + \bar{b})\alpha/r$ ) should be a little bigger than  $\alpha$  in order to get the desired exponentially small upper bound with the exponent  $\alpha$ . The fourth reduction (that is, from  $\check{a} = \bar{b}$  to 0) is also small, since  $\bar{b}$  can be taken as small as necessary.

This decreasing positive sequence of analyticity strip widths is associated to a similar sequence of analyticity radii. To be precise, we will construct a sequence of the form

$$b < \check{b} < \bar{b} \leq b_m < b_2 < b_*, \quad \check{b} := b + b^2 < \bar{b}\sqrt{r}.$$

The inequality  $\bar{b} \leq b_m$  does not correspond to a true reduction, but to a restriction on the size of  $\bar{b}$ . Note that we have consumed all the analyticity strip width after the last reduction, but we still keep a positive analyticity radius  $b$ .

We split the proof in the eight steps.

*Step 1: Control of the Fourier norm.* If the analytic map  $f$  satisfies the properties (i)–(iii) listed in Lemma 1, then the map  $F_2 = A + G_2 := f$  is real analytic and has the intersection property on the cylinder  $\mathbb{T} \times (-b_*, b_*)$ , can be extended to the complex domain  $D_{a_*, b_*}$ , and has the form (7). The Fourier norm  $\|G_2\|_{a_*, b_*, 2}$  may be infinite, but  $\|G_2\|_{a_2, b_2, 2} < \infty$  for any  $a_2 \in (0, a_*)$  and  $b_2 \in (0, b_*)$ .

*Step 2: Application of the averaging lemma.* Once fixed an integer  $m \geq 6$  and any  $a_m \in (0, a_2)$ , we know from Lemma 10 that there exist an analytical radius  $b_m \in (0, b_2)$  and a change of variables of the form  $\Phi_m = I + \Psi_m$  for some  $\Psi_m \in \mathcal{X}_{a_m, b_m, 1}$  such that the transformed map  $F_m = \Phi_m^{-1} \circ F_2 \circ \Phi_m$  is real analytic, has the intersection property on the cylinder  $\mathbb{T} \times (-b_m, b_m)$ , and has the form  $F_m = A + G_m$  for some  $G_m \in \mathcal{X}_{a_m, b_m, m}$ .

*Step 3: Application of Theorem 11.* Let  $r \in (0, 1)$  be the number that appears in (33). Set  $\bar{F} = A + \bar{G} = A + G_m$ ,  $\bar{a} = a_m$ , and  $\bar{d} = \|\pi^*(\bar{G})\|_{\bar{a}, b_m, m} + (1 + c_2(r))\|\pi^\bullet(\bar{G})\|_{\bar{a}, b_m, m}$ .

Let  $\bar{b} = \bar{b}(m, \bar{a}, \bar{d}, r) > 0$  be the constant stated in Theorem 11. We can assume that  $\bar{b} \leq b_m$  and the condition (33) holds, by taking a smaller  $\bar{b} > 0$  if necessary. Let  $b'_* \in (0, b_*)$  be defined by  $b'_* + (b'_*)^2 = \bar{b}\sqrt{r}$ . Fix any  $b \in (0, b'_*)$ . Set  $\check{a} = \bar{a} - (1 + \bar{b})\alpha/r$  and  $\check{b} = b + b^2 \leq b'_* + (b'_*)^2 = \bar{b}\sqrt{r}$ .

If  $\bar{d}^*$  and  $\bar{d}^\bullet$  are the norms defined in (19), then  $\bar{d}^* + (1 + c_2\bar{d}^\bullet) \leq \bar{d}$ . Hence, we can apply Theorem 11 to obtain a change of variables  $\check{\Phi} = I + \check{\Psi}$ , with  $\check{\Psi} \in \mathcal{X}_{\check{a}, \check{b}, m}$  and  $|\check{\Psi}|_{\check{a}, \check{b}, m-1} \leq c_1\bar{d}^\bullet \leq c_1\bar{d}$ , and a transformed map  $\check{F} = A + \check{G} = \check{\Phi}^{-1} \circ \bar{F} \circ \check{\Phi}$ , with  $\check{G} \in \mathcal{X}_{\check{a}, \check{b}, m}$ ,  $|\pi^\bullet(\check{G})|_{\check{a}, \check{b}, m} \leq \|\pi^\bullet(\check{G})\|_{\check{a}, \check{b}, m} \leq c_3e^{-2\pi r(\bar{a}-\check{a})/\check{b}}\bar{d}^\bullet \leq c_3e^{-2\pi\alpha(1+\bar{b})/\check{b}}\bar{d} \leq c_3e^{-2\pi\alpha/b}\bar{d}$ , and  $|\pi^*(\check{G})|_{\check{a}, \check{b}, m} \leq \|\pi^*(\check{G})\|_{\check{a}, \check{b}, m} \leq \bar{d}$ .

*Step 4: Uniform estimates on the change  $\Phi = \Phi_m \circ \check{\Phi}$ .* By construction,  $\check{\Phi} = I + \check{\Psi}$ , with  $\check{\Psi} \in \mathcal{X}_{\check{a}, \check{b}, m-1}$  and  $|\check{\Psi}|_{\check{a}, \check{b}, m-1} \leq \check{M}$ , where the constant  $\check{M} := c_1\bar{d}$  does not depend on  $b$ . Thus,

$$\check{\Phi}(x, y) = \left( x + y^{m-1}\check{\psi}_1(x, y), y + y^m\check{\psi}_2(x, y) \right)$$

for some functions  $\check{\psi}_j(x, y)$  analytic on  $D_{\check{a}, \check{b}} = D_{\bar{b}, b+b^2}$  such that  $|\check{\psi}_j|_{\bar{b}, b+b^2} \leq \check{M}$ . The Cauchy estimates imply that

$$|\check{\psi}_j(x, y)| \leq \check{M}, \quad |\partial_1\check{\psi}_j(x, y)| \leq \bar{b}^{-1}\check{M}, \quad |\partial_2\check{\psi}_j(x, y)| \leq b^{-2}\check{M},$$

for all  $(x, y) \in \mathbb{T} \times B_b$  and, in particular, for all  $(x, y) \in \mathbb{T} \times (-b, b)$ . Hence,

$$\check{\Phi}(x, y) = (x + O(y^{m-1}), y + O(y^m)), \quad \det [\check{\Phi}(x, y)] = 1 + O(y^{m-2}),$$

for all  $(x, y) \in \mathbb{T} \times (-b, b)$ , where the  $O(y^{m-2})$ ,  $O(y^{m-1})$ , and  $O(y^m)$  terms are uniform in  $b$ . We recall that  $m \geq 6$  and the change  $\Phi_m$  satisfies properties (17), so the complete change  $\Phi = \Phi_m \circ \check{\Phi}$  satisfies the properties stated in Theorem 2.

*Step 5: Exponentially small bound on the remainder  $G$ .* After all these changes of variables, we have the map  $F = A + G := \check{F}$ , with  $G := \check{G} \in \mathcal{X}_{\check{a}, \check{b}, m}$ ,  $|\pi^*(G)|_{\check{a}, \check{b}, m} \leq \bar{d}$  and  $|\pi^\bullet(G)|_{\check{a}, \check{b}, m} \leq c_3\bar{d}e^{-2\pi\alpha/b}$ . We can bound  $G^* = \pi^*(G)$  by using the bound on  $G^\bullet = \pi^\bullet(G)$  and the intersection property of  $F$  on the cylinder  $\mathbb{T} \times (-b, b)$ . We recall that if

$$G(\xi, \eta) = (\eta^m g_1(\xi, \eta), \eta^{m+1} g_2(\xi, \eta)),$$

for some  $g_1, g_2 \in \mathcal{X}_{a, b}$ , then

$$G^*(\xi, \eta) = (0, \eta^{m+1} g_2^*(\eta)), \quad G^\bullet(\xi, \eta) = (\eta^m g_1(\xi, \eta), \eta^{m+1} g_2^\bullet(\xi, \eta)),$$

where  $g_2^*(\eta)$  is the average of  $g_2(\xi, \eta)$  and  $g_2^\bullet = g_2 - g_2^*$ . Fix any  $\eta_0 \in (-b, b)$ . We know that

$$F(\mathbb{T} \times \{\eta_0\}) \cap (\mathbb{T} \times \{\eta_0\}) \neq \emptyset.$$

Therefore, there exists  $\xi_0 \in \mathbb{T}$  such that  $g_2^*(\eta_0) + g_2^\bullet(\xi_0, \eta_0) = 0$ , and so

$$|g_2^*(\eta)| \leq \sup_{\mathbb{T} \times B_b} |g_2^\bullet| \leq |G^\bullet|_{\check{a}, \check{b}, m} \leq c_3\bar{d}e^{-2\pi\alpha/b}, \quad \forall \eta \in (-b, b).$$

This implies that  $|g_j(\xi, \eta)| \leq |G|_{\check{a}, \check{b}, m} \leq |G^*|_{\check{a}, \check{b}, m} + |G^\bullet|_{\check{a}, \check{b}, m} \leq 2c_3\bar{d}e^{-2\pi\alpha/b}$  for all  $(\xi, \eta) \in \mathbb{T} \times (-b, b)$ .

*Step 6: Exponentially small bounds on some derivatives of the remainder.* We recall that

$$\max \{|g_1(\xi, \eta)|, |g_2^\bullet(\xi, \eta)|\} \leq |G^\bullet|_{\check{a}, \check{b}, m} \leq c_3 \bar{d} e^{-2\pi\alpha/b}$$

for all  $(\xi, \eta) \in D_{\check{a}, \check{b}} = D_{\bar{b}, b+b^2}$ . Thus, we get from  $\partial_1 g_2 = \partial_1 g_2^\bullet$  and the Cauchy estimates that

$$\begin{aligned} |\partial_1 g_j(\xi, \eta)| &\leq c_3 \bar{d} \bar{b}^{-1} e^{-2\pi\alpha/b}, \\ \max \{|\partial_2 g_1(\xi, \eta)|, |\partial_2 g_2^\bullet(\xi, \eta)|\} &\leq c_3 \bar{d} \bar{b}^{-2} e^{-2\pi\alpha/b}, \end{aligned}$$

for all  $(\xi, \eta) \in \mathbb{T} \times (-b, b)$ .

*Step 7: A crude bound on the derivative of  $g_2^*$ .* We recall that  $G^*(\xi, \eta) = (0, \eta^{m+1} g_2^*(\eta))$ , so

$$|g_2^*(\eta)| \leq |G^*|_{\check{a}, \check{b}, m} \leq \bar{d}, \quad \forall \eta \in B_{\check{b}} = B_{b+b^2}.$$

Therefore, the Cauchy estimates imply that  $|(g_2^*)'(\eta)| \leq b^{-2} \bar{d}$  for all  $\eta \in B_b$  and, in particular, for all  $\eta \in (-b, b)$ .

*Step 8: Computation of the constant  $K$ .* By combining the inequalities obtained in Steps 5–7, we get that  $|g_j(\xi, \eta)| \leq K e^{-2\pi\alpha/b}$  and  $|\partial_i g_j(\xi, \eta)| \leq K b^{-2}$  for all  $\mathbb{T} \times (-b, b)$ , provided

$$K = \bar{d} \max \{2c_3, c_3 + 1\}.$$

This ends the proof of Theorem 2.

## 6. Proof of Theorem 3

### 6.1. A space of matrix functions

Henceforth, let  $I_b = (-b, b) \subset \mathbb{R}$  and  $S_b = \mathbb{T} \times I_b$  with  $b > 0$ . Let  $S$  be any compact subset of  $S_b$ . Let  $\mu \in \mathbb{N}$ . Let  $\mathcal{M}_{S, \mu}$  be the set of all matrix functions  $\Gamma : S \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{R})$  of the form

$$\Gamma(\xi, \eta) = \begin{pmatrix} \eta^\mu \gamma_{11}(\xi, \eta) & \eta^{\mu-1} \gamma_{12}(\xi, \eta) \\ \eta^{\mu+1} \gamma_{12}(\xi, \eta) & \eta^\mu \gamma_{22}(\xi, \eta) \end{pmatrix},$$

for some continuous functions  $\gamma_{ij} : S \rightarrow \mathbb{R}$ . The set  $\mathcal{M}_{S, \mu}$  is a Banach space with the norm

$$\|\Gamma\|_{S, \mu} = \max \{|\gamma_{ij}(\xi, \eta)| : (\xi, \eta) \in S, 1 \leq i, j \leq 2\}.$$

**Lemma 18.** *Let  $S \subset S_b$ ,  $\Gamma \in \mathcal{M}_{S, \mu}$ ,  $\Delta \in \mathcal{M}_{S, \nu}$ , and  $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ . Let  $k \in \mathbb{N}$  and  $j \in \mathbb{Z}$ .*

*Let  $f : S \rightarrow S$  be a map of the form (12) with  $|g_2(\xi, \eta)| \leq K_0$  for all  $(\xi, \eta) \in S$ . Then:*

- (i)  $\Gamma \Delta \in \mathcal{M}_{S, \mu+\nu}$  and  $\|\Gamma \Delta\|_{S, \mu+\nu} \leq 2\|\Gamma\|_{S, \mu} \|\Delta\|_{S, \nu}$ .
- (ii)  $A^k \Gamma \in \mathcal{M}_{S, \mu}$  and  $\|A^k \Gamma\|_{S, \mu} \leq (1 + bk)\|\Gamma\|_{S, \mu}$ .
- (iii)  $\Gamma A^k \in \mathcal{M}_{S, \mu}$  and  $\|\Gamma A^k\|_{S, \mu} \leq (1 + bk)\|\Gamma\|_{S, \mu}$ .
- (iv)  $\Gamma \circ f^j \in \mathcal{M}_{S, \mu}$  and  $\|\Gamma \circ f^j\|_{S, \mu} \leq (1 + K_0 b^m)^{(\mu+1)|j|} \|\Gamma\|_{S, \mu}$ .

*Proof.* It is a straightforward computation. □

## 6.2. A technical lemma

Let  $f$  be an analytic map of the form (12). Let  $p$  and  $q$  be two relatively prime integers. There exist two curves  $R = \text{graph } \zeta$  and  $\hat{R} = \text{graph } \hat{\zeta}$  and two RICs  $R_{\pm} = \text{graph } \zeta_{\pm}$  with Diophantine rotation numbers  $\omega_- < p/q < \omega_+$ , all four contained in a small neighbourhood of  $\mathbb{T} \times \{p/q\}$ , such that  $f^q$  projects  $R$  onto  $\hat{R}$  along the vertical direction and  $R$  and  $\hat{R}$  are contained in the strip of the cylinder enclosed by the RICs  $R_{\pm}$ . Following Birkhoff [4, Section VI] and Arnold [2, Section 20], all  $(p, q)$ -periodic points of  $f$  are contained in  $R \cap \hat{R}$ . Besides, we will see later on that the (geometric) area enclosed between  $R$  and  $\hat{R}$  is an upper bound of the quantities  $\Delta^{(p,q)}$ . These are the reasons for the study of  $R$  and  $\hat{R}$ .

Let us prove that these four curves exist for big enough periods  $q$ . In this case, “big enough” only depends on the size of the nonintegrable terms of  $f$ , the size of the neighbourhood of  $\mathbb{T} \times \{p/q\}$ , the exponent  $m$ , and the winding number  $p$ . On the contrary, it *does not depend* on the particular map at hand. Therefore, every time that we ask  $q$  to be “big enough” along the proof of the following lemma, it only depends on the quantities  $K_0 > 0$ ,  $c > 1$ ,  $m \geq 4$ ,  $p \in \mathbb{Z} \setminus \{0\}$ , and  $q'_* \in \mathbb{N}$  fixed at the first line of the next statement.

**Lemma 19.** *Let  $K_0 > 0$ ,  $c > 1$ ,  $m \geq 4$ , and  $p, q'_* \in \mathbb{N}$ . Let  $q \geq q'_*$  be an integer relatively prime with  $p$ . Set  $b = c^2 p/q$ . Let  $f : S_b \rightarrow \mathbb{T} \times \mathbb{R}$  be an analytic map of the form (12) such that  $|g_j(\xi, \eta)| \leq K_0$  and  $|\partial_i g_j(\xi, \eta)| \leq K_0 b^{-2}$  for all  $(\xi, \eta) \in S_b$ . Let  $(\xi_q, \eta_q) = f^q(\xi, \eta)$ . Let  $I = (p/c^2 q, c^2 p/q)$ ,  $I_- = (p/c^2 q, p/cq)$ , and  $I_+ = (cp/q, c^2 p/q)$ . There exists  $q''_* = q''_*(K_0, c, m, p, q'_*) \geq q'_*$  such that, if  $q \geq q''_*$ , the following properties hold:*

(i) *The map  $f$  has two RICs  $R_{\pm} \subset \mathbb{T} \times I_{\pm} \subset S_b$  whose internal dynamics is conjugated to a rigid rotation of angles  $\omega_{\pm} \in I_{\pm}$ , respectively;*

(ii) *If  $S$  is the compact subset of  $S_b$  enclosed by  $R_-$  and  $R_+$ , then*

$$\frac{\partial \xi_q}{\partial \eta}(\xi, \eta) > 0, \quad \forall (\xi, \eta) \in S; \quad (34)$$

(iii) *There exist two unique analytic functions  $\zeta : \mathbb{T} \rightarrow I$  and  $\hat{\zeta} : \mathbb{T} \rightarrow I$  such that*

$$f^q(\xi, \zeta(\xi)) = (\xi, \hat{\zeta}(\xi)), \quad \forall \xi \in \mathbb{T}, \quad (35)$$

*and all the  $(p, q)$ -periodic points of the restriction  $f|_S$  are contained in  $\text{graph } \zeta$ .*

*The same statement holds if  $p$  is a negative integer,  $b = c^2 |p|/q$ ,  $I = (c^2 p/q, p/c^2 q)$ ,  $I_- = (c^2 p/q, cp/q)$ , and  $I_+ = (p/cq, p/c^2 q)$ .*

*Proof.* Let us assume  $p > 0$ . The case  $p < 0$  is analogous.

First, the existence of the RICs  $R_-$  and  $R_+$  follows from some quantitative estimates in KAM theory established by Lazutkin [21, Theorem 2]. To be precise, Lazutkin proved that there exists  $b'_* = b'_*(K_0) > 0$  such that if  $\omega \in (-b'_*, b'_*)$  satisfies the Diophantine condition

$$2|\omega - i/j| \geq |i|j^{-4} \quad (36)$$

for all integers  $j \geq 1$  and  $i$ , then  $f$  has a RIC  $R = \{\eta = \omega + O(\omega^m)\}$  whose internal dynamics is  $C^l$ -conjugated to a rigid rotation of angle  $\omega$ , for a suitable  $l \geq 1$ . The conjugation is  $O(1/q^{m-1})$ -close to the identity. Item (i) follows directly from this estimate, because there

exist some real numbers  $\omega_+ \in (c^{4/3}p/q, c^{5/3}p/q) \subset I_+$  and  $\omega_- \in (p/c^{5/3}q, p/c^{4/3}q) \subset I_-$  satisfying the Diophantine condition (36), provided  $q$  is big enough.

Second, let us check that the power map  $f^q$  satisfies (34). The compact subset  $S \subset S_b$  is invariant by  $f$ , because it is delimited by RICs. Thus, all powers  $(\xi_j, \eta_j) = f^j(\xi, \eta)$  are well defined on  $S$ . We write  $Df(\xi, \eta) = A + \Gamma(\xi, \eta)$ , where  $A$  was introduced in Lemma 18. Next, we compute the differential of the power map:

$$Df^q = (A + \Gamma_q) \cdots (A + \Gamma_1) = A^q + \Delta_1 + \cdots + \Delta_q, \quad (37)$$

where  $\Gamma_j = \Gamma \circ f^j$  and  $\Delta_l$  is the sum of all the products of the form  $A^{k_1}\Gamma_{j_1} \cdots A^{k_l}\Gamma_{j_l}A^{k_{l+1}}$ , with  $k_i \geq 0$ ,  $q \geq j_1 > j_2 > \cdots > j_l \geq 1$ , and  $q = l + \sum_{i=1}^{l+1} k_i$ . These products are elements of  $\mathcal{M}_{S,ml}$ , because  $\Gamma \in \mathcal{M}_{S,m}$ . Indeed, if  $C = \|\Gamma\|_{S,m}$ , then

$$\begin{aligned} \|A^{k_1}\Gamma_{j_1} \cdots A^{k_l}\Gamma_{j_l}A^{k_{l+1}}\|_{S,ml} &\leq 2^{l-1} \|A^{k_1}\Gamma_{j_1}\|_{S,m} \cdots \|A^{k_{l-1}}\Gamma_{j_{l-1}}\|_{S,m} \|A^{k_l}\Gamma_{j_l}A^{k_{l+1}}\|_{S,m} \\ &\leq 2^{l-1} C^l \prod_{i=1}^{l+1} (1 + bk_i) \prod_{i=1}^l (1 + K_0 b^m)^{(m+1)j_i} \\ &\leq 2^{l-1} C^l \exp\left(\sum_{i=1}^{l+1} bk_i + \sum_{i=1}^l (m+1)K_0 b^m j_i\right) \\ &\leq 2^{l-1} C^l e^{c^2 p + (m+1)K_0 c^{2m} p^m} = C'(2C)^l, \end{aligned}$$

where we have used Lemma 18, inequality  $1 + x \leq e^x$  for  $x \geq 0$ ,  $\sum_{i=1}^{l+1} k_i \leq q$ ,  $l \leq q$ ,  $j_i \leq q$ ,  $b = c^2 p/q$ , and  $m \geq 4$ . We have also defined  $C' = e^{c^2 p + (m+1)K_0 c^{2m} p^m} / 2$ .

The matrix  $\Delta_l$  is the sum of the products with precisely  $l$  factors  $\Gamma_j$ . This shows that there are  $\binom{q}{l}$  terms inside  $\Delta_l$ . Therefore,  $\Delta_l \in \mathcal{M}_{S,ml}$  and

$$\|\Delta_l\|_{S,ml} \leq \binom{q}{l} \|A^{k_1}\Gamma_{j_1} \cdots A^{k_l}\Gamma_{j_l}A^{k_{l+1}}\|_{S,ml} \leq C'(2C)^l. \quad (38)$$

The element of the first row and second column of  $A^q$  is equal to  $q$ , so

$$\begin{aligned} \left| \frac{\partial \xi_q}{\partial \eta}(\xi, \eta) - q \right| &\leq C' \sum_{l=1}^q (2C)^l b^{ml-1} \leq \frac{C'}{b} \sum_{l=1}^q (2K_0 q b^{m-2})^l \\ &\leq 4C' K_0 q b^{m-3} \leq 4C' K_0 c^{2(m-3)} p^{m-3}, \end{aligned}$$

for all  $(\xi, \eta) \in S \subset S_b$ , which implies the twist condition (34) provided that  $q$  is big enough. Here, we have used relation (37), bound (38),  $b = c^2 p/q$ , and  $m \geq 4$ . We have also used that  $C \leq K_0 b^{-2}$  and  $2K_0 q b^{m-2} \leq 1/2$ , provided  $q$  is big enough.

Third, we establish the existence of the functions  $\zeta, \hat{\zeta} : \mathbb{T} \rightarrow I$ . We know from Lazutkin [21] that  $R_{\pm} = \text{graph } \zeta_{\pm}$  for some differentiable functions  $\zeta_{\pm} : \mathbb{T} \rightarrow I_{\pm}$ . We work with the lifts  $F, \Xi_q$ , and  $Z_{\pm}$  of the objects  $f, \xi_q$ , and  $\zeta_{\pm}$ . The RICs are invariant, so  $F^q(\xi, Z_{\pm}(\xi)) = (\Xi_{\pm}(\xi), Z_{\pm}(\Xi_{\pm}(\xi)))$  for some differentiable functions  $\Xi_{\pm} : \mathbb{R} \rightarrow \mathbb{R}$ . If we prove that there exist two unique analytic 1-periodic functions  $Z, \hat{Z} : \mathbb{R} \rightarrow I$  such that

$$F^q(\xi, Z(\xi)) = (\xi + p, \hat{Z}(\xi + p)), \quad \forall \xi \in \mathbb{R}, \quad (39)$$

then item (iii) follows. Since the dynamics of  $F^q$  on  $R_{\pm}$  is  $C^l$ -conjugated to a rigid rotation of angle  $q\omega_+$  through a  $O(1/q^{m-1})$ -close to the identity conjugation,

$$\Xi_+(\xi) = \xi + q\omega_+ + O(1/q^{m-1}) \geq \xi + cp + O(1/q^{m-1}) > \xi + p$$



provided that  $q$  is big enough. Analogously, we obtain  $\Xi_-(\xi) < \xi + p$ . That is,

$$\Xi_q(\xi, Z_-(\xi)) = \Xi_-(\xi) < \xi + p < \Xi_+(\xi) = \Xi_q(\xi, Z_+(\xi)), \quad \forall \xi \in \mathbb{R}.$$

Since  $\Xi_q(\xi, \eta)$  is analytic and strictly increasing for  $\eta \in (Z_-(\xi), Z_+(\xi)) \subset I$ , we deduce that there exists a unique function  $Z : \mathbb{R} \rightarrow I$  such that  $\Xi_q(\xi, Z(\xi)) = \xi + p$ .

The function  $S \ni (\xi, \eta) \mapsto G(\xi, \eta) := \Xi_q(\xi, \eta) - \xi - p$  is analytic and  $\frac{\partial G}{\partial \eta}(\xi, \eta) > 0$ , so  $Z$  is analytic by the Implicit Function Theorem. The 1-periodicity of  $Z$  follows from the uniqueness and the property  $F^q(\xi + 1, \eta) = F^q(\xi, \eta) + (1, 0)$ . Function  $\hat{Z} : \mathbb{R} \rightarrow I$  is defined by means of relation (39). Finally, functions  $\zeta, \hat{\zeta} : \mathbb{T} \rightarrow I$  are the projections of  $Z, \hat{Z} : \mathbb{R} \rightarrow I$ .  $\square$

### 6.3. Proof of Theorem 3: Case $(m, n) = (0, 1)$

If  $(m, n) = (0, 1)$ , by hypothesis, the map  $g : \mathbb{T} \times I \rightarrow \mathbb{T} \times I, (s, r) \mapsto (s_1, r_1)$ , is an analytic exact twist map with a  $(a_*, b_*)$ -analytic  $(0, 1)$ -resonant RIC, such that  $\varrho_- \leq 0 \leq \varrho_+$ . The map  $f = g^n = g$  satisfies the properties (i)–(iii) listed in Lemma 1 in some suitable coordinates  $(x, y)$ . Let  $(s, r) = \tilde{\Phi}(x, y)$  be the associated change of variables. Let  $\tilde{f} = \tilde{\Phi}^{-1} \circ f \circ \tilde{\Phi}$  be the new map defined in the domain (8). Note that the  $(a_*, b_*)$ -analytic  $(0, 1)$ -resonant RIC is  $C \equiv \{y = 0\}$  in the  $(x, y)$  coordinates.

Let  $p$  be an integer such that  $1 \leq |p| \leq L$ . Let  $c \in (1, 2)$  such that  $\alpha < c^2\alpha < a_*$ . We take  $c^2\alpha$  as the  $\alpha$  appearing in Theorem 2,  $m = 4$ , and  $b = c^2|p|/q$ , provided that  $q$  is relatively prime with  $p$  and is large enough so that  $c^2|p|/q < b'_* = b'_*(\alpha)$ . That is,  $q > q'_* := c^2|p|/b'_*$ .

Hence, there exist  $K_0, K_1 > 0$ , both independent of  $q$ , and a change of coordinates  $(x, y) = \Phi(\xi, \eta)$  such that  $\tilde{f} = \Phi^{-1} \circ \tilde{f} \circ \Phi : S_b \rightarrow \mathbb{T} \times \mathbb{R}$  is an analytic map of the form (12) such that  $|g_j(\xi, \eta)| \leq K_0 e^{-2\pi c^2\alpha/b} = K_0 e^{-2\pi\alpha q/|p|} \leq K_0$ ,  $|\partial_i g_j(\xi, \eta)| \leq K_0 b^{-2}$ , and  $\sup\{|\det[D\Phi(\xi, \eta)]|\} \leq K_1$  for all  $(\xi, \eta) \in S_b$ .

The map  $\tilde{f} : S_b \rightarrow \mathbb{T} \times \mathbb{R}$  satisfies the hypotheses of Lemma 19 for any  $q \geq q'_*$ . Let  $q_*$  be the maximum value of  $q''_*$  among the integers  $0 < |p| \leq L$ . Let  $R_\pm$  be the RICs with rotation numbers  $\omega_\pm$  given in Lemma 19. Let  $S$  be the compact subset of  $S_b$  enclosed by  $R_-$  and  $R_+$ . Since  $f$  is globally twist and  $\varrho_- < \omega_- < p/q < \omega_+ < \varrho_+$ , all the Birkhoff  $(p, q)$ -periodic orbits of  $f$  are contained in  $S$ . By Lemma 19, any  $(p, q)$ -periodic orbit in  $S$  lies on  $R = \text{graph } \zeta$ . Let  $\Omega \subset S$  be the domain enclosed by the curves  $R = \text{graph } \zeta$  and  $\hat{R} = \text{graph } \hat{\zeta}$ . Let  $B = (\tilde{\Phi} \circ \Phi)(\Omega)$ . Let  $K_2$  be the supremum of  $|\det[D\tilde{\Phi}]|$  in the compact set  $\mathbb{T} \times [-b'_*, b'_*]$ . Let  $K = 4K_0K_1K_2L(b'_*)^3$ . Then, following the arguments contained in Subsection 2.2 about the difference of periodic actions, we get that

$$\begin{aligned} \Delta^{(p,q)} &\leq \text{Area}[B] \leq K_1K_2 \text{Area}[\Omega] = K_1K_2 \int_{\mathbb{T}} \left| \hat{\zeta}(\xi) - \zeta(\xi) \right| d\xi \\ &\leq K_1K_2qb^4 K_0 e^{-2\pi\alpha q/|p|} \leq K e^{-2\pi\alpha q/|p|}, \end{aligned} \quad (40)$$

for all relatively prime integers  $p$  and  $q$  with  $1 \leq |p| \leq L$  and  $q \geq q_*$ . We have used expression (12),  $b = c^2|p|/q \leq b'_*$ ,  $c^2 < 4$ , the bounds on the nonintegrable terms  $g_j(\xi, \eta)$ , and the bounds on the Jacobians of the changes of variables  $\Phi$  and  $\tilde{\Phi}$ .

This ends the proof of Theorem 3 when  $(m, n) = (0, 1)$  and  $q \geq q_*$ . By redefining the constant  $K$ , the same bound holds for all  $q \geq 1$ .

#### 6.4. Proof of Theorem 3: General case

We reduce the general case to the previous one. We split the argument in four steps.

*Step 1: About the rational rotation numbers.* If  $C$  is a  $(m, n)$ -resonant RIC and  $(s, r)$  is a  $(p, q)$ -periodic point of  $g$ , then  $C$  is a  $(m, 1)$ -resonant RIC and  $(s, r)$  is a  $(p', q')$ -periodic point of the power map  $f = g^n$ , where

$$p' = \frac{np}{\gcd(n, q)}, \quad q' = \frac{q}{\gcd(n, q)}.$$

By taking the suitable lift  $F$  of  $f$ , we can assume that  $C$  is a  $(0, 1)$ -resonant RIC and  $(s, r)$  is a  $(p'', q'')$ -periodic point of  $f$ , with  $p''/q'' = p'/q' - m$ . That is,

$$p'' = p' - mq' = \frac{np - mq}{\gcd(n, q)}, \quad q'' = q' = \frac{q}{\gcd(n, q)}. \quad (41)$$

If  $p$  and  $q$  are relatively prime integers such that  $1 \leq |np - mq| \leq L$  and  $q \geq q_*$ ,  $p''$  and  $q''$  are relatively prime integers such that  $|p''| \leq L/\gcd(n, q) \leq L$  and  $q'' \geq q_*/\gcd(n, q) \geq q_*/n$ .

*Step 2: About the Lagrangians.* Let  $G$  and  $F = G^n$  be the lifts of  $g$  and  $f = g^n$  we are dealing with. If  $G^*\lambda - \lambda = dh$ , then

$$F^*\lambda - \lambda = \sum_{j=0}^{n-1} [(G^{j+1})^*\lambda - (G^j)^*\lambda] = \sum_{j=0}^{n-1} d(h \circ G^j) = d\left(\sum_{j=0}^{n-1} h \circ G^j\right),$$

so  $\ell(s_0, s_n) := h(s_0, s_1) + h(s_1, s_2) + \dots + h(s_{n-1}, s_n)$  is a Lagrangian of  $f$ . This Lagrangian is well defined in a neighbourhood of the resonant RIC  $C$ , because  $f$  is twist on  $C$ .

*Step 3: About the periodic actions.* Let  $O$  be the  $(p, q)$ -periodic orbit of  $g$  through the point  $(s, r)$ , being  $W^{(p, q)}[O]$  its  $(p, q)$ -periodic action. Let  $O''$  be the  $(p'', q'')$ -periodic orbit of  $f$  through the same point, being  $W^{(p'', q'')}[O'']$  its  $(p'', q'')$ -periodic action. We deduce from the previous steps and a straightforward computation that

$$W^{(p'', q'')}[O''] = \frac{n}{\gcd(n, q)} W^{(p, q)}[O]. \quad (42)$$

*Step 4: Final bound.* The result follows directly from the bound (40) taking into account relations (41) and (42). We just note that

$$e^{-2\pi\alpha q''/|p''|} = \exp\left(-\frac{2\pi\alpha q}{|np - mq|}\right).$$

This ends the proof of Theorem 3.

## Acknowledgments

The authors were supported in part by CUR-DIUE Grant 2014SGR504 (Catalonia) and MINECO-FEDER Grant MTM2012-31714 (Spain). Some of the results of this paper were obtained while A T was in the University of Warwick as a visiting researcher. Thanks to Vassily Gelfreich, Robert MacKay, Laurent Niederman, Sônia Pinto-de-Carvalho, and Tere Seara for useful and stimulating conversations. We are also indebted to the referee for several detailed suggestions that encouraged us to improve the paper.

## References

- [1] K. G. Andersson and R. B. Melrose. The propagation of singularities along gliding rays. *Invent. Math.*, 41(3):197–232, 1977.
- [2] V. I. Arnol'd and A. Avez. *Ergodic Problems of Classical Mechanics*. Translated from the French by A. Avez. W. A. Benjamin, Inc., New York-Amsterdam, 1968.
- [3] Y. Baryshnikov and V. Zharnitsky. Sub-Riemannian geometry and periodic orbits in classical billiards. *Math. Res. Lett.*, 13(4):587–598, 2006.
- [4] G. D. Birkhoff. *Dynamical Systems*. With an addendum by Jurgen Moser. American Mathematical Society Colloquium Publications, Vol. IX. American Mathematical Society, Providence, R.I., 1966.
- [5] P. Boyland. Dual billiards, twist maps and impact oscillators. *Nonlinearity*, 9(6):1411, 1996.
- [6] Y. Colin de Verdière. Sur les longueurs des trajectoires périodiques d'un billard. In *South Rhone seminar on geometry, III (Lyon, 1983)*, Travaux en Cours, pages 122–139. Hermann, Paris, 1984.
- [7] L. Coutinho. *Bilhares em superfícies de curvatura constante*. PhD thesis, Universidade Federal de Minas Gerais, 2014.
- [8] V. Cyr. A number theoretic question arising in the geometry of plane curves and in billiard dynamics. *Proc. Amer. Math. Soc.*, 140(9):3035–3040, 2012.
- [9] M. Day. Polygons circumscribed about closed convex curves. *Trans. Amer. Math. Soc.*, 62:315–319, 1947.
- [10] A. Delshams and R. de la Llave. KAM theory and a partial justification of Greene's criterion for nontwist maps. *SIAM J. Math. Anal.*, 31(6):1235–1269, 2000.
- [11] R. Douady. *Applications du théorème des tores invariantes*. Thèse, Université Paris VII, 1982.
- [12] E. Fontich and C. Simó. The splitting of separatrices for analytic diffeomorphisms. *Ergodic Theory Dynam. Systems*, 10(2):295–318, 1990.
- [13] J. M. Greene. A method for determining a stochastic transition. *J. Math. Phys.*, 20:1183–1201, 1979.
- [14] E. Gutkin. Capillary floating and the billiard ball problem. *Journal of Mathematical Fluid Mechanics*, 14(2):363–382, 2012.
- [15] E. Gutkin and A. Katok. Caustics for inner and outer billiards. *Comm. Math. Phys.*, 173(1):101–133, 1995.
- [16] N. Innami. Convex curves whose points are vertices of billiard triangles. *Kodai Mathematical Journal*, 11(1):17–24, 1988.
- [17] M. Kac. Can one hear the shape of a drum? *Amer. Math. Monthly*, 73(4, part II):1–23, 1966.
- [18] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.
- [19] O. Knill. On nonconvex caustics of convex billiards. *Elem. Math.*, 53(3):89–106, 1998.
- [20] V.V. Kozlov and D.V. Treschev. *Billiards: A Genetic Introduction to the Dynamics of Systems with Impacts*, volume 89 of *Translations of Mathematical Monographs*. Amer. Math. Soc., Providence, RI, 1991.
- [21] V. F. Lazutkin. Existence of caustics for the billiard problem in a convex domain. *Izv. Akad. Nauk SSSR Ser. Mat.*, 37:186–216, 1973.
- [22] R. S. MacKay. Greene's residue criterion. *Nonlinearity*, 5(1):161–187, 1992.

- [23] R. S. MacKay, J. D. Meiss, and I. C. Percival. Transport in Hamiltonian systems. *Phys. D*, 13(1-2):55–81, 1984.
- [24] P. Martín, R. Ramírez-Ros, and A. Tamarit-Sariol. Exponentially small asymptotic formulas for the length spectrum in some billiard tables. To appear in *Experimental Mathematics*.
- [25] S. Marvizi and R. Melrose. Spectral invariants of convex planar regions. *J. Differential Geom.*, 17(3):475–502, 1982.
- [26] J. N. Mather. A criterion for the nonexistence of invariant circles. *Inst. Hautes Études Sci. Publ. Math.*, (63):153–204, 1986.
- [27] J. D. Meiss. Symplectic maps, variational principles, and transport. *Rev. Modern Phys.*, 64(3):795–848, 1992.
- [28] J. Moser. Is the solar system stable? *Math. Intelligencer*, 1(2):65–71, 1978/79.
- [29] A. I. Neishtadt. Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions. *J. Appl. Math. Mech.*, 45(6):1016–1025, 1981.
- [30] A. I. Neishtadt. The separation of motions in systems with rapidly rotating phase. *Prikl. Mat. Mekh.*, 48(2):197–204, 1984.
- [31] S. Pinto-de-Carvalho and R. Ramírez-Ros. Non-persistence of resonant caustics in perturbed elliptic billiards. *Ergodic Theory Dynam. Systems*, 33(6):1876–1890, 2013.
- [32] J. Pöschel. The concept of integrability on Cantor sets for Hamiltonian systems. *Celestial Mech.*, 28(1-2):133–139, 1982.
- [33] R. Ramírez-Ros. Break-up of resonant invariant curves in billiards and dual billiards associated to perturbed circular tables. *Phys. D*, 214(1):78–87, 2006.
- [34] C. L. Siegel and J. K. Moser. *Lectures on Celestial Mechanics*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Translated from the German by C. I. Kalme, Reprint of the 1971 translation.
- [35] A. Sorrentino. Computing Mather’s  $\beta$ -function for Birkhoff billiards. *Discrete Contin. Dynam. Systems*, 35(10):5055–5082, 2015.
- [36] S. Tabachnikov. Billiards. *Panor. Synth.*, (1):vi+142, 1995.
- [37] S. Tabachnikov. On the dual billiard problem. *Adv. Math.*, 115(2):221–249, 1995.
- [38] S. Tabachnikov. Dual billiards in the hyperbolic plane. *Nonlinearity*, 15(4):1051–1072, 2002.